AN ANALYSIS OF THE COHERENCE-BASED GENEALOGICAL METHOD USING PHYLOGENETICS

by

ANDREW CHARLES EDMONDSON

A thesis submitted to the
University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of Theology and Religion
School of Philosophy, Theology and Religion
College of Arts and Law
University of Birmingham
September 2018

Abstract

The *Novum Testamentum Graecum: Editio Critica Maior* is the first major critical edition of the New Testament for a century, aiming to document the New Testament's textual history through its first millennium. To date, two of the six volumes have been published. As part of this project the Institut für neutestamentliche Textforschung in Münster has developed the Coherence-Based Genealogical Method (CBGM), a computer-aided method designed to handle complete sets of textual evidence and to identify their initial text and textual history. The CBGM is widely held to be difficult to understand and its results are treated with scepticism.

Phylogenetics is the study of relationships between groups of organisms and their evolutionary history. Phylogenetics and the CBGM (and wider textual criticism) have many commonalities. This thesis provides a thorough examination of the CBGM using phylogenetics.

Part One documents the literature surrounding the CBGM and includes a worked example of the process. Part Two explores the ECM data for John's Gospel and identifies appropriate methods for applying phylogenetics to it. Part Three compares the results of phylogenetics and the CBGM. It concludes that the CBGM is producing valid results from the data, but could be improved in a number of ways.

Acknowledgements

I am indebted to many people for helping and supporting me during my studies. I have had the privilege of being supervised by two directors of the Institute for Textual Scholarship and Electronic Editing (ITSEE). First I would like to thank my main supervisor Professor David Parker for his immeasurable help and guidance during the last six years, even into his stepped retirement. He has found a perfect balance of giving me autonomy yet not letting me go off track during my research. My thesis is built on David's work editing John's Gospel for the *Novum Testamentum Graecum: Editio Critica Maior*, and could not exist without it. Following David's partial retirement Professor Hugh Houghton took over as director of ITSEE and has also been my co-supervisor for the second half of my research. I am very grateful to Hugh for his many insightful comments on my thesis, and for all the help he has offered me in my time at Birmingham. I would also like to thank Dr Catherine Smith for her help in ITSEE, and for introducing me to BlueBEAR (the University of Birmingham's supercomputer) which radically reduced the time required for my experiments.

I have received expert help from many places during my research. I would particularly like to mention Dr Klaus Wachtel from the Institut für Neutestamentliche Textforschung in Münster. Klaus has discussed the Coherence-Based Genealogical Method (CBGM) many times with me, and has read drafts of my work and provided many valuable comments and corrections. This thesis explores the CBGM partially by making use of phylogenetics, and I am very grateful to Professor Jean-Baptiste Cazier (director of the University of Birmingham's Centre for Computational Biology) and Dr Albert Nobre de Menezes for validating the method, interpreting the results and making many valuable suggestions. Albert introduced me to Fluxus Engineering's *Network* program, which has proved itself to be excellent for analysing my data.

Many thanks go to Rev Dr Lee Rayner, then leader of Malvern Evangelical Church, for introducing me to New Testament Greek and encouraging me to go to Bible College to study Theology where I was taught Greek by Dan Yarnell. Dan’s enthusiasm for the Bible, and its Greek text in particular, continues to inspire me. His comment, following a lesson on textual criticism, to the effect that he could see me spending the rest of my career on that topic may well be the reason that I am where I am today.

There are many other people I should thank, but space will not permit. I will nevertheless mention Dr John Owen and Carol Sandys for supporting my research during my time in the Advanced Research Computing team. Thanks also to Jonathan Bell, leader of Churchcentral in Birmingham, for providing opportunities for me to teach people in our church network about the things I have learned during my research.

I could not have completed this research without the support of the scholarship awarded to me by the College of Arts and Law for the second half of my studies.

I’m very grateful to my parents and parents-in-law, who have helped and supported me many, many times during this journey. Finally, and most importantly, I want to thank my wife Karen, who has sacrificed a lot over the last six years. We started this task as a couple, and end it as a family of four.

Table of Contents

PART ONE

THE COHERENCE-BASED GENEALOGICAL METHOD

PART TWO

PHYLOGENETICS

PART THREE

ANALYSING THE CBGM

Introduction

This thesis sets out to explore, analyse and improve the Coherence-Based Genealogical Method (CBGM). This task is approached directly by studying and using the CBGM, and indirectly by exploring the underlying data using phylogenetics. The core idea is that phylogenetics is conceptually similar to the CBGM – they both provide a computer-aided analysis of the evolution of items: texts (CBGM) or species (in Biology) – yet they are independent methods and therefore the results of one may be able to validate the results of the other.

The thesis is arranged in three parts. Part One provides a detailed introduction to the CBGM. Chapter 1 introduces the CBGM, its history and its relationship to the *Novum Testamentum Graecum: Editio Critica Maior* (ECM). Chapter 1 also provides an overview of its primary literature, i.e. that produced by the Institut für neutestamentliche Textforschung (INTF) in Münster, where the CBGM was created and continues to be developed. Chapter 2 sets out a summary of how the CBGM has been received by the wider scholarly community. Chapter 3 is a thorough, step-by-step description of applying the CBGM to John 6:21-24 based on all majuscules and papyri of John's Gospel included in the ECM. Beginning with the task of collating the witnesses and ending with the creation of the global stemma, it allows a transparent and complete understanding of how to use the CBGM for such verses. Chapter 3 also includes a brief comparison of the results of phylogenetic software and the CBGM on this small dataset.

Part Two explores phylogenetics as applied to manuscript traditions, and evaluates various different methods and software. Chapter 4 is an exploration of transcriptions of John from the International Greek New Testament Project (IGNTP), visualising the data and applying phylogenetics to it. Two phylogenetic software packages are found to be highly valuable, and are applied in detail to the manuscript tradition of the Gospel of John: *MrBayes* and Fluxus Engineering's *Network*.[1] Chapter 5 applies *MrBayes* to the transcriptions of all 1,559 extant Greek manuscripts of John 18, identifying some well known groups and families, and finding some hitherto unknown groups. Chapter 6 is a brief excursus seeking to optimise the use of *MrBayes* for the IGNTP data for John 2, resulting in a reduction of the time required from an average of eleven and a half hours to thirty eight minutes. This is *eighteen times faster*, considering the averages.

Part Three contains the analysis and suggested improvements for the CBGM. Chapter 7 argues that using and showing only the rank for connectivity is insufficient, and it would be better to include a minimum coherence percentage. This is explored by changing the CBGM's output in various ways. Chapter 8 highlights the importance of the strength of textual flow, which is not currently included in standard textual flow diagrams, and discusses alternative display methods and algorithms. Chapter 9 contains a number of shorter criticisms of the CBGM, for example suggesting an alternative and "more objective" algorithm for calculating pre-genealogical coherence. Chapter 10 sets out a detailed comparison of the results of the CBGM and phylogenetic software using the full data from David Parker's work on the ECM of John's Gospel. It concludes that the CBGM and phylogenetics both identify similar groups and structures in the data, using independent methods: therefore the CBGM's results may be treated with confidence.

Much of the research described in this thesis uses my own implementation of the CBGM, written in Python. This software can be found on my GitHub page, https://github.com/edmondac/, or by following the DOIs provided in the references

---

[1] For *MrBayes* see http://mrbayes.sourceforge.net; Huelsenbeck and Ronquist, 'MRBAYES'; Ronquist and Huelsenbeck, 'MrBayes 3'. For *Network* see http://www.fluxus-engineering.com; Bandelt, Forster, and Rohl, 'Median-Joining Networks for Inferring Intraspecific Phylogenies'.

throughout. All software has bugs, and it would be foolish to claim that this software was any different. However, it has been thoroughly tested and I believe that there are no bugs that could significantly affect the results. Each chapter or section will note the version(s) of the software used, to aid reproducibility.

PART ONE

THE COHERENCE-BASED GENEALOGICAL METHOD

Chapter 1: The CBGM

Introduction

In this chapter I will introduce the Coherence-Based Genealogical Method (CBGM), its history, its relationship to the *Novum Testamentum Graecum: Editio Critica Maior* (ECM) and provide an overview of its primary literature. While publications about the CBGM continue to appear regularly, the need to impose a limit for the purposes of this thesis means that it only considers literature published before 1 January 2018.

History of the CBGM and the ECM

> *"The ECM represents a new level of scientific research on the text of the Greek New Testament and offers a text newly established on this basis."*[1]

In 1997 the Institut für neutestamentliche Textforschung (INTF) in Münster published the first instalment (James) of the first volume (IV: Catholic Epistles) of a major new project: *Novum Testamentum Graecum: Editio Critica Maior,* more commonly known as the ECM. Building on years of previous research at the INTF, the ongoing goal of the ECM is to document the New Testament's textual history through its first millennium. In 2000 the INTF published the second instalment (the Letters of Peter), in 2003 the third (the First Letter of John) and in 2005 the fourth (the Second and Third Letter of John, the Letter of Jude). They continued working on the Catholic Epistles, and in 2013 published the second edition of volume IV. In 2017 they published volume III: Acts, after which they will turn to Mark.[2] The INTF has entered into agreements with other institutions to complete the ECM: the

---

[1] Holger Strutwolf, Foreword to Aland et al., *NA28* (unnumbered page).

[2] Note that the volume numbers have changed from those anticipated in volume IV, where Acts would have been volume II. They are now defined as follows: I. Synoptic Gospels; II. Gospel of John; III. Acts; IV. Catholic Letters; V. Pauline Letters; VI. John's Apocalypse. See Strutwolf et al., *ECM III (Acts)*, pt 1, 18*.

International Greek New Testament Project (IGNTP) is editing the Gospel of John and will then move on to Paul's letters.[3] The Apocalypse of John is currently being edited by the Institut für Septuaginta- und biblische Textforschung (ISBTF) in Wuppertal[4].

A major claim of the ECM is that it "provides the full range of resources necessary for scholarly research in establishing the text and reconstructing the history of the New Testament text during its first thousand years."[5] However, the ECM text is not without its critics, or rather the ECM method and its resulting text has drawn criticism even from some who would traditionally stand by the Nestle-Aland text. The core of this new method is the Coherence-Based Genealogical Method (CBGM).

The CBGM was not used to create the first instalment of the ECM in 1997.[6] The second instalment in 2000 contained a note that a new method was in use.[7] By the third instalment in 2003 the CBGM had risen to prominence, and the preface to that instalment cites it by name as being used to "explore the relationship among the witnesses".[8] For the second edition of volume IV of the ECM in 2013 the CBGM was applied to the combined data for all the Catholic Epistles, resulting in a small number of textual changes. It provided the text for the Catholic Epistles in NA28 (the 28th edition of the Nestle-Aland hand edition of the Greek New Testament) and as such its text and that of its successors will be found on countless bookshelves around the world for decades to come. Likewise, ECM volume III will provide the text for Acts in NA29.[9]

Created by Gerd Mink in Münster, the CBGM's objective is "to establish a

---

[3] The IGNTP is a transatlantic committee overseeing work on critical editions of the Greek New Testament see http://www.igntp.org/. The work on Paul's letters is scheduled to be completed in the 2030s.

[4] See Kirchliche Hochschule Wuppertal-Bethel, http://www.kiho-wuppertal-bethel.de/.

[5] Aland et al., *ECM IV 2nd ed. (Catholic Letters)*, 21*.

[6] See Aland et al., *ECM IV.1 (James)*.

[7] See Aland et al., *ECM IV.2 (Peter)*.

[8] Aland et al., *ECM IV.3 (1 John)*, 37*; The preface is reprinted in Aland et al., *ECM IV 2nd ed. (Catholic Letters)*, XV.

[9] The intention is that this trend will continue and the ECM will eventually provide the entire text for the NA.

comprehensive hypothesis for the genealogical structure of the textual tradition [and] to examine the validity of textual decisions”.[10] Creating a stemma and a critical text is not new – Lachmann was doing this in the nineteenth century – but the CBGM now makes a startling claim: “<u>All the evidence is taken into account</u>. Statements about the relationships between copies of the text are based on an evaluation of all variants in all manuscripts that are judged to be useful for editing the text and reconstructing its history.”[11] This statement is key, in that the CBGM is able to manage the huge amount of data involved in considering the complete textual evidence of large numbers of manuscripts.

In addition to the basic challenge of the volume of the data, the New Testament is also a highly contaminated tradition. Consider two manuscripts X and Y. It is common that there will be places in the text where they differ, attesting readings x^1 and y^1 and it is determined that x^1 developed from y^1 (perhaps as a grammatical correction, for example). At the same time, at another place in the text they might read x^2 and y^2, yet here it is determined that y^2 developed from x^2 (perhaps mistakenly omitting a word). In other words neither manuscript will simply present itself as the ancestor of the other. Mink describes this phenomenon by saying that “[Textual flow] in both directions can be demonstrated for almost every pair of witnesses.”[12]

In Mink’s seminal work of 2004, ‘Problems of a highly contaminated tradition’, he “attempts to show a new way (i) of finding and evaluating the genealogical data that can be used to construct a stemma of such texts, and (ii) of constructing a stemma that reflects all genealogical data.”[13] This new way is the CBGM, and David Parker states that because of it: “We are at last able to make Lachmannian stemmatics workable in complex textual

---

[10] Mink, ‘The Coherence-Based Genealogical Method (CBGM) – Introductory Presentation, 1.0’, 9.

[11] Spencer, Wachtel, and Howe, ‘The Greek Vorlage of the Syra Harclensis’, para. 14 (my emphasis).

[12] Mink, ‘Contamination, Coherence, and Coincidence in Textual Transmission’, 144.

[13] Mink, ‘Problems of a Highly Contaminated Tradition’, 13.

traditions."[14]

One key point of divergence, however, from traditional stemmatics is unavoidable in the face of the levels of contamination present in the New Testament: in a contaminated tradition, the stemma of witnesses is not straightforward. Contamination requires that some witnesses have more than one parent. Indeed, in a tradition with contamination and also many missing manuscripts, the effect of this is magnified.

The Method

> *Gerd Mink, creator of the CBGM: "I want the consequences of textual decisions for the whole of the textual history to be understood."*[15]

At the heart of the CBGM is the idea of applying genealogical methods to the individual variant unit not just to the text as a whole, and from that to calculate information about the relationships between texts of manuscripts. As Mink states, "...the CBGM derives genealogical relationships between witnesses from genealogical relationships between their variants. The gist of this method is a way to map genealogical relationships between variants into coherent fields within a global stemma of witnesses."[16] Below is a brief introduction to the terminology and steps of the method.

Peter Head states that "The Coherence-Based Genealogical Method is difficult to summarise briefly..."[17] Stephen Carlson, in his introduction to the CBGM says "Owing to the complexity of the Method, it is very possible that I have not completely understood various aspects of it..."[18] These statements are representative of a widely held view that the CBGM is

---

[14] Parker, *Textual Scholarship and the Making of the New Testament*, 84. See Chapter 4 for a discussion of Lachmannian Stemmatics.

[15] Mink, 'Editing and Genealogical Studies: The New Testament', 52.

[16] Mink, 'Contamination, Coherence, and Coincidence in Textual Transmission', 150f.

[17] Head, 'Editio Critica Maior: An Introduction and Assessment', 143.

[18] Carlson, 'Coherence-Based Genealogical Method', para. 1.

difficult to understand. With this common response in mind I have supplemented the following information with a full worked example of the CBGM in Chapter 3, aimed at demystifying the method by using simple data.

Texts not Manuscripts

Consider the following facts of the New Testament tradition: there are many missing manuscripts; high levels of contamination are evident; there is frequent uncertainty of manuscript provenance; and a manuscript can only provide an upper limit on the date of its own text. Thus it is not possible to reconstruct the actual history of transmission or the genealogy of the manuscripts.

A crucial difference between traditional stemmata and those produced by the CBGM is that the CBGM considers texts and not manuscripts. A claim of the CBGM is that in reconstructing the history of the first millennium of the text much crucial evidence is provided by manuscripts from the second millennium, where there is evidence in later manuscripts of earlier text forms – potentially even earlier than our earliest extant manuscripts.[19]

A second consequence of considering texts not manuscripts is that the stemmata produced by the CBGM do not require the traditional concept of hyparchetypes: “In a conventional stemma, hyparchetypes take the place of lost manuscripts because the stemma is meant to show which manuscripts were copied from which. A stemma that does not try to exhibit relations between manuscripts but between texts can do without hyparchetypes because the aim is to show the pattern of change that has taken place in the tradition as represented by preserved manuscripts.”[20] Somewhat confusingly, although no better idea presents itself, the stemmata produced by the CBGM still have nodes labelled using

---

[19] See Mink, ‘Contamination, Coherence, and Coincidence in Textual Transmission’, 146.

[20] Spencer, Wachtel, and Howe, ‘The Greek Vorlage of the Syra Harclensis’, para. 14.

manuscript references – because they are the easiest way of referring to the text to which they bear witness.

Wachtel offers an important clarification, that "a textual flow diagram shows a structure that integrates states of text and must not be confused with an outline of actual historical processes or a stemma of manuscripts."[21] And again "Textual history, according to the CBGM approach, deals with texts carried by manuscripts, not with the manuscript traditions as such. The relative chronology of the textual development is quite independent of the history of manuscripts."[22] Despite such statements by members of the INTF over many years, it is not uncommon for other scholars to misunderstand this important premise.[23] This is perhaps unsurprising, when the introduction to ECM volume III contains the following footnote, highlighting the problem nicely: "When the term "manuscript" is used in the following, it means the text transmitted in the manuscript."[24] Only those who read that footnote will note the redefinition of the term.

Assumptions

The CBGM is based on four assumptions. Within the bounds of parsimony, the following are asserted to hold true *more often than their opposites*:[25]

1. Scribes attempted to copy faithfully
2. Scribes did not invent readings
3. Scribes used few sources
4. A scribe's sources were closely related

---

[21] Wachtel, 'The Coherence Method and History', 1.
[22] Ibid., 5.
[23] See Chapter 2, Secondary Literature / Other Authors.
[24] Strutwolf et al., *ECM III (Acts)*.1, 28*.
[25] See Mink, 'Contamination, Coherence, and Coincidence in Textual Transmission', 153ff.

A further underlying hypothesis of the CBGM is that every variant reading can be said to derive from a reading in other witnesses (including the hypothetical initial text). Existing readings may be combined, shortened, lengthened, changed, deleted and so on. Parsimony dictates that the solution using the fewest parent readings is to be preferred here.

Several years later, the editorial team of ECM Acts wrote a paper expanding these assumptions combined with their experience from Acts into nine guidelines for creating local stemmata. See "INTF Publications" below and Chapter 3 for further details. The guidelines should be read by anyone seeking to apply the CBGM.[26]

Terminology

The Initial Text (*Ausgangstext*)

One methodological innovation of Mink and the CBGM is the introduction of the concept of the initial text (or *Ausgangstext* in German – hence its label "A"). The initial text is a hypothetical text from which all extant witnesses are descended. It is likely to be different from the authorial text (if there is such a thing as a single authorial text) and is also different from the idea of a hyparchetype since the CBGM can delve deeper into the textual history than a hyparchetype of the oldest available manuscripts. In Mink's words, "This initial text may be hypothesized as representing the starting point of the tradition at each passage."[27]

Wachtel offers a useful clarification that "the initial text is neither the same as the archetype nor the authorial text. The archetype marks the beginning of the manuscript tradition, being itself a manuscript now lost... The reconstruction of the initial text can never be more than a hypothesis which may be convincing at most places, while serious doubts

---

[26] See Gäbel et al., 'The CBGM Applied to Variants from Acts: Methodological Background', 3.

[27] Mink, 'Editing and Genealogical Studies: The New Testament', 52.

remain at some."[28]

Coherence

The idea of coherence is fundamental to the CBGM (hence the name). It comes in three types (pre-genealogical, genealogical and stemmatic) and crucially forms new external criteria for text-critical decisions. Pre-genealogical coherence describes the level of agreement between two witnesses as a percentage reflecting the number of agreements divided by the total number of variant units found in both witnesses.[29] It does not have a direction. This simple formula will be discussed further in Chapter 9 and compared to more objective measures. It stands to reason that agreement of readings in closely related witnesses may be assumed to be non-coincidental.

Genealogical coherence combines this coherence percentage with knowledge of the relationships between the variant readings, providing a direction to the coherence. The witness with the most prior readings (as opposed to posterior) becomes a potential ancestor of the other and thus the direction of predominant "textual flow" is discovered.[30] Textual flow is the key to the construction of the CBGM's stemmata.

Stemmatic coherence "is found between ancestors and descendants in the global stemma, which presents the simplest hypothesis for which ancestors are necessary to explain the full text of each descendant at every point of variation."[31]

These new external criteria are not introduced in an attempt to diminish the role of traditional criteria. However, as Mink points out: "My intention is to reduce subjective evaluation of readings and witnesses and to strengthen the external criteria. But internal

---

[28] Wachtel, 'The Coherence Method and History', 2.

[29] A variant unit is a section of the text where the different manuscripts attest different readings of the text.

[30] As well as prior and posterior, some readings are classed as unrelated (e.g. cousins in a family tree).

[31] Mink, 'The Coherence-Based Genealogical Method (CBGM) – Introductory Presentation, 1.0', 476.

criteria are important now as before. They are not really restrained, but applied in a new methodical environment. The external criteria become more powerful and reliable if we pay close attention to the principle of coherence and to the role each witness plays in the textual flow."[32]

Local Stemmata

The idea to create a stemma for each variant unit (a local stemma) is described by Parker as "an essentially simple but brilliant stroke".[33] These local stemma are generally small and easy to understand and are created using traditional text-critical criteria, along with the extra information provided by coherence. As discussed above, genealogical relationships between variant readings are used to derive genealogical relationships between witnesses. Equally important, however, is the opposite concept where already determined coherence values are used to help to determine the relationships between variant readings. Thus if there are two readings (α, β) for a variant unit then not only is the traditional text-critical work undertaken to consider the relationships between the readings (e.g. β probably derives from α because they seem related and it is the *lectio difficilior*) but also the relationships between extant witnesses of each reading are considered. If the witnesses of α are not potential ancestors of witnesses of β then it is considered unlikely that β could derive from α. If, conversely, the witnesses attesting β are potential ancestors of those attesting α then the CBGM suggests that it is more likely that α derives from β, and so this must be considered when making this local stemma.

Overview of the Method

Mink, the creator of the CBGM, states: "The Coherence-Based Genealogical Method

---

[32] Mink, 'Editing and Genealogical Studies: The New Testament', 56.

[33] Parker, *An Introduction to New Testament Manuscripts and Their Texts*, 169.

makes no textual decisions. It merely reveals an image of the tradition which emerges from a text-critical philological study of all the variants. The iterative process of the method helps the text-critical philological hypotheses to confirm their plausibility."[34] The iterative nature of the CBGM means that later findings are used to reconsider earlier decisions. For example, a local stemma may have made perfect sense when it was created, but once enough such local stemmata are available, and so genealogical coherence data has been calculated, the original local stemma may no longer hold true. It would then have to be changed, which in turn would subtly affect the balance of the genealogical coherence. Enough local changes can flip the prevailing direction of textual flow between witnesses and thus the direction of genealogical coherence. It seems reasonable to assume that the results will stabilise after enough iterations have been completed.

The starting point for the CBGM is to establish as many uncontroversial local stemmata as possible. When constructing a local stemma all the attestations of each reading must be considered. The pre-genealogical coherence of these attesting witnesses is used to inform decisions about the coincidental emergence (or not) of multiple attestations of the same reading. For example, to find two witnesses with very low pre-genealogical coherence (say 75%) that share an otherwise unattested reading suggests that the reading emerged independently twice.[35] Alternatively, two witnesses of pre-genealogical coherence 98% sharing a rare reading that could easily have independently emerged (considering traditional text-critical criteria) would nevertheless suggest that the attestation is probably related. As these local stemmata are created the genealogical coherence between witnesses begins to take shape as in each case prior and posterior readings are established.

---

[34] Mink, 'The Coherence-Based Genealogical Method – What Is It About?', para. 21.

[35] Note that the average agreement between any given pair of witnesses of the Catholic Epistles is 87.6%. See Mink, 'Contamination, Coherence, and Coincidence in Textual Transmission', 157. Nevertheless, relationships between the text from very early manuscripts (in a much sparser tradition than later evidence) often yield considerably lower coherences than the average.

Genealogical coherence, and the textual flow diagrams that are created using it, are then used to check and revise local stemmata, and to define hitherto undefined local stemmata. Local stemmata will either hold true in the light of textual flow, or not – for example a local stemma is held to be false if no witness attesting a parent reading is an appropriately highly ranked ancestor of witnesses attesting the child reading. See Chapter 3 for a detailed example of defining initial local stemmata, and applying genealogical coherence to define more, and refine existing, local stemmata. This process can now be iterated, until the scholar is satisfied with the local stemmata.

Once such local stemmata have been created work begins on creating the global stemma. Mink highlights a problem, however: "The final task seems to be very easy: all the trees mirroring the connections inherent in the local stemmata must be calculated, and we must select only the tree with the fewest branches. But this is not practicable, for the number of trees reflecting the local stemmata is enormous, so we must reduce the possibilities at the beginning and during the work without cutting off essential ones."[36] The reduction of possibilities is achieved by creating an optimal substemma for each witness and then combining all these substemmata into one or more global stemmata. An optimal substemma is a stemma containing the fewest ancestors necessary to explain all the readings in the witness. Each reading is either present in an ancestor, or emerged as a new variant originating from a reading found in an ancestor. Note: a variant is considered "connective" when coincidental emergence is unlikely or coherence is high between witnesses. Connective variants are important when creating an optimal substemma, as they will affect which witnesses need to be included.

Because of contamination, it is sometimes borderline which of a pair of witnesses is

---

[36] Mink, 'Editing and Genealogical Studies: The New Testament', 55f.

considered the ancestor and which the descendant in a substemma. The overall textual flow is used to decide, but in that case altering the decision of only a handful of variant passages might change the overall decision. Such changes are to be expected in subsequent iterations of the method, and even subsequent editions of the ECM. The difference in the number of prior readings in each of the pair of witnesses is referred to as the strength of the textual flow. See Chapter 8 for suggested improvements concerning the strength of textual flow.

A Controversial Example: 1 Peter 4:16/24-28

The INTF have provided a list of the changes between the NA27 and NA28 texts of the Catholic Letters.[37] Peter Head discusses the list and points out a surprising choice for the initial text (*Ausgangstext*) of 1 Peter 4:16/24-28, explored here in detail:[38]

- *a* ECM, NA28: … τω μερει τουτω (let him glorify God in *this matter*)
- *b* NA26, NA27, UBS4: … τω ονοματι τουτω (let him glorify God by *this name*)

UBS4 (the fourth edition of the United Bible Societies Greek New Testament) doesn't even mention a textual variant here. NA28 lists the support for *a* as "P 307. 642. 1448. 1735. Byz." (no extant manuscripts earlier than the ninth century) and *b* as "𝔓72 א A B Ψ 5. 33. 81. 436. 442. 1175. 1243. 1611. 1739. 1852. 2344. 2492. latt sy co; Cyr". Holmes, in the SBLGNT, includes *b* in the text with the following in the apparatus: "ὀνόματι WH Treg NIV] μέρει RP", showing that *b* is preferred by Westcott and Hort, Tregelles and the New International Version (NIV) while *a* is found in Robinson-Pierpont (RP = Byz.) alone.[39] The

---

[37] See Wachtel, 'Textual Changes in NA28'.

[38] This is an ECM reference where 24 refers to the word number at the start of the variant unit and 28 the word at the end. Head, 'Editio Critica Maior: An Introduction and Assessment', 144.

[39] Holmes, *The Greek New Testament. SBL Edition*, 463 (SBLGNT); For Holmes' abbreviations, see ibid., ix–x; Westcott and Hort, *The New Testament in the Original Greek* (WH); Tregelles, *The Greek New Testament, Edited from Ancient Authorities, with Their Various Readings in Full, and the Latin Version of Jerome* (Treg); Goodrich and Lukaszewski, *A Reader's Greek New Testament* (NIV); Robinson and Pierpont, *The New Testament in the Original Greek* (RP).

ECM editors state that *a* is the harder reading and print it in their critical text, although on the face of it there is a strong case for *b* as harder instead.[40] The ECM lists three more variants (τω μερει τουτου; τουτω τω μερει; and τω μερει τουτω η τω ονοματι τουτω) but we need say no more about them as they are all considered posterior to *a* or *b*.

So how has the ECM's committee come to choose such a reading over one with such obvious support from so many "good" manuscripts? Sceptical of this, Head says that if true, this "would be an example where the CBGM allows the critic to recognise the Ausgangstext in one late strand of the manuscript tradition."[41] This ability, of course, is a major claim of the CBGM.

This decision can be interrogated using the "Genealogical Queries" programs (first edition).[42] First, Figure 1 shows the textual flow diagram for variant *a* created by the "Coherence in Attestations" program with connectivity 10 (average) and setting the initial text to *a* (i.e. the text as printed in the ECM). Coherence is perfect here – i.e. for each witness to variant *a* one of its ten most likely potential ancestors also reads *a*, back to 468 and then the initial text A. In fact, in all but five places in this diagram each witness descends from its highest ranking potential ancestor and even in those five cases the lowest ranked ancestor to appear is third.

Figure 2 shows the textual flow diagram for reading *a*, with initial text set to *b* (i.e. the text printed in NA26, NA27, UBS4, …). This time reading *a* is shown as having emerged from *b* twice – in 025 and 2298 with 025 being the reason for every other attestation (via 468). Again there are five instances of a non-first ranked ancestor being required to explain the reading but this time the lowest rank seen is seven for 1735, and in this case the first six

---

[40] Mink, 'The Coherence-Based Genealogical Method (CBGM) – Introductory Presentation, 1.0', 218.

[41] Head, 'Editio Critica Maior: An Introduction and Assessment', 144.

[42] INTF, '"Genealogical Queries" v1 (for the 1st Edition of ECM Vol IV)'.

most likely potential ancestors of 1735 read *b*.[43] It is worth noting that if the connectivity had been set to 6 rather than 10 then this would be counted as a third emergence of the variant. This numerical ranking system has the potential to be misleading, however, as the percentage difference in coherence between the first (02, 90.365%) and seventh (642, 89.635%) potential ancestors is only 0.73%. Compare this to the difference between the first and second most likely potential ancestors of, say, manuscript 1832: 876 at 98.547% and 424 at 92.586% - a difference of almost 6%. Incidentally, 1832 and 876 read *b*, while 424 reads *a*. The validity of using these rankings versus percentages will be explored in Chapter 7.

---

[43] In the first diagram 1735 is shown immediately below A as its second ranked ancestor.

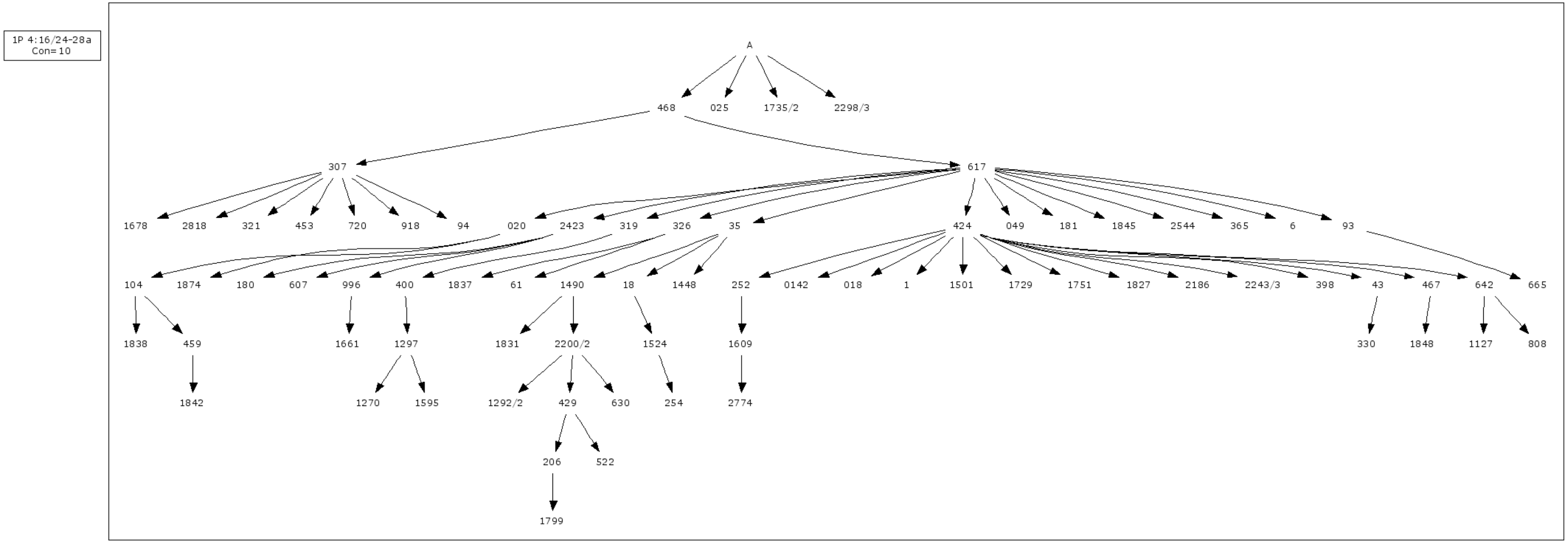


Figure 1: Coherence in Attestations, variant a with initial text a for 1 Pet 4:16/24-28

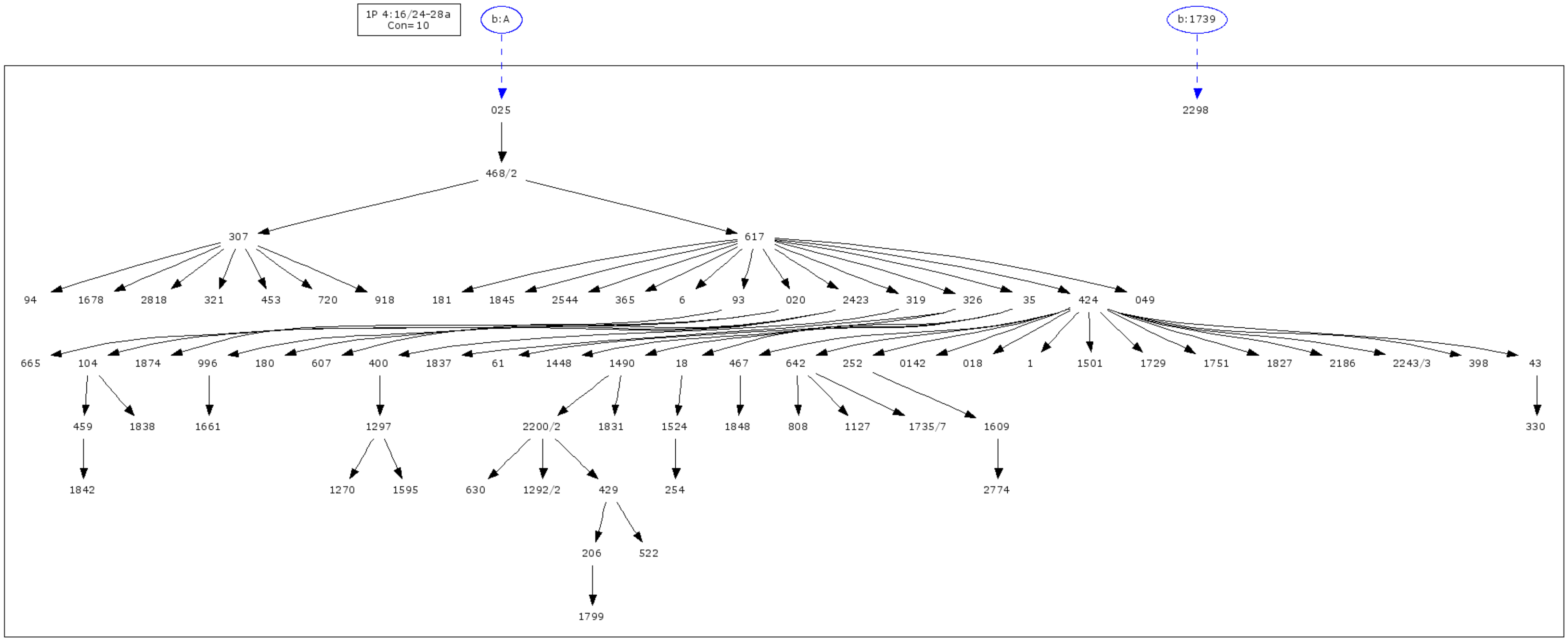


Figure 2: Coherence in Attestations, variant a with initial text b for 1 Pet 4:16/24-28

Figure 3 shows the textual flow diagram for reading *b*, with initial text set to *a.* This time reading *b* is a case of multiple emergence from four different ancestors, all attesting *a*. The question here, then, is how likely does it seem that this fourfold emergence would occur from reading *a* found only in later extant manuscripts, to *b* found in 𝔓72 and other early manuscripts?

Finally, Figure 4 shows the textual flow diagram for reading *b*, with initial text set to *b.* Again reading *b* is a case of fourfold emergence: in the initial text and three further times, all originating from reading *a* – even though *b* is asserted to be the initial text. This is the crucial problem for including reading *b* in the initial text – it must have been changed to *a* and then changed back to *b* at least three times in the tradition's history. Note that 2492 only just creeps in with its tenth ranked potential ancestor, otherwise there would be yet another emergence to deal with. It is also worth noting that the three non-initial text emergences form three relatively small groups of manuscripts with the lion's share of witnesses to reading *b* descending unbroken from the initial text.

The most likely of these scenarios, it was decided by the ECM editors, is the first one: the initial text reads *a* – and the coherence evidence above does seem to justify this choice. Figure 5 shows the resulting local stemma. Mink states that "The editors of the Editio Critica Maior consider variant a to be the more difficult reading and assume that variant b has developed from variant a. Even if variant b was the original, variant b must have been derived several times from variant a coincidentally. Variant b remains a case of multiple origins."[44]

---

[44] Mink, 'Introductory Presentation', 218.

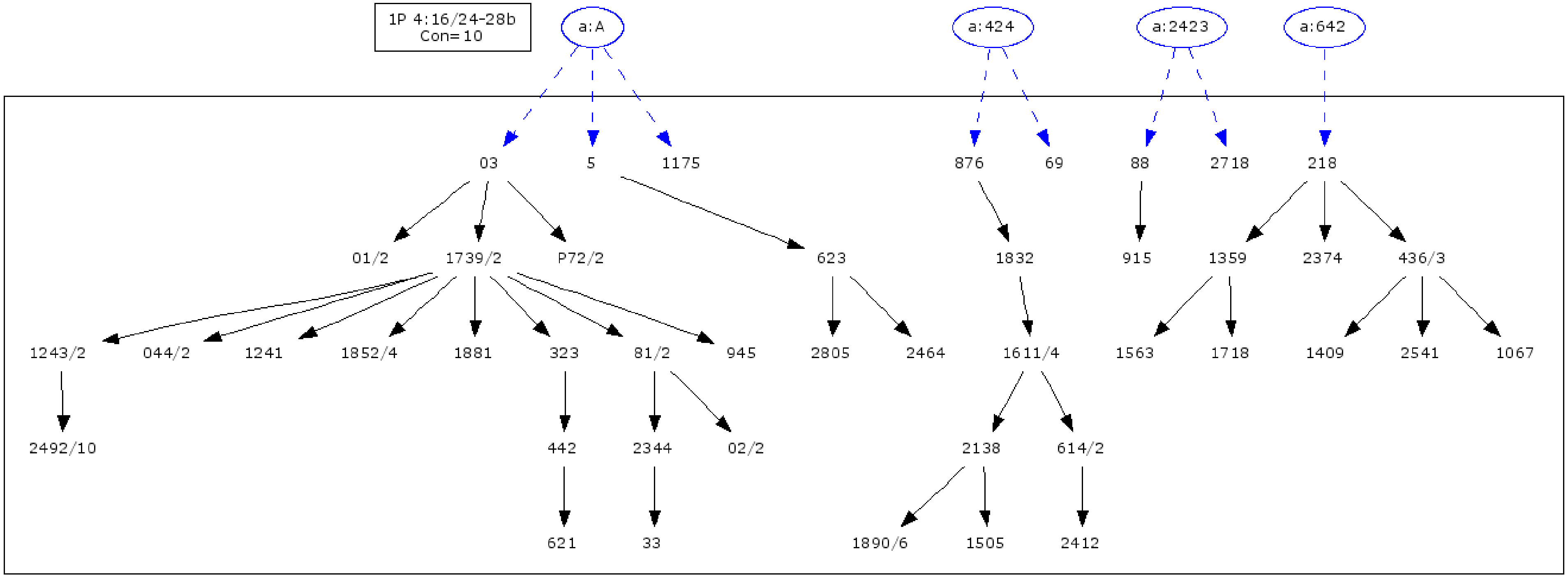


Figure 3: Coherence in Attestations, variant b with initial text a for 1 Pet 4:16/24-28

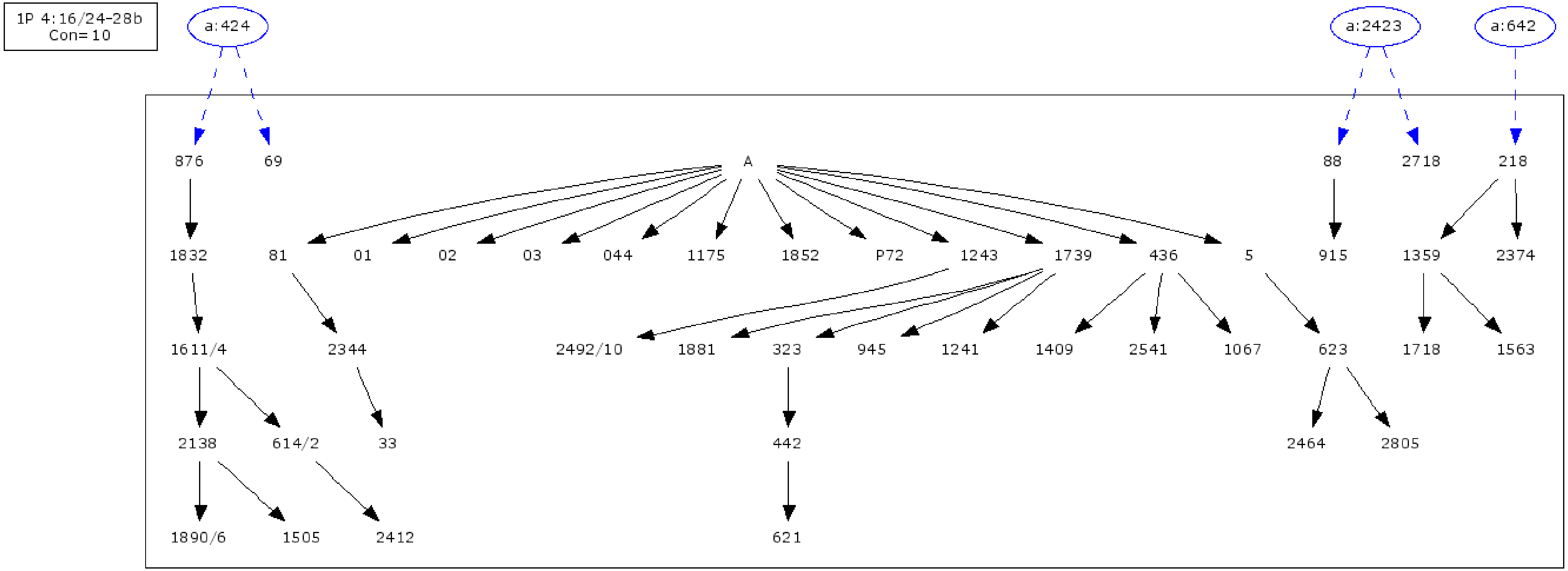
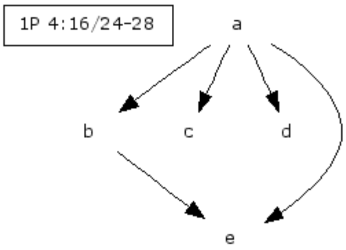


Figure 4: Coherence in Attestations, variant b with initial text b for 1 Pet 4:16/24-28



a) τω μερει τουτω
b) τω ονοματι τουτω
c) τω μερει τουτου
d) τουτω τω μερει
e) τω μερει τουτω η τω ονοματι τουτω

Figure 5: Local stemma for 1 Pet 4:16/24-28

It is hard to argue with the CBGM data here, although perhaps easier to argue with the editors' assertion that reading *a* is the *lectio difficilior.* Even arguing that *b* were the most difficult, and showing that it has the "best" and oldest manuscript support, it is hard to accept the resultant story depicted by the textual flow diagrams above (with *b* as the initial text). Incidentally, for the equivalent of Figure 4 to show perfect coherence (all witnesses to reading *b* appear in a single tree), it is necessary to increase the connectivity setting to 40 and perfect coherence cannot be achieved at all for Figure 2. Nevertheless, the hypothesis that *b* is the initial text (as has been asserted in previous NA editions) is not a lost cause. All that is required is for reading *a* to change (be corrected back?) to *b* in three manuscripts – descendants of 424, 2423, 642 – each forming a relatively rarely copied tradition.

Primary Literature

Introduction and Genre

For the purposes of this chapter, all works on the ECM or CBGM produced by the INTF in Münster are considered primary literature and all others secondary. This section sets out a history of the development of the primary literature concerning the CBGM.

The primary literature does not attempt, it would appear, to introduce the CBGM to non-philologists. On the one hand this might seem reasonable – as who but a philologist would need to apply the CBGM? But the potential impact of the method on the *ongoing* transmission of the text of the Greek New Testament (for example in the NA28, NA29 and so on) has led to a kindling of interest far outside the editorial committees of such editions. Thus the interested, perhaps even concerned, readers find themselves confronted with works in a genre that they are not easily able to digest. This might go some way to explaining the misunderstandings and not uncommon negative reaction to the CBGM from some quarters.[45]

In 2004 Mink published an article '"Problems of a highly contaminated tradition: the New Testament – Stemmata of variants as a source of a genealogy for witnesses' in the journal *Studies in Stemmatology II.* Wachtel and Holmes, in their introduction to *The Textual History of the Greek New Testament* refer to this as Mink's "seminal study" and as is implied by this. and citations in several other articles and books, 'Problems of a highly contaminated tradition' is Mink's major publication on the CBGM, its terminology and its use.[46] Mink puts forward his theory on contamination in the New Testament tradition: that "if contamination occurs, it emerges from those texts which were at the disposal of the scribe, i.e. texts in his

---

[45] See Chapter 2. For example, the discussion of Porter, *How We Got the New Testament*.

[46] Wachtel and Holmes, *The Textual History of the Greek New Testament*, 6.

direct environment, i.e. texts which are, for the most part, closely related with each other."[47] He describes the decision only to include firsthand readings in continuous text in the ECM, therefore "not counting variants by correctors, in marginal text and commentaries."[48] This decision will be discussed further and challenged in Chapters 3 and 9. He argues that later manuscripts contain evidence of a "very old text, only slightly altered in the course of time" thus introducing the key methodological point of distinguishing between manuscripts and the text to which they bear witness: "It is impossible to find the genealogical relationship between manuscripts if most of them have not survived. It is merely possible to uncover the genealogical structure of the preserved texts. In this context, therefore, the text is the witness and not the manuscript."[49]

The rest of the paper is given over to a description of the method and discussions of cases where a stemma produced by the standard CBGM procedures would be in some way incorrect, incompatible or "false".[50] Such local stemmata are not uncommon due to the complexity of the data, for example: "usually a number of the variants of the ancestor in a contaminated tradition are posterior to the corresponding variants of the descendant, and a number of the variants of the descendant are prior to those of the ancestor."[51] The detailed discussions of such cases, and the procedure to follow for each, are a major reason why this paper is so often referenced.

INTF Publications

We will now trace the INTF's remaining publications on the CBGM chronologically.

---

[47] Mink, 'Problems of a Highly Contaminated Tradition', 14.

[48] Ibid., 78.

[49] Ibid., 24.

[50] For true vs. false, consider the example of the global stemma. It is "true" only if it is compatible with each local stemma, all the coherence data at every place of variation and all non-coincidental emergence of variant readings. See Ibid., 30.

[51] Ibid., 51.

In 1993 Mink published a paper 'Eine umfassende Genealogie der neutestamentlichen Überlieferung' showing some of his early thoughts on the genealogy of the New Testament, contamination and his fledgling method – which he presents "nur in Grundzügen und in vereinfachter Form" ("only in outline and in simplified form"). This paper includes Mink's earliest discussion (of which I am aware) of the need to consider a genealogy of texts and not manuscripts.

In 2000 Mink wrote an article describing the development and use of the CBGM for the letter of James in the ECM: "Editing and genealogical studies: The New Testament" in *Literary and Linguistic Computing*.[52] There are no examples and little help to the uninitiated reader but this article contains early insights into the development of the CBGM. For example, the need for the middle step of optimal substemmata is explained here: because the gap between the local and global stemmata was too great to be bridged. Interestingly, Mink also explains his frustrations with trying to use the traditional apparatus in the Nestle-Aland edition for computer-aided analysis. This led to the development of the databases and applications that became fundamental to the CBGM and ECM.

Proceedings of a conference in Munster in 2001 (published in 2003) include another early paper by Mink exploring genealogical coherence: "Was verändert sich in der Textkritik durch die Beachtung genealogischer Kohärenz?" This paper was written while the CBGM was in active (early) development. In his conclusion he considers what this method changes, for example that some witnesses will be considered of less importance than before as their posterior nature has now been identified.

The INTF website hosts several resources relating to the ECM and CBGM. The CBGM overview page (dated 2002 but revised later) 'The Coherence-Based Genealogical

---

[52] I believe this is the earliest article about the CBGM in English, although the article does not name the method as the CBGM.

Method – What is it about?' is surely intended as the first thing the web user will find regarding the CBGM.[53] It sets out (in only fifteen hundred words) the reasons and core concepts of the CBGM but, as is typical of the corpus of work on the CBGM, it is not presented in such a way as to easily introduce the topic.

In 2002 Klaus Wachtel, in collaboration with two biologists from Cambridge (Matthew Spencer and Christopher J. Howe), published a paper comparing the CBGM and the cladistic maximum parsimony (MP) method from evolutionary biology.[54] They point out that MP is "one of the most popular phylogenetic methods, and has often been used in the study of manuscript traditions".[55] They set out to compare the methods by looking at a group of manuscripts of the Epistle of James whose Greek text is very similar to their text made by translating the Syriac Harclensis back into Greek. After describing each method and the results gained from both, they set out to make a comparison – but in the end they conclude that the methods produce different results and should be used in different circumstances, or with different goals in mind. For example, MP makes no use of (so doesn't require any) philological reasoning and doesn't try to find an original text – but rather just calculates an overall representation of the relationships between texts. The CBGM, in contrast, exists first to identify the initial text. Wachtel briefly revisited this project in his 2007 paper in *Editing the Bible* (published in 2012), using it to highlight characteristics of the CBGM such as this emphasis on philological input.[56]

In 2004, Spencer, Wachtel and Howe joined forces again to apply the reduced median network method from biology to the Greek textual tradition of James. The reason for this

---

[53] The article on this page is dated as 2002 although at least the bibliography has been updated more recently.

[54] Essentially, in MP, the smallest tree that explains the data is considered the best tree. Spencer, Wachtel, and Howe, 'The Greek Vorlage of the Syra Harclensis'.

[55] Ibid., para. 1.

[56] Wachtel, 'The Coherence-Based Genealogical Method: A New Way to Reconstruct the Text of the Greek New Testament'.

study was that "Many manuscripts of the Greek New Testament are influenced by multiple pathways of textual flow. This makes it difficult to reconstruct a stemma using traditional methods."[57] They build on their 2002 paper which, they argue, showed that the kind of stemma produced by MP could not properly represent the relationships between the witnesses. This second paper only refers to the CBGM once by name, but CBGM-core concepts such as textual flow and dealing with a contaminated tradition are present throughout. They argue that the reduced median method allowed them to "represent multiple pathways of transmission, to reconstruct the readings in hypothetical ancestral manuscripts, and to determine whether there are characteristic readings identifying groups of manuscripts." The resultant stemma contained eighty-two real and 8,517 hypothetical manuscripts to represent the textual flow of the tradition. This method can cope with contamination and large numbers of manuscripts (unlike some of the alternatives) and creates a network that, they argue, is consistent with theories about the transmission of the New Testament. No philological reasoning is used and there are restrictions that, at face value, make this method seem inappropriate for the job at hand: for example the method can only cope with binary splits, i.e. each parent has exactly two children. The output of this method is not one single stemma, in the traditional sense, as they argue: "When a text tradition is heavily contaminated, there may be many equally parsimonious trees. The solution is to produce a multi-dimensional network (the median network) that contains all of these trees simultaneously. However, such a network will often be too complex to display, in which case we can reduce its complexity by including only the most likely pathways. The result is known as a reduced median network..."[58] This then, "is a network in which contamination is explicitly represented."[59] They briefly discuss alternative

---

[57] Spencer, Wachtel, and Howe, 'Representing Multiple Pathways of Textual Flow in the Greek Manuscripts of the Letter of James Using Reduced Median Networks', 1.

[58] Ibid., 2f.

[59] Ibid., 1.

network methods, saying that none is currently suitable for textual criticism but research is ongoing. Nevertheless, "The theoretical advantage of network methods is that they can produce explicit hypothetical descriptions of the history of heavily contaminated text traditions."[60] See Part Two (Chapters 4, 5 and 6) for information and analysis about using phylogenetics to analyse manuscript traditions.

At the annual meeting of the SNTS (Studiorum Novi Testamenti Societas) in 2005 in Halle, Wachtel and Parker presented a paper introducing the IGNTP/INTF's joint venture to create the ECM volume for John's gospel and also an electronic critical edition of John – using the CBGM. Parker provided an overview and a discussion of the concept of the initial text. Wachtel presented an introduction to the CBGM based on an example from James. He made several important points arguing that the CBGM is a tool under the control of the textual critic and not the other way round, for example: "Philological reasoning has a clear preponderance over stemmatological procedures. At every point it can be determined what the software we use is doing. There is no 'stemmatological black box'."[61] Talking about genealogical coherence and making text-critical decisions, Wachtel highlights the principle that "One can opt against prevalent tendencies where ever one sees good reasons against them. But it is an integrated part of the method that such tendencies cannot be ignored."[62] He goes on to argue that the CBGM has highlighted that "in many cases we do not know 'wie es wirklich gewesen -- how it was in reality', – that we are dealing with probabilities rather than with a fixed 'Urtext'."[63]

At the 5th Birmingham Colloquium on the Textual Criticism of the New Testament in 2007 Wachtel presented a paper 'Towards a Redefinition of External Criteria: The Role of

---

[60] Ibid., 12.
[61] Wachtel and Parker, 'The Joint IGNTP/INTF Editio Critica Maior of the Gospel of John', 9.
[62] Ibid.
[63] Ibid.

Coherence in Assessing the Origin of Variants.’[64] This paper contains another basic introduction to the CBGM, showing further insight into the start of the process: “passages that don’t pose text-critical problems are used to collect evidence that the assessment of genealogical coherence can build on.”[65] He also includes an interesting example from Jude 5, five words that have yielded thirty variants: “εἰδότας ὑμᾶς ἅπαξ πάντα ὅτι Ἰησοῦς”.[66] The ECM prints this text (supported by 03), while NA27 prints “εἰδότας [ὑμᾶς] πάντα ὅτι [ὁ] κύριος ἅπαξ” (a reading not found in the ECM apparatus although its component readings are present). The discussion focuses on the subject Ἰησοῦς, κύριος or θεος – where the majority of the NA27 editorial committee argued, according to Metzger, that “[despite] the weighty attestation supporting Ἰησοῦς... the reading was difficult to the point of impossibility, and explained its origin in terms of transcriptional oversight”.[67] Wachtel describes the decision-making process that led to the ECM reading “and in this case it was the coherence of the attestation that tipped the scale...”[68] This is an example of a Byzantine NA27 reading being replaced in the ECM with an “older” reading.

Klaus Wachtel presented another introduction to the CBGM in 2007 at the Conference on Editorial Problems in Toronto in his paper ‘The Coherence-Based Genealogical Method: A New Way to Reconstruct the Text of the Greek New Testament.’ Updated conference proceedings were published in 2012 by SBL as *Editing the Bible: Assessing the Task Past and Present*.[69] Justifying a key assumption of the CBGM, that “All surviving witnesses are related to each other and there is coherence within the entire tradition”, Wachtel points out that in the Catholic Letters, only 39 pairs of manuscripts have less than 80% coherence, most

---

[64] This was published in 2008 in Houghton and Parker, *Textual Variation: Theological and Social Tendencies?*

[65] Wachtel, ‘Towards a Redefinition of External Criteria’, 113.

[66] Ibid., 121ff. Wachtel claims 30, yet I count 31 variants in the ECM apparatus.

[67] Metzger, *A Textual Commentary on the Greek New Testament*, 657.

[68] Wachtel, ‘Towards a Redefinition of External Criteria’, 123.

[69] See Kloppenborg and Newman, *Editing the Bible*.

pairs have greater than 85%, all but three pairs have a closest potential ancestor greater than 90%, and those three have greater than 87%.[70] He goes on to describe the method, includes an example of a local stemma and the philological reasoning used in its creation and shows how it was revised in light of coherence data from the whole of the Catholic Letters.[71] The resulting local stemma is quite different following the revision, although the root is unchanged. He argues that the CBGM process lets the critic keep track of decisions, and also claims that “What we gain is an external criterion for assessing textual variation that is far more discerning that the old text-type model.”[72]

Holger Strutwolf also presented a paper at the Toronto conference, ‘Scribal Practices and the Transmission of Biblical Texts: New Insights from the Coherence-Based Genealogical Method.’ He begins by saying “The most important task of textual criticism is to reconstruct the original text, or to be more modest: to establish a sound and well-argued hypothesis about the initial text of the transmission of a certain piece of literature that was handed down to posterity via manuscripts.”[73] His point is that establishing an initial text cannot be done away with in favour solely of a study of the history of the text and therefore the various classical text-critical criteria are needed. These criteria can be seen as attempts to formalise the outputs of scribal behaviour. He goes on to argue that using the data produced by the CBGM for the Catholic Letters it is now possible to consider a manuscript’s potential ancestors and so see the scribe’s behaviour in a new light: “We can reconstruct the scribal habits much better than was possible before.”[74] He gives several examples, for example in 1 Peter 2:18/32 the longer and smoother reading is the ancestor of a shorter, rougher one.[75] He

---

[70] Wachtel, ‘The Coherence-Based Genealogical Method: A New Way to Reconstruct the Text of the Greek New Testament’, 129.
[71] See ibid., 131ff.
[72] Ibid., 136.
[73] Strutwolf, ‘Scribal Practices and the Transmission of Biblical Texts’, 139.
[74] Ibid., 148.
[75] Recall this is an ECM reference where 32 refers to the word number at the start of the variant unit.

argues that the scribes were not intending to make a rougher reading but succumbed to a common error; the accidental omission of small words. Therefore, as a result of the CBGM, even the *internal* criteria of textual criticism need re-evaluation – it does not just add extra external criteria. Referring to Hort's famous quotation on "Knowledge of documents" he concludes that "judgement on readings must always be combined with the knowledge of the overall picture of transmission" – something which the CBGM claims to provide.[76]

Then in 2008 Mink presented a lengthy session introducing the CBGM at the colloquium on *The Textual History of the Greek New Testament: Changing Views in Contemporary Research* in Münster. The conference proceedings were published in a book of the same title by SBL in 2011, edited by Wachtel and Holmes. The book contains eight papers, but roughly two thirds of its pages are taken up by this paper by Mink: 'Contamination, coherence, and coincidence in textual transmission: The Coherence-Based Genealogical Method (CBGM) as a complement and corrective to existing approaches.' Jongkind states that "the real reason for the conference was to initiate a discussion of the Coherence Based Genealogical Method."[77] He goes on to say that, at that time, this paper was "by far the best introduction to the CBGM available."[78] This paper is an extensive introduction to the CBGM, its assumptions, terminology, method and also the "Genealogical Queries" software. For example, there are sections discussing potential ancestors versus stemmatic ancestors, the use of textual flow diagrams etc.. There is an interesting and detailed example of creating an optimal substemmata for a witness, and interrogating the results.[79] Perhaps attempting to counter some of the mistrust growing about the CBGM, Mink

---

[76] Strutwolf, 'Scribal Practices and the Transmission of Biblical Texts', 160. The quote from Hort is: "Knowledge of documents should precede final judgement upon readings", from Westcott and Hort, *The New Testament in the Original Greek, Introduction and Appendix*, 31. (Note, the introduction was written by Hort alone. See ibid 18.)

[77] Jongkind, 'Review of "The Textual History of the Greek New Testament", Wachtel, Klaus and Michael Holmes Eds.', 147.

[78] Ibid.

[79] See Mink, 'Contamination, Coherence, and Coincidence in Textual Transmission', 189ff.

emphasises several times the strong place of philological reasoning within the method. Similarly, in the conclusion of this book, Wachtel states that "The CBGM is not the philosopher's stone that produces a reconstruction of the initial text automatically. Yet it makes visible and evaluates *coherence* – a class of evidence that could not be reliably gathered and surveyed before the adoption of database technology."[80]

The first bibliography entry on the web page 'The Coherence-Based Genealogical Method – What is it about?' (mentioned above as claiming to have been last updated in 2002) is a large presentation from 2008 that provides the first detailed examples of using the CBGM: 'The Coherence-Based Genealogical Method. Introductory Presentation by Gerd Mink. Release 1.0.' This presentation is thorough and very useful for understanding the method. Appropriately for a method with a reputation for being hard to understand, it is several hundred pages long and still refers the reader to Mink's 'Problems of a Highly Contaminated Tradition' where further discussion is required.

The INTF website also hosts the "Genealogical Queries" suite of programs, and a guide to using them (also written in 2008). This guide explains how to use the five modules of "Genealogical Queries" intended for "investigation of genealogical relationships inherent in the New Testament textual tradition" and also how to interpret their results.[81] In addition, it contains brief but important comments on textual flow and a discussion about using connectivity settings to decide on stemmatic ancestors of a witness.

The second edition of the ECM (volume IV) was published by the INTF in 2013. Its introduction contains a description of how all the available manuscripts were collated by the INTF in ninety-eight test passages and many manuscripts were found to witness to the majority text in over ninety percent of the test passages. From these a small number was

---

[80] Wachtel and Holmes, *The Textual History of the Greek New Testament*, 221.

[81] Mink, 'Guide to "Genealogical Queries"', para. 1 (preface).

chosen for inclusion in the edition. In addition, all manuscripts with less than ninety percent agreement were included, making a total of 183. They claim that this "*guarantees* reliably that the critical apparatus contains all the known readings which have appeared in the history of the text from its earliest beginnings through the formation and final establishment of the Byzantine text."[82] To this they added twenty-one of the 400 Apostolos lectionaries... which "demonstrated once more that the lectionary text in no way represents a tradition independent of the Byzantine text."[83] Quotations (and rarely allusions) from all Greek church fathers up to the seventh/eighth centuries were included and two from the ninth/tenth. They included, as quotations, anything where the wording is identical to a known variant from the manuscript tradition. Back-translations of critical editions into Greek were made from Latin, Coptic and Syriac versions, wherever possible. They state that the Armenian, Georgian, Old Church Slavonic, and Ethiopian versions' relationships to the Greek text needed more research before they could similarly be included. They also included some readings not attested by any of the selected Greek manuscripts.[84]

The section 'Notes on the text of the second edition of the Catholic Letters' includes notes on a number of key concepts behind the ECM/CBGM: for example, coherences (pre-genealogical and genealogical), potential ancestors, textual flow diagrams etc.[85] It is by no means, however, an introduction to the CBGM. What it does include is a description of the aims of the ECM and makes some important claims about it. The phrase "initial text" has sparked much scholarly debate, and this section sets out what this means for the ECM: "For the Catholic Letters the aim of our reconstruction is a hypothesis about the respective

---

[82] Aland et al., *ECM IV 2nd ed. (Catholic Letters)*, 22* (my emphasis).

[83] Ibid.

[84] Ibid., 29* explains that additional readings were included in the supplement (Part 2) from Aland, *Text Und Textwert Der Griechischen Handschriften Des Neuen Testaments*; Soden, *Die Schriften des Neuen Testaments*.

[85] See Aland et al., *ECM IV 2nd ed. (Catholic Letters)*, 30*ff.

authorial texts."[86] They do *not* claim that this initial text will be identical with the authorial text. They relate how, in the 1980s, various techniques were applied to the New Testament tradition, "including cluster analysis, approaches based on graph theory, and formal concept analysis."[87] What emerged was an understanding that philological reasoning was a necessary part of any such work and just using statistical methods was not good enough. This, clearly, was a fundamental step in the process that eventually produced the CBGM in its current form. Once the data for James was available, the CBGM was developed and "aimed at developing an overview of the relationships between all witnesses involved".[88] They highlight an important claim of the CBGM – that this output provides a "comprehensive picture" of the decisions taken by the editors which in turn allows them to challenge their own decisions.[89] It was not until the third instalment (the First Letter of John) of the ECM that the CBGM was able to be used fully, as the genealogical data for James and the Letters of Peter was not available before publication.[90] For the second edition, the genealogical data from the whole of the Catholic Letters could be applied to the text of each individual letter and this led to a number of changes against the first edition.

'Notes on the text' also sets out some conclusions of the INTF that have caused scholarly disagreement. One such conclusion is that "In the course of our work on the Catholic Letters it soon became clear that the old text-type terminology is not useful for describing the evidence."[91] They justify this claim by describing the network of interrelationships between manuscripts (both Byzantine and other) that are themselves closely related to the initial text: "we have come to realise that several witnesses containing the

---

[86] Ibid., 30*.
[87] Ibid., 31*.
[88] Ibid.
[89] Ibid.
[90] See ibid., 32*.
[91] Ibid.

Byzantine text in pure form feature high rates of agreement... with the initial text A. This is surprising, because these witnesses support the Byzantine text in nearly all passages where it clearly differs from our reconstruction of the initial text."[92] In other words, the various traditional text-types are not distinct groups as was previously thought but in fact the reality is much more complex. Their final claim about the CBGM is that it "offers tools which do not impose any bias on the text-critical decision. It rather confronts the editors with the objective data of pre-genealogical coherence and with a summary of all text-critical decisions they have made..."[93] This claim will be tested in Chapter 9.

A morning session on "The Genealogical Method" was held at the Annual Meeting of the Society of Biblical Literature, San Diego, 22 November 2014. The introduction and four articles subsequently appeared as a special feature entitled 'The Coherence Based Genealogical Method' in the online journal TC, Volume 20 (2015). Here we will discuss the INTF's contribution to the special feature.[94] The first article is 'The Coherence Method and History' by Klaus Wachtel. In this short paper he offers some useful clarifications about the CBGM's concepts, such as the distinctions between pre-genealogical coherence (objective) and genealogical coherence (subjective, dependent on the editor's decisions); or the nature of the initial text.[95] See Chapter 2 for an example of the effect on the initial text of altering the underlying collation and Chapter 9 for a suggestion to make pre-genealogical coherence genuinely objective. Wachtel also points out the importance of the distinction between a textual flow diagram and a stemma, namely that textual flow diagrams are "just graphs visualizing the relationships between witnesses and their potential ancestors ranked by

---

[92] Ibid., 34*.

[93] Ibid., 35*.

[94] For the remaining contributions see Chapter 2.

[95] Rather than 'objective', Hüffmeier describes pre-genealogical coherence as "more or less non-subjective" in Hüffmeier, 'The CBGM Applied to Variants from Acts', 1.

percentages of agreement."[96] His final brief conclusion is also worth noting: "The CBGM is not the tool to use for dating manuscripts or texts."[97]

The second INTF article from SBL 2014 is 'The CBGM Applied to Variants from Acts: Methodological Background', by the editorial team of ECM Acts (Gäbel, Hüffmeier, Mink, Strutwolf, Wachtel). This more narrative article is based on the experience of applying the CBGM to Acts, and again offers some helpful clarifications. For example, the authors set out their argument for the "obvious" supremacy of "a combination of transcriptional probability and coherence analysis" (i.e. the CBGM's approach to textual criticism) over the more traditional use of intrinsic probability, related to "our knowledge about the language and style of the author."[98] The implication of their point is that the CBGM must produce more consistent (and accurate?) results than traditional textual criticism. The second section of the Acts team's paper then expands the four CBGM assumptions combined with the experience from Acts into nine guidelines for creating local stemmata.[99] They include, for example: advice on deciding when to include a majority reading in the initial text; and a note that singular readings are normally derivative. See Chapter 3 for a further discussion of these guidelines.

The third INTF SBL 2014 paper is Wachtel's 'Constructing Local Stemmata for the ECM of Acts: Examples.' These examples show that the Acts team has formalised a way of working that applies genealogical coherence (GC) and transcriptional probability (TP) separately for each variant unit and lists their findings as such. Then the results of GC and TP are compared. Helpfully, the guidelines from the previous paper are indexed against each example. This potentially highly valuable resource is hindered by appearing too much like a

---

[96] Wachtel, 'The Coherence Method and History', 6.

[97] Ibid.

[98] They say it is "obvious", and I agree. Gäbel et al., 'The CBGM Applied to Variants from Acts: Methodological Background', 2.

[99] See "Assumptions" above.

handout and we are left wishing we had the text of the presentation alongside it. Wachtel had presented preliminary versions of these two papers (‘Methodological Background’ and ’Examples’) at the sixty ninth General Meeting of the SNTS (Studiorum Novi Testamenti Societas) in Szeged the same year, in a talk entitled ‘Constructing Local Stemmata for the Editio Critica Maior of Acts.’

The fourth and final INTF SBL 2014 paper is ‘The CBGM Applied to Variants from Acts’, by Hüffmeier. In contrast to Wachtel’s paper of examples, Hüffmeier has turned her presentation into a stand-alone paper including a much fuller explanation of each of her examples.[100] This paper, again, makes frequent reference to the guidelines in the ‘Methodological Background’ paper from the same SBL session. These guidelines and the PC (pre-genealogical coherence), GC, TP notation have clearly become central to the working practices of the editorial team of Acts. She ends with a comment pointing to the core iterative nature of the CBGM: “Perhaps at the end of phase 2 we will have gained enough additional material to enable us in phase 3 to revise our decisions.”[101]

Wachtel presented at the DiXiT convention in 2016, in the ‘Digital Scholarly Editing and Textual Criticism’ workshop. His talk was entitled ‘Towards a Global Stemma of the Greek New Testament Textual Tradition: Methodological Approach’, the slides of which are available online, constituting a straightforward overview of the CBGM.[102] The DiXiT presentations were then published in 2017 in the form of “extended abstracts” in *Advances in Digital Scholarly Editing: Papers presented at the DiXiT conferences in The Hague, Cologne, and Antwerp*, edited by Boot et al., where Wachtel’s paper became called ‘A Stemmatological Approach in Editing the Greek New Testament: The Coherence-Based

---

[100] Hüffmeier’s first example takes almost three pages to discuss Acts 13:14/33 compared to Wachtel’s paper which fits two or three examples on each page.

[101] Hüffmeier, ‘The CBGM Applied to Variants from Acts’, 12.

[102] Wachtel, ‘Towards a Global Stemma’.

Genealogical Method.'[103] In this paper Wachtel briefly introduces the problem of contamination and proposes the CBGM as its solution: "...the NT tradition is highly, some say hopelessly, contaminated. However, the Coherence-Based Genealogical Method... does offer a remedy against contamination. The remedy is the result of an analysis and interpretation of coherence which is the feature of the NT tradition balancing contamination or mixture."[104] In his description of the method, he emphases an important point, in the form of a rule: "Do not try to reconstruct the manuscript tradition but focus on the states of text preserved in the manuscripts."[105] It may be useful to print such a statement on the CBGM's packaging, so to speak, as this difference between the CBGM's results and real manuscript copying relationships have caused significant confusion and suspicion from other scholars, as will be seen in Chapter 2.

In 2017 the INTF published volume III of the ECM: Acts, the introduction of which briefly explains the CBGM process as applied to Acts. There is also a textual commentary describing all variant units where the ECM Acts text differs from NA28 – something that was long promised for volume IV but never materialised. The existence of this textual commentary for volume III is very welcome. The INTF is working with the Cologne Center for eHumanities (CCeH) to create a new user interface for the CBGM, to replace the existing "Genealogical Queries". It has been released in the final phase of producing the ECM of Acts.[106] It has many advantages over the old system, and over time is likely to incorporate several of the recommendations made in this thesis.

The Acts team describe, at a high level, that they divided the definition of local

---

[103] See Wachtel, 'A Stemmatological Approach in Editing the Greek New Testament'.

[104] Ibid., 2.

[105] Ibid., 4.

[106] See INTF and Cologne Center for eHumanities, '"Genealogical Queries" Phase 4 (for ECM Vol III)'; Strutwolf et al., *ECM III (Acts)*.

stemmata into three phases.[107] In phase one they identified important witnesses – namely the witnesses closest to the NA28 text and witnesses supporting the Western text. “The editorial team paid particular attention to these witnesses, or rather to the variants transmitted by them, in the first phase of work on the local stemmata.”[108] They defined local stemmata using these important witnesses and also pre-genealogical coherence (to indicate coherence of attestations and multiple emergence of variants). They state that “In the first phase of constructing local stemmata for all variant passages, the larger number of A-related witnesses and pre-genealogical coherence provided far more sophisticated external criteria than have hitherto been available in eclectic textual criticism.”[109] At this stage they either agreed with the NA28 or left the initial text undefined. In phase two they refined the local stemmata using genealogical coherence, using the new “Genealogical Queries” software, and in particular textual flow diagrams, as expected. They then iterated over the local stemmata in a third phase, again using genealogical coherence. It was only at this stage that the Acts team’s initial text diverged from NA28.

It is important to note that the Acts team’s decision not to diverge from NA28 in phase one is not a core part of the CBGM as defined in the prior literature. Indeed, Parker has not applied the CBGM in this way in his in-progress ECM edition of John. In the first iteration of defining local stemmata Parker used traditional text-critical criteria, along with pre-genealogical coherence to define local stemmata and did not restrict himself to following NA28 for the initial text. The Acts team’s approach and Parker’s approach are both correct, according to the CBGM’s literature and analysis described in the current thesis.

At SBL 2017, Wachtel gave a paper entitled ‘Reconstructing the Initial Text of Acts:

---

[107] See Strutwolf et al., *ECM III (Acts)*, 28*-30*.
[108] Ibid., 28*f.
[109] Ibid., 29*.

Principles and Criteria', where he compared the CBGM with conventional reasoned eclecticism. He used the new CCeH user interface for the first significant time to create the diagrams for his presentation. He concluded with "Five Principles of NT Textual Criticism in Light of the CBGM".[110] They range from a rejection of text types (not news, for those familiar with the CBGM) to a strong statement in favour of majority variants: "We should not hesitate to opt for a majority variant if it[s] read[ing i]s less common or even more difficult than an early rival."[111] This is likely to be another cause of contention with more traditional textual critics.

---

[110] Wachtel, 'Reconstructing the Initial Text of Acts: Principles and Criteria', 25.

[111] Ibid.

Chapter 2 – The CBGM in Wider Scholarship

Introduction

This chapter sets out a summary of how the CBGM has been received in wider scholarship, i.e. outside the INTF in Münster. First, publications by the INTF's close collaborators David Parker, the International Greek New Testament Project (IGNTP) and the Institute for Textual Scholarship and Electronic Editing (ITSEE) are discussed, followed by the wide ranging opinions of other scholars. As in Chapter 1, this thesis only considers literature published before 1 January 2018.

David Parker, the IGNTP and ITSEE

David Parker is Executive Editor for the ECM of the Gospel of John entrusted to the IGNTP. As such, he works closely with the INTF, as do his colleagues at ITSEE, which he founded in 2005 to work on digital editions. Wallace states: "David Parker is probably Great Britain's best NT textual critic today. What he says helps to define the discipline."[1] As befits this description, Parker has contributed significantly to the literature surrounding the CBGM. Contributions by Parker, the IGNTP and ITSEE to the literature on the CBGM could conceivably have been included in the primary literature in Chapter 1, due to their authors' close working relationship with the INTF. Nevertheless, I have designated them as "secondary" literature partly for convenience, and partly because Parker's use of the CBGM has been somewhat independent of the INTF's implementation, up to 2018.

In the early 2000s Parker wrote a paper 'Through a Screen Darkly: Digital Texts and the New Testament' for the JSNT.[2] At this time the IGNTP and the INTF were already

---

[1] Wallace, 'Challenges in New Testament Textual Criticism for the Twenty-First Century', 82.

[2] Parker, 'Through a Screen Darkly'. Published in 2003, this was before the CBGM rose to prominence, and it

collaborating on “shared methods and standards in making digital transcriptions.”[3] Parker writes about the changes brought about by the use of modern computer technology, which is “changing for ever the relationship between the scholar, the text which the scholar is studying, and the text which the scholar is creating.”[4] He points out the importance of *not* just accepting “hugely detailed statistical analyses of manuscript relationships” unless accompanied by traditional methods such as “codicological and palaeographical evidence, historical information, proper philological scholarship, and common sense.”[5] He warns against the danger of textual critics not being properly prepared and informed about such emerging methods, claiming that “the current revolution is as significant as that brought about by the Complutensian Polyglot and Erasmus’s Novum Instrumentum. It is to be hoped that it will be received with less confusion.”[6] In the following years, and perhaps with these concerns in mind, Parker has written at length about this emerging discipline.

In 2008 a supplement to Novum Testamentum was published: *Jesus, Paul, and Early Christianity Studies in Honour of Henk Jan de Jonge.* This book contains a paper by Parker called ‘Making Editions of the New Testament Today’ in which Parker again discusses modern technological advances for textual criticism. He again sets out the importance of using computers appropriately, for example: “the blind decisions of a computer are the worst possible way of making a critical text. Indeed, the critical text by its very nature represents everything that a computer can never achieve”; and “The computer is not a substitute for critical thinking... The computer is a means of [doing text-critical work] much more effectively and, having done it, to present the results to the public in a useful manner.”[7] This

---

is not mentioned in his paper.

[3] Ibid., 397.

[4] Ibid., 395.

[5] Ibid., 397.

[6] Ibid., 402.

[7] Parker, ‘Making Editions of the New Testament Today’, 349; 361.

more positive quote is representative of Parker's viewpoint, and, discussing the CBGM, he argues that it allows editors to be more "in control of their material, and more self-critical, than has ever been possible before."[8] He describes the CBGM as essential for editors "in controlling and monitoring their own editorial activity."[9]

Also in 2008, Cambridge published what is described on the back cover as the "first major English-language introduction to the earliest manuscripts of the New Testament to appear for over forty years": Parker's *An Introduction to the New Testament Manuscripts and their Texts.*[10] This comprehensive introduction contains a section charting "Two Hundred Years of Textual Criticism" in which Parker takes the reader from Lachmann through to the CBGM.[11] Parker highlights some problems from twentieth century textual criticism (for example methods that ignored all singular readings, or those too wedded to traditional text types). He also outlines three major problems that have made it impossible to apply Lachmannian stemmatics to the New Testament and then goes on to compare them with those faced by geneticists:

1. Contamination in a stemma of manuscripts is like genetic cross-mutation
2. Coincidental emergence of readings is like independent mutations
3. Later-reverted variant readings are like a mutation undoing another mutation's change.

These similarities have meant that phylogenetic methods and software developed by geneticists can be adapted and used by textual critics. Parker then argues that the CBGM overcomes all three problems (contamination, coincidental emergence and reversion).

Parker highlights Mink's idea of creating local stemmata for each variant unit,

---

[8] Ibid., 357.

[9] Ibid.

[10] Parker, *An Introduction to New Testament Manuscripts and Their Texts*, back cover.

[11] Ibid., 159ff.

describing it as “brilliant”.[12] He briefly outlines the CBGM’s steps and makes an interesting aside: “in a tradition with none of the three problems which have dogged traditional stemmatology, every local stemma should be identical.”[13] However, these three problems are widespread in the New Testament, and each local stemma must therefore be considered individually. This is, perhaps, one of the best ways to explain the size of the problem facing textual critics. Even though the CBGM is seen as complicated, none of the literature, from the INTF or ITSEE, quite prepares the reader for the magnitude of the task of actually applying the CBGM to even a small text, which will become clear in Chapter 3.

Parker repeats an idea often found in the CBGM’s primary literature, namely that the CBGM is not wedded to any particular text-critical criteria and these local stemmatic choices “may be made on any grounds.”[14] This widely made point, however, is only true within certain constraints, as will be shown below. For example Gurry has pointed out that it is not suitable for use by thoroughgoing eclectics, and unnecessary for those favouring Byzantine priority.[15]

Hugh Houghton wrote a paper outlining ‘Recent Developments in New Testament Textual Criticism’ in 2011.[16] When discussing the ECM, he discusses Wasserman’s edition of Jude compared to that of the ECM: Wasserman’s edition is based on all 560 continuous-text Greek manuscripts, compared to the ECM’s 140, and “brought to light numerous new readings as well as further support for poorly-attested variants.”[17] See Chapter 7 for an experiment including all the extant manuscripts of John 18 in the CBGM. Houghton

---

[12] Ibid., 169.

[13] Ibid., 170.

[14] Ibid., 169; See, for example, Mink, ‘Contamination, Coherence, and Coincidence in Textual Transmission’, 142.

[15] See Gurry, *A Critical Examination of the Coherence-Based Genealogical Method*, 101ff.

[16] In 2017 Houghton succeeded Parker as director of ITSEE.

[17] Houghton, ‘Recent Developments in New Testament Textual Criticism’, 2; See Wasserman, *The Epistle of Jude*.

highlights a key point for the application of the CBGM: at first, local stemmata must be made for those cases where there is no problem. Only with that body of evidence can more difficult cases be tackled. Houghton, too, argues that the editor is in control of the method, and not vice versa. He also argues that the conclusions drawn from the ECM and CBGM are not only making traditional text types redundant, but also other relatively recent methods such as the Claremont Profile Method.

In 2012 Parker's 2011 Lyell Lecture series was published by Oxford as *Textual Scholarship and the Making of the New Testament.* These lectures are built upon the idea that "the scribe has been forgotten, and as a result we have come to misunderstand what the [New Testament] is."[18] In the third lecture (entitled "Understanding How Manuscripts Are Related") Parker describes the CBGM (referring to it as the "Münster Method").[19] He again shows that old theories built on the use of singular readings have been shown to be flawed by such modern methods. He makes an interesting point about texts as dealt with by the CBGM: a text leading to contamination (for example) no longer has to be found in a manuscript, but could equally well exist in a scribe's memory. As such a written text and a memorised one can both be considered as texts to be analysed when considering a tradition's transmission.[20]

In the fourth lecture he points out that the ECM is still choosing manuscripts using test passages. This fact sheds light on the claim by Wachtel et al. that the CBGM's results are "based on an evaluation of all variants in all manuscripts *that are judged to be useful* for editing the text and reconstructing its history."[21] Test passages are the mechanism, it may be inferred, by which a manuscript's usefulness is judged. This is pragmatic, yet feels unsatisfactory. Consider Wasserman's experiment described above using all 560 manuscripts

---

[18] Parker, *Textual Scholarship and the Making of the New Testament*, 2.
[19] Ibid., 84ff.
[20] See ibid., 97f.
[21] Spencer, Wachtel, and Howe, 'The Greek Vorlage of the Syra Harclensis', para. 14 (my emphasis).

of Jude, and the new readings they contained. It has to be hoped that, in time, it will be possible to include all readings from all witnesses at the start of the process, before excluding the many whose readings are all found elsewhere.

In 2013 Parker presented a paper at the Byzantine Studies Symposium at the University of North Carolina at Greensboro, entitled 'New Testament Textual Traditions in Byzantium.' The conference proceedings were published in 2016.[22] In it he discusses the Byzantine text, or rather the text as it was used and developed in the Byzantine empire, and the methods in use today in wider textual scholarship. He (again) points out that text types have had their day: "the vague concept of incompatible text-types has been replaced by a methodology [i.e. the CBGM] that reconstructs the textual flow of the tradition... [this reveals] that the tradition is as broad as it is long: there are as many differences in the catholic letters between the two fourth-century codices Vaticanus and Sinaiticus as there are between the reconstructed critical text and a typical Byzantine witness."[23] Here Parker's implication is that the CBGM has been accepted, text-types are gone, and the CBGM is a standard tool for modern textual criticism.

Houghton and Smith's 2016 paper 'Digital Editing and the Greek New Testament' describes the process by which the ECM is being made, and includes a brief description of the CBGM.[24] This paper offers insights into the many steps involved in creating a critical edition using digital means, from selecting manuscripts, through transcription, regularisation and collation, to printing the critical text and apparatus. They also highlight that the CBGM has

---

[22] See Parker, 'New Testament Textual Traditions in Byzantium'.

[23] Ibid., 28.

[24] See Houghton and Smith, 'Digital Editing and the Greek New Testament', 121–22.

"sounded the death knell" for geographical text-types.[25]

Other Authors

> *"Problems arise here in abundance... We are introduced to a novel theory of stemmata studies soon to overwhelm most areas of New Testament critical study: The Coherence-Based Genealogical Method."*[26]

The CBGM has been met with varying levels of concern and even alarm by scholars more distant from the INTF or ITSEE. The quote above is from Sutton's review in 2014 of the NA28, which includes extended remarks on the CBGM. His review is aimed at non-text-critical readers, and so is a useful window into the CBGM's wider reception. He suggests that using stemmata (at all) will lessen the time given to studying the manuscripts themselves – which is a useful reminder for text critics although risks throwing the baby out with the bath water. He goes on to say "The methods in practice now among New Testament critics lag behind other fields of ancient studies by about seventy-five years" - although he neglects to explain how he has arrived at this conclusion.[27] He then (with no irony) argues that "the same or better results could have been derived from theories used by editors of previous editions [of the NA]".[28] This is generally, however, a thoughtful review coming down firmly against the CBGM, perhaps more because the CBGM is not sufficiently explained in the primary literature than because the reviewer has found real flaws with it. Consider, for example, what Sutton says when discussing the NA28 text: "Is there an article or essay which provides a sensible reason behind the alteration of Ja. 2: 3? I am unable to appreciate the logic of it all."[29]

---

[25] Ibid., 122.

[26] Sutton, 'CAVEAT EMPTOR. Remarks on the Nestle/Aland Novum Testamentum Graece', 10.

[27] Ibid., 11. For information about why modern textual scholars are doing the things they do, see, for example, Parker, 'New Testament Textual Traditions in Byzantium', 27f.

[28] Sutton, 'CAVEAT EMPTOR. Remarks on the Nestle/Aland Novum Testamentum Graece', 11.

[29] Ibid., 12.

Other Authors

We will now proceed chronologically through publications from other authors. In 2006 Paul Foster wrote a report for JSNT on a conference in Edinburgh on recent developments in textual criticism. The sessions were led by Holger Strutwolf and Klaus Wachtel, beginning with a session on the CBGM by Wachtel. With the exception of Foster's report, it would appear that these sessions were not published and therefore are not included in the primary literature above. Foster's paper must be considered as his interpretation of the sessions rather than Strutwolf's or Wachtel's. Oddly, Foster translates *Ausgangstext* as "base text" (rather than "initial text") although his explanation seems sound.[30] He reports that Wachtel seemed confident that the *Ausgangstext* and original text are virtually identical. Assuming that Wachtel's explanation is in keeping with his argument in other places, Foster misinterprets Wachtel as saying that missing manuscripts and contamination mean no global stemma can be produced. While this is true for a stemma of manuscripts, the CBGM *does* claim to be able to produce a global stemma of states of text. Foster's description of pre-genealogical coherence talks of "grouping manuscripts together" and "coherence between a subset of manuscripts" - again suggesting that he may have misunderstood the method.[31] This is not about groups or subsets, but about coherence relationships between pairs of manuscripts intended to find which may be potential ancestor or descendant. Further confusion follows, for example "On pre-genealogical grounds it may have already been determined that text A is prior to the state of text represented by B. In this case A is viewed as the ancestor for B."[32] Pre-genealogical coherence is *undirected* and so Foster's statement cannot be true. Interestingly, Foster reports Wachtel (once again) as being "keen to emphasize" that the CBGM is not a "black box".[33] The second session was a plea by Strutwolf to abandon the idea

---

[30] Foster, 'Recent Developments and Future Directions in New Testament Textual Criticism: Report on a Conference at the University of Edinburgh', 229.

[31] Ibid., 230.

[32] Ibid., 231.

[33] Ibid.

of local text-types, using examples generated by the CBGM to justify his argument. There followed an introduction to the digital NA28 and a question and answer session.

In 2009, Stephen Carlson (Duke University, North Carolina) posted to his blog a short introduction to the CBGM that he had produced for a seminar. He states that the CBGM was made to counter the inherent circular reasoning that "witnesses are good because of their good readings, but readings are good because of their good witnesses"[34]. He explains the CBGM well (Wachtel replied to the post with a comment calling it an "excellent summary of the CBGM"), including some key points often overlooked – for example that uncontroversial local stemmata are drawn up first thus creating a pool of trustworthy genealogical coherence data.[35] Concluding, he describes the CBGM as "more rigorous and precise about evaluating external evidence than the heuristic rules of thumb currently employed in reasoned eclecticism."[36] He goes on to point out that the CBGM has yet to be applied much outside of the Catholic letters and that "its operation is still poorly understood outside of Muenster... It is still very much an unproven method; time will have to tell as more people gain experience with it."[37]

In 2010, Peter Head of Tyndale House, Cambridge published an article in the Tyndale Bulletin, entitled: 'Editio critica maior: an introduction and assessment.' In this paper he gives a useful overview of the ECM, followed by a discussion of some of its more controversial features and textual decisions. He discusses the CBGM in principle but does not engage with the methodology *per se* due to its complexity. He is generally positive about the ECM, but more circumspect about the way the CBGM is sometimes used to support traditionally

---

[34] Carlson, 'Coherence-Based Genealogical Method', para. 3.

[35] See http://hypotyposeis.org/weblog/2009/12/coherence-based-genealogical-method.html#comment-186 [Accessed 6th March 2018] for Wachtel's comment replying to Carlson, 'Coherence-Based Genealogical Method'.

[36] Ibid., para. 11.

[37] Ibid.

unsupported textual decisions.[38] Perhaps crucially, Head seems unconvinced that a CBGM-only approach will yield a better text than that of, for example, NA27 where a committee made up of members with different approaches made the decisions.

Also in 2010, Larry Hurtado visited the INTF as part of a funding review. After the "informative" visit he states that "[It] now seems to me that the basic aim and the fundamental concept are not that different from what Hort proposed and offered (1881). But computing technology now makes it possible to take a MUCH larger body of manuscript data into account."[39] This highlights two simple yet important points about the CBGM: first it is not a major step away from traditional textual criticism (as some seem to think). Secondly – and this is useful as an explanation of the CBGM's somewhat negative reception in wider scholarship – the primary literature does not explain the method very accessibly when compared to a face-to-face explanation from the INTF team.

In 2012 Carlson completed his PhD at Duke University, North Carolina. His thesis is entitled 'The Text of Galatians and Its History' and documents his use of cladistic methods to create a stemma for Galatians. Carlson's thesis was published in 2015 under the same title and only "lightly revised".[40] The section on the CBGM is clearly based on his 2009 work and it is, perhaps, interesting to consider some of the changes. In his thesis in 2012 he added a comment that "[the CBGM's] practitioners see this method as an advance over the use of text-type[s] and categories that dominated the twentieth century."[41] He also adds a footnote saying that the CBGM is "incompatible with a purely documentary approach, such as the majority-text method."[42] This seems in contradiction with the INTF's assertion that the CBGM is

---

[38] See "Controversy" above,

[39] Hurtado, 'Muenster', para. 1.

[40] Carlson, *The Text of Galatians (Book)*, VII.

[41] Carlson, 'The Text of Galatians (Thesis)', 54.

[42] Ibid., 55.

methodologically agnostic (a meta-method).[43] He includes a completely new paragraph here showing the differences between stemmata produced by the CBGM and classical stemmata, for example the lack of hyparchetypes. He appears to confuse textual flow diagrams with the global stemma (a confusion that seems to happen a lot in this field). His concluding paragraph has been extensively changed with the removal of any concerns about lack of widespread testing and instead a comment that the CBGM's initial text "may reflect the authorial text better [than using traditional methods], and this benefits those who are interested in the text as it was composed by an author within its historical context."[44] He then states that because the CBGM creates a stemma of states of text and not of manuscripts then this is *not* a reconstruction of the tradition's history and so of less interest to historians. Indeed, Carlson states that "the textual history is scrapped in favor of a highly abstract diagram of textual flows that does not convey the textual state of any lost exemplar other than the initial text."[45] This argument seems to be lacking in nuance: surely some textual history could be learned from the textual flow diagrams (albeit by applying some detective methods).

In 2013 Dirk Jongkind wrote four book reviews in JSNT, one of which was of *The Textual History of the Greek New Testament* (eds. Wachtel and Holmes). He describes the book as "the reference point for further evaluation of the CBGM".[46] Then at SBL 2013 in Baltimore he presented a paper entitled 'On the Weighing and Counting of Variants: The Coherence Based Genealogical Method, Potential Ancestors, and Statistical Significance.' This paper is significant as it is the first to offer detailed criticisms of the CBGM. Jongkind, as is now typical when scholars outside the INTF discuss the CBGM, begins by saying "[the

---

[43] See, for example, Mink, 'Contamination, Coherence, and Coincidence in Textual Transmission', 142. Note that Gurry agrees with Carlson's view, see Gurry, *A Critical Examination of the Coherence-Based Genealogical Method*, 101ff.

[44] Carlson, 'The Text of Galatians (Thesis)', 58.

[45] Ibid., 72.

[46] Jongkind, 'Review of "The Textual History of the Greek New Testament", Wachtel, Klaus and Michael Holmes Eds.'

CBGM] is not easy to understand and, even when understood, is not easily replicable in the classroom or department. Therefore the following contribution is offered with the caveat that the author may be mistaken in the two or three question marks that are put next to the CBGM."[47] He then goes on to set out two main objections: The CBGM attempts a more detailed analysis than the data allows; and the CBGM "will misrepresent some important historical scenarios."[48] Before these, however, he raises the question of how pre-genealogical coherence is calculated (the number of variant units in which the two witnesses agree divided by the number of variant units where they are both extant – as a percentage). He points out this means that the pre-genealogical coherence between any two manuscripts is contingent on the extant tradition – and, although he does not say so, how it has been collated. This will be discussed further in Chapter 9 and compared to more objective measures.

Jongkind compares two witnesses (33 and 2344) in 1 Peter and makes his own analysis. He makes only a few different decisions from the CBGM's authors, but this results in a reversal of the predominant textual flow. He points out that if the boundaries are moved between variant units then the final counts of difference or similarity change. He gives an example of two variant units that he argues should be counted as only one. Therefore, he concludes, the data is subjective and not solid enough to provide a base for quantitative analysis. See Chapter 3 for an example of the impact of changing the collation.

He also challenges the CBGM's use of only original firsthand readings. While acknowledging that this is methodologically consistent, he argues that this "only introduces the appearance of objectivity by suspending judgement and knowledge of the manuscript."[49] Here Jongkind is somewhat mistaken, as the CBGM actually uses the text as it left the scribe,

---

[47] Jongkind, 'On the Weighing and Counting of Variants', 1.

[48] Ibid., 2.

[49] Ibid., 6.

including any firsthand corrections. This is not made clear in the documentation, but has been personally clarified to me by Klaus Wachtel when he reviewed an early draft of Chapter 3 of this thesis.[50] This is totally at odds with Mink's position: "The reason for disregarding corrections generally is that the basic data used for the ECM do not differentiate between immediate correctors of the first hand… and other corrections."[51] It seems likely that Wachtel's ongoing use of the CBGM has moved on from Mink's earlier position. The inclusion or exclusion of correctors is discussed further in Chapters 3 and 9. It is important to note that, as of 2017, Parker (when using the CBGM for the ECM of John's Gospel) takes this further and includes scriptorium corrections where possible, and this approach has also now been adopted by the INTF.[52]

Mink states that "A higher proportion of prior variants indicates an older text."[53] Jongkind challenges this argument (citing Eldon Epp's analysis) specifically for cases where the number of prior and posterior readings are close.[54] Jongkind also quotes Mink as naming the difference between prior and posterior readings the "stability value", but Jongkind says "the implications of this instability are not discussed".[55] Mink indeed talks about strength and stability of textual flow in a number of papers but it is unclear how such strength and stability values affect the CBGM's output.[56] As an example, Jongkind argues, using statistical methods, that the CBGM's textual flow diagram for 1 John 4:4/6a contains eight (out of seventy-three) suspect relationships, many of which occupy key locations in the textual flow

---

[50] Personal email from Klaus Wachtel, 8 December 2014.

[51] See Gurry, *A Critical Examination of the Coherence-Based Genealogical Method*, 202. Gurry quotes here from Mink's unpublished response paper from the meeting in January 2014 where the INTF invited Dirk Jongkind and other scholars (including Gurry) to discuss Jongkind's recent SBL paper criticising the CBGM.

[52] That scriptorium corrections are now being included by the INTF was confirmed to me in a personal conversation with Klaus Wachtel in January 2018.

[53] Mink, 'The Coherence-Based Genealogical Method (CBGM) – Introductory Presentation, 1.0', 110.

[54] See Epp, 'Textual Clusters'.

[55] Mink, 'Problems of a Highly Contaminated Tradition', 56; Jongkind, 'On the Weighing and Counting of Variants', 8.

[56] For their definitions see Mink, 'Problems of a Highly Contaminated Tradition', 57.

diagram. He suggests that his method of evaluating the strength of textual flow provides a stronger basis for concluding that one text is older than another. For further discussion of strength and stability of textual flow see Chapter 9. Discussing the consequences of his argument for the creation of textual flow diagrams and the global stemma and the conclusions on which they are built, Jongkind argues that "these conclusions cannot always be sustained by the nature of the actual data."[57]

Jongkind's second (and more serious) major objection is that the CBGM will "misrepresent some historical scenarios" where a single text exerts influence at multiple stages in a stemma.[58] He gives an artificial example of five manuscripts A-1-2-3 and R, where R is a copy of A and drip-feeds its readings into 1, 2 and 3 as they are copied from one another. In this example the CBGM concludes that the stemma is a linear A-1-2-3-R – and Jongkind concludes that the CBGM is at fault. It does seem, however, that the CBGM has made a perfectly reasonable decision based on the evidence at hand – unless there is also some non-textual evidence available showing that R had this persistent influence. Jongkind argues that this kind of influence "must have happened frequently" in reality and he therefore casts doubt on the CBGM's ability to chart "the correct development of texts" in the New Testament.[59] Jongkind claims that the CBGM decision to work with texts and not manuscripts invalidates the CBGM's claim as representing the history of transmission of the New Testament: "The text flow becomes the object of study, which is in essence an artificial continuum of witnesses in which they are positioned on the basis of their proportion of prior readings."[60]

Nevertheless, his final conclusion in this paper is important, given that it is a simple

---

[57] Jongkind, 'On the Weighing and Counting of Variants', 13.

[58] Ibid.

[59] Ibid., 15 and 16.

[60] Ibid., 16.

fact that could counter much of the anti-CBGM sentiment that abounds: “[In] the end, the Editio Critica Maior of the Catholic Epistles is a text produced by scholars; not by a method, but with the help of a method.”[61]

In 2013 Elliott wrote a review of the ECM for JTS 64.1. He points out the comprehensiveness of the ECM by the example that it has 750 variant units in James, whereas the United Bible Society’s edition has only twenty-three. He names the CBGM as being the methodology behind the primary line of text and calls it “home-spun”.[62] He welcomes the publication of the second edition and its related “Genealogical Queries” online programs, saying “we have our chance now to assess [the CBGM’s] effectiveness as a text critic’s tool in practice.”[63] Unfortunately, he takes hold of the list of the CBGM’s highly ranked witnesses in the ECM (those with A as their first-ranked potential ancestor) and seems to treat them as a new set of “good” manuscripts, questioning why readings were chosen that go against the majority of these witnesses.[64] This points to a misunderstanding, or perhaps merely a wishful simplification of the CBGM’s concepts.

Also in JTS 64.1, Elliott wrote an article on NA28 entitled ‘A new edition of Nestle–Aland, Greek New Testament.’ This paper includes a section on the Catholic Epistles (the only section with textual differences compared to NA27). Discussing the CBGM’s use, he argues that the ECM’s editors treat manuscripts as “mere bearers of the tradition” - and while he is pointing out that a later manuscript may carry an early text his language suggests that he thinks the editors have discarded much useful data by taking this approach.[65] Again he says “particular manuscript support is of no significance when CBGM is applied”, but here surely Elliott is mistaken since local stemmata are created using standard text-critical methods *plus*

---

[61] Ibid., 18.
[62] Elliott, ‘Review of “Novum Testamentum Graecum: Editio Critica Maior”, INTF (ed.)’, 636.
[63] Ibid., 637.
[64] See ibid., 638.
[65] Elliott, ‘A New Edition of Nestle-Aland, Greek New Testament’, 57.

the new external criteria of coherence.[66] He poses a question which implies that he believes this method is flawed: "It remains to be seen what the editors do when at some stage they take a manuscript's date into account. What changes will they make if they see that there is a consistent preference for the textual flow between two manuscripts running from x to y if manuscript x is younger than manuscript y?" He pleads for the publication of the promised companion textual commentary to NA28.[67]

Again in 2013 a second edition was published of *The Text of the New Testament in Contemporary Research: Essays on the Status Quaestionis*, edited by Bart Ehrman and Michael Holmes. This revised edition of the 1995 volume contains updates to, and sometimes replacements of, the original articles. It hopes to "provide informed discussions of the current state of knowledge with respect to a wide range of important text-critical topics..." and contains twenty-eight articles in over 800 pages.[68] Thomas Geer contributes chapter 19: 'Analyzing and Categorizing New Testament Greek Manuscripts.' He begins by discussing Colwell and his legacy, for example the Claremont Profile Method. He then moves on to talk about new avenues, including the CBGM, where he states (surprisingly) that "[the] whole process is admittedly driven by the concern to eliminate MSS of the late Byzantine tradition from consideration in publishing a major critical edition of the Greek NT."[69] He criticises the continued use of test passages for selecting manuscripts for inclusion. Again, somewhat surprisingly, he claims that "The [CBGM] is a splendid tool to assist in mapping the history of the transmission of the text. Because it necessarily works from a limited number of variant readings in order to keep the stemmata manageable, it may nevertheless not be the best instrument for estimating the degree of relationship among MSS."[70]

---

[66] Ibid., 59.

[67] Ibid., 57.

[68] Ehrman and Holmes, *The Text of the New Testament in Contemporary Research*, ix.

[69] Racine and Geer, 'Analyzing and Categorizing New Testament Greek Manuscripts', 505.

[70] Ibid., 507.

Chapter 20 of the same book is by Eldon Jay Epp: ‘Textual Clusters:Their Past and Future in New Testament Textual Criticism.’ He briefly considers the CBGM and its impact on textual clusters / text types and in this critique he raises a concern that the CBGM is all about counting and not weighing variants. Epp, however, goes on to say that “apparently, however, this is not the case.”[71] Then, beginning with the expected caveat “Many will admit that the functioning of the CBGM is not always easily grasped...” he goes on to give a concise and clear (if very brief) description of the method, describing the creation of local stemmata as “simply NT textual criticism as all of us practise it.”[72]

Chapter 21 is by Tommy Wasserman: ‘Criteria for Evaluating Readings in New Testament Textual Criticism.’ Again we find the CBGM has a bearing on the topic, and Wasserman includes examples from the Catholic letters. He argues that “The main advantage of the CBGM is that it allows scholars with their different emphases on criteria to correlate their textual decisions in a single variation unit with the decisions throughout a whole book or corpus.”[73] He offers three potential criticisms of the CBGM: first, he suggests that in other New Testament books in which there is a strong “Western” group of witnesses, the tradition may not be sufficiently coherent for the CBGM to produce “valid results”.[74] Secondly, discussing pre-genealogical coherence, he states that “scholars may have different opinions about the assessment of the character of the textual witnesses in the initial stages.”[75] Finally, he joins the argument that the use of test passages to exclude witnesses could present a problem.

Stanley Porter published a monograph in 2013, *How We Got the New Testament*. In a book divided into three similar-length sections, Porter describes the New Testament’s text, its

---

[71] Epp, ‘Textual Clusters’, 550.
[72] Ibid.
[73] Wasserman, ‘Criteria for Evaluating Readings in New Testament Textual Criticism’, 605.
[74] Ibid., 606.
[75] Ibid.

transmission and its translation. The book is based on the Hayward lecture series delivered by Porter in 2008 at Acadia Divinity College in Wolfville, Nova Scotia. He addresses the question of the goal of textual criticism, the history of the text and hot topics such as those raised by Bart Ehrman's recent theories. He briefly discusses the CBGM a number of times, stating problems he has identified with the method. Interestingly, he refers to it as "scientific" but complains that it is explicitly based on assumptions.[76] When first mentioning the CBGM, he takes the time to point out that it is "not the same as the Alands' local-genealogical method".[77] He is correct about that, of course, but it is not clear why he feels the need to offer such a clarification – although this perhaps highlights a problem with Porter's critique of the CBGM: he has preconceived notions about the terminology, and Mink uses it differently. Elliott, reviewing this book, says "The new-fangled coherence-based genealogical methodology is viewed cautiously" correctly indicating Porter's opinion of the CBGM and offering an intriguing insight into Elliott's.[78] After complaining that the CBGM needs significant resources (databases and the like) to be useful and is therefore, he implies, currently not usable for external scholars, Porter states "There is also the difficulty with what is meant by "coherence" (is it merely statistical?)..."[79] He is sure this is a problem with the CBGM, and the reason is exposed, perhaps, while he is discussing text-types: "The issue is perhaps not coherence, which indicates the cognitive process of appropriation, so much as cohesion, the textual indices that create a text and hold it together."[80] In other words, Porter has a particular understanding of "coherence" – and it is not the CBGM's definition (i.e. the level of agreement between two witnesses expressed as a percentage). Porter also dislikes the ambiguity present in Mink's definition of the "initial text". He cites Mink as saying the

---

[76] Porter, *How We Got the New Testament*, 32 (footnote 74). He uses quotation marks around the word "scientific" but does not explicitly explain why.

[77] Ibid., 32.

[78] Elliott, 'Review of "How We Got the New Testament: Text, Transmission, Translation." Porter, S.E.', 672.

[79] Porter, *How We Got the New Testament*, 59.

[80] Ibid., 60.

"initial text" could be authorial, or a redactor's text, or the archetype of the tradition. Porter says "[this] is part of the ambiguity of the method" but in fact Mink, Wachtel, et al. have left little room for ambiguity.[81] To be fair to Porter, some key material has been published since he delivered the lectures – although it was available before his book was published. His final criticism of the CBGM is that it does not accept that different variants may have different significance. He fails to explain this criticism, simply stating it in a footnote and hinting at it in parentheses in his text, saying: "[the CBGM] allows all variants in all extant manuscripts to be recorded (without differentiation of variants)..."[82] This is probably linked to his concerns about the statistical nature of coherence in the CBGM – and there is a valid question surrounding the allocation of the status "(best) potential ancestor" based on a percentage.

Also in 2013, Ryan Wettlaufer published a book entitled *No Longer Written: The Use of Conjectural Emendation in the Restoration of the Text of the New Testament, the Epistle of James As a Case Study*. In an excursus he uses the CBGM to investigate the extent of manuscript loss in the New Testament – citing it as a mechanism for calculating this more empirically than has hitherto been possible. Interestingly, in his introduction, he states that the CBGM's emphasis on texts and not manuscripts is not new, although it has never been consistently applied before. He cites Metzger as arguing that the date of the text of a manuscript is far more important than the date of the manuscript itself. Indeed, Metzger, in 1964, went on to say that a minuscule (say 1739) could bear a more important (older?) text than some majuscules – and this is strikingly similar to the reasoning given by the CBGM's authors for the inclusion of second-millennium manuscripts in the ECM.[83] Houghton and Smith provide further insight into this by saying: "Given the trend towards uniformity in the

---

[81] Ibid., 33; See, for example, Mink, 'Editing and Genealogical Studies: The New Testament', 52; Wachtel, 'The Coherence Method and History', 2.

[82] Porter, *How We Got the New Testament*, 33 (continued footnote 74); ibid., 32. Concerning the significance of different variants see the discussion on Alexanderson below,

[83] See Metzger, *The Text of the New Testament. Its Transmission, Corruption, and Restoration*, 209.

textual tradition of the Greek New Testament, witnesses which deviate from the norm are more likely to preserve earlier forms of text, regardless of the age of the document."[84]

In 2014 Bengt Alexanderson published a book entitled *Problems in the New Testament: Old Manuscripts and Papyri, the New Genealogical Method (CBGM) and the Editio Critica Maior (ECM)*. In this book he sets out his position on textual criticism and it is clear that he stands quite far from Münster: he does not agree with Barbara Aland's opinions on old papyri and their text, and he lambastes the CBGM. For example in his introduction he describes his third chapter as: "... a fairly thorough discussion of the Coherence-Based Genealogical Method (CBGM) which has been fundamental to establishing the text of the Editio Critica Maior of the New Testament. *The method is found to be of no or little value.* This is serious, considering the fact that this very ambitious new edition will in all probability have an impact on future work on the New Testament and generally on editing classical and medieval texts."[85] He is correct, that a seriously faulty CBGM could lead to a ECM of greatly diminished value. Sadly his book is replete with misunderstandings – sometimes because he disagrees with the terminology and sometimes because he has simply missed an important point. There is not space here to engage with all the issues in this book, but several key ones will be addressed below.

It is perhaps fair, however, to begin with a positive point. His criticism of the iterative process by which earlier decisions are revisited in the light of the emerging textual flow is not without merit.[86] That the feedback loop could become forced and self-fulfilling is a genuine concern. See Chapter 9 for a fuller discussion of this.

He explains his comment about "little or no value" by saying "One reason is that it is

---

[84] Houghton and Smith, 'Digital Editing and the Greek New Testament', 112.

[85] Alexanderson, 'Problems in the New Testament: Old Manuscripts and Papyri, the New Genealogical Method (CBGM) and the Editio Critica Maior (ECM)', 3 (my emphasis).

[86] Ibid., 82.

footed on a faulty use of statistics, giving the same weight to accidental changes as to deliberate ones, another that the ideas of how one reading develops into another are often highly improbable, a third that the method consistently undervalues interpolation."[87] He frequently picks up the point about accidental changes, for example "I prefer not to pay attention to what I regard as such variants as show no relationships because they may appear anywhere unintentionally... My point is that what may be unintentional should be considered unintentional."[88] He does admit that "Taking a reading as intentional or unintentional is of course a matter of personal judgement and personal experience of texts."[89]

Similarly, when discussing contamination, he states "We shall see many times that the CBGM prefers coincidence (multiple emergence), in my opinion too much so."[90] Now, the CBGM uses all non-regularised variation to rule on coincidence (multiple emergence) but Alexanderson thinks that unintentional changes (i.e. coincidental ones) should be ignored – and so he takes an orthogonal approach. He cares about intent, to the extent where only intentional scribal changes (i.e. not mistakes) matter to him, and by implication they are the only ones that can carry significant genealogical information.[91] An important counter to this argument comes from Andrews, documenting an unrelated experiment using artificial texts. She concludes that philologists are not good at determining which variants are genealogically significant, and "that 'insignificant' variation is really not that insignificant at all."[92] See Chapter 9 for a discussion of applying weights to variant units.

Alexanderson's brief discussion of prior and posterior variants is superficial and seems to indicate that he does not understand that the CBGM includes methods for dealing

---

[87] Ibid., 8.
[88] Ibid., 45.
[89] Ibid., 47.
[90] Ibid., 62.
[91] See ibid., 62–64.
[92] Andrews, 'Analysis of Variation Significance in Artificial Traditions Using Stemmaweb', 538.

with contamination, coincidental agreement etc. In fact, since this is one of the CBGM's big claims (that it can cope with contamination better than other approaches) Alexanderson seems to have missed something significant here.[93]

In contrast to Wettlaufer, Alexanderson appears to have completely missed the importance of texts versus manuscripts. Many of his arguments are built on the idea that the stemmata, textual flow diagrams etc. created by the CBGM are intended to show a real manuscript to manuscript genealogy. They are not – they are intended to show a likely textual flow. For example, "The fundamental idea of the CBGM concerning ancestors and descendants is on the whole misleading. To simplify somewhat: Manuscript A may to 55% have a more original text than manuscript B and to 45% a less original one. To the CBGM this means that A is the ancestor of B, that there is flow from A to B, but the same manuscript is in fact both ancestor and descendant, both father and son."[94] Alexanderson has evidently not understood the CBGM's approach to combating contamination at all.

Similarly, when discussing the CBGM's use of statistics, he says "What is a 'reading', a 'variant', a 'passage', a 'place of variation'? From a practical point of view, I think that these terms all mean the same thing, and I think they do so to the CBGM."[95] However, in reality, there is a distinct difference between, for example, a variant and a reading. An orthographic reading or a nonsense reading will not be treated as a variant, and will therefore be excluded from the CBGM's data.

He dislikes the idea of applying phylogenetics to texts (by extension from this comment): "A very dangerous simile is that of the sequences of variants being compared to DNA chains. This sounds very scientific, but all mammals have two biological parents, no

---

[93] See Alexanderson, 'Problems in the New Testament: Old Manuscripts and Papyri, the New Genealogical Method (CBGM) and the Editio Critica Maior (ECM)', 61–62.

[94] Ibid., 69.

[95] Ibid., 64.

more, no less, whereas the document may have many parents, none of them biological."[96] Parker, in his review of Alexanderson's book, points out that this is "a false analogy, since the better comparison than with mammals is with the development of the 'flu' virus."[97] Parker also highlights many of Alexanderson's misunderstandings regarding the CBGM.

Interestingly, Alexanderson (while talking about the ECM, but seemingly commenting on all modern textual criticism and methods) says: "The oldest witnesses are unreliable, the method is faulty. We are on a slippery ground indeed. Our knowledge of the Greek language is unsatisfactory; so is our knowledge of the life and traditions of early Christian congregations; we do not know much about copying in antiquity and in the Middle Ages. What we have is un[s]atisfactory knowledge and a highly fallible judgement."[98] It seems he doubts much of modern textual criticism's ability to do its job – and I would expect that many textual scholars would take serious issue with his comments (in this quote and elsewhere in the book). When considering the ECM, he states that "It is in fact astonishing that using a poor method, the result is as good as it is".[99] It is interesting that Alexanderson does not appear to have considered that, therefore, perhaps the CBGM is not actually as flawed as he thinks. Alexanderson states that "I do not think that the CBGM can be rescued."[100] In contrast, Parker ends his review by saying: "[The CBGM] certainly deserves to be properly tried and tested and this review concludes with a challenge to those who condemn it: try using it to make a text, and only then decide whether it works or not."[101]

2014 also saw publication of *Texts and Traditions. Essays in Honour of J. Keith*

---

[96] Ibid., 70.

[97] Parker, 'Review of Alexanderson, Problems in the New Testament: Old Manuscripts and Papyri, the New Genealogical Method (CBGM) and the Editio Critica Maior (ECM)', 2.

[98] Alexanderson, 'Problems in the New Testament: Old Manuscripts and Papyri, the New Genealogical Method (CBGM) and the Editio Critica Maior (ECM)', 9.

[99] Ibid., 8.

[100] Ibid., 141.

[101] Parker, 'Review of Alexanderson, Problems in the New Testament: Old Manuscripts and Papyri, the New Genealogical Method (CBGM) and the Editio Critica Maior (ECM)', 4.

Elliott, edited by Doble and Kloha. The ECM and CBGM make only one notable appearance in this *Festschrift*, in Epp's essay 'In the Beginning was the New Testament Text, but Which Text?' Epp discusses at length the concept and terms "Ausgangstext" and "initial text". He points out that the text had no single beginning but it is instead "highly multiple and multileveled", and notes Elliott's proposal of the plural term *Ausgangstexte*.[102] An important point is that unlike the "newly minted" German term, the English term "initial text" seems to imply an equivalence with the original text and it is necessary to consider how the term will be understood by readers rather than simply to seek to define it *in vacuo*.[103] He concludes that the ambiguity of the terms makes it difficult to know how it will be understood, and instead welcomes the "teachable moments" created by the ongoing discussions.

The SBL session on the CBGM in 2014 saw, in addition to the papers from the INTF, contributions from Wasserman and Jongkind and responses from Carlson and Morrill.[104] Wasserman and Carlson's papers were published with the INTF's in the special feature entitled 'The Coherence Based Genealogical Method' in the online journal *TC*, Volume 20 (2015). Jongkind declined to have his paper published in *TC* and although it has not subsequently been published anywhere, he briefly shared it publicly in the following week with the caveat that it was not ready for publication.[105] Morrill chose not to publish his response paper as Jongkind had not published his, and it is not available.

Wasserman's paper, 'Historical and Philological Correlations and the CBGM as Applied to Mark 1:1', is an attempt to analyse Mark 1:1 using the CBGM's techniques – but

---

[102] Epp, 'In the Beginning Was the New Testament Text, but Which Text?', 60.

[103] Ibid., 68. Epp interestingly notes the similar words *Ausgangspunkt*: starting point; *Ausgangsfrage*: initial question; *Ausgangskapital*: original investment; *Ausgangsmaterial*: source material, or original material; *Ausgangssprache*: source language; *Ausgangsstellung*: initial or starting position.

[104] See Chapter 1 for the INTF's papers.

[105] See TC Editors, 'Special Feature: The Coherence-Based Genealogical Method. Editorial Introduction.' Jongkind's post (no longer available) sharing the paper states: "it needs some work, but take it as it is for now."

without the supporting CBGM data (pre-genealogical coherence data is not yet available for Mark). He has some interesting insights. He discusses counting and weighing variants and hails the definition of a variant as connective or not to be an important example of weighing.

A brief excursus into a work by Fee in 1993 is helpful here. Fee discusses types of variation and their weighing in ‘On the Types, Classification, and Presentation of Textual Variation’. He points out that orthographical variation is normally irrelevant – and indeed the CBGM is normally applied after such variants have been regularised away, along with nonsense readings (another category listed by Fee).[106] Fee defines the weighing of variants as “a classification which considers the degree (or lack of it) of the possible genetic significance of agreements in variation.”[107] A non-connective variant in which two distantly-related witnesses agree, “multiple emergence” in the CBGM’s parlance, is neatly described by Fee (before the CBGM existed) as “a non-genetic accidental agreement in variation between two MSS which are not otherwise closely related”.[108]

So Wasserman’s statement about connectivity being an important example of weighing is supported by Fee’s definitions. To expand on Fee’s definition, in light of Wasserman’s statement, we may say that agreement in a variant between two otherwise distantly-related witnesses is an example of multiple, accidental emergence if the variant is not connective; and an example of contamination if the variant is connective.

There is an important corollary to this, it seems to me, regarding the definition and re-evaluation of local stemmata for connective variants in the CBGM: such a local stemma which does not agree with the textual flow of witnesses must nevertheless be allowed to remain. Thus a genuine genetic relationship between two otherwise unrelated witnesses would

---

[106] See Fee, ‘On the Types, Classification, and Presentation of Textual Variation’, 66f.
[107] Ibid., 67.
[108] Ibid.

be encoded in the data. Yet this connectivity setting is not encoded into the data at all for future use – i.e. no numerical “weight” is assigned that can have an effect on the statistics. A connective variant has exactly the same weight, in statistical terms, as any other variant yet this local stemma is allowed to be defined contrary to the textual flow (and thus exerts its “weight” on the process). See Chapter 3, section “Textual Flow and Consistency” for an example of textual flow and optimal substemmata in a situation where the source of a reading is only found in non-ancestor witnesses.

Later Wasserman discusses singular readings, and argues that there should be an option to ignore them when using the CBGM. He correctly points out that singular readings (in particular in the case of Codex Sinaiticus, 01) “potentially distort the genealogical analysis.”[109] He argues that the CBGM should include an option to exclude singular readings. Now, it is true that including singular readings can dramatically lower the coherence percentage between two witnesses, but surely that is nevertheless a true representation of the data. Combine this with the arbitrary nature of singular readings and I am unconvinced by this idea, since a discovery of just one more witness attesting the reading makes it no longer singular. Witnesses are still being found in libraries, archaeological digs, private collections etc. with enough regularity to make this a real possibility. Drawing a line between a singular reading (attested by one witness) and a non-singular reading (attested by perhaps just two witnesses) is an arbitrary decision based simply on a fluke of history.

Wasserman notes Wachtel’s paper from SNTS 2014 saying “Wachtel reports that there are 7,638 variant passages in Acts among the included witnesses in the ECM of Acts. In the first stage, decisions were made in no less than 7,213 passages (ca. 94%) based on pre-genealogical evidence, whereas merely 425 passages (ca. 6%) were left to the second phase

---

[109] Wasserman, ‘Historical and Philological Correlations and the CBGM as Applied to Mark 1:1’, 5.

where genealogical evidence will be taken into account.”[110] This is a very welcome statistic, and implies that 94% of the local stemmata in Acts were decided using only objective statistical data and (presumably) traditional text-critical reasoning – a statistic that may help to sway those who view the CBGM with suspicion.[111] Indeed Wasserman’s final conclusion is “the traditionally accepted philological principles of textual criticism and the dominant view of the textual history of the New Testament exert considerable control in the application of the [CBGM].”

Carlson’s response to this SBL session is published alongside the other papers as ‘Comments on the Coherence-Based Genealogical Method.’ In his introduction he refers to the concern among wider scholarship regarding the CBGM, and that it “could create a priesthood within a priesthood, where only those behind the curtain are privy to its mysteries.”[112] He goes on to say that the INTF has been attempting to “lift the veil on the CBGM” and is helping others in their work relating to it.[113] This SBL session is a case in point. He comments on the CBGM’s ability to find early readings in late manuscripts, comparing it to his own work on Galatians.[114] He explains (with more clarity than perhaps the INTF) the difference of considering the text of a manuscript (as a potential ancestor) and the manuscript itself. For example a text (from a later manuscript) may have an early place in a textual flow diagram, but that does not mean that every reading in the manuscript is early.

---

[110] Wachtel’s paper is not generally available, but was a preliminary version of two papers presented at SBL 2014 (See Chapter 1, and Gäbel et al., ‘The CBGM Applied to Variants from Acts: Methodological Background’; Wachtel, ‘Constructing Local Stemmata for the ECM of Acts: Examples’.). Wasserman cites it as Klaus Wachtel, “Constructing Local Stemmata for the Editio Critica Maior of Acts” (paper presented at the 69th General Meeting of Studiorum Novi Testamenti Societas in Szeged, 8 August 2014)”. See Wasserman, ‘Historical and Philological Correlations and the CBGM as Applied to Mark 1:1’, 11.

[111] See the INTF’s papers from SBL 2014 in Chapter 1 regarding the objectivity of pre-genealogical coherence. Namely, Gäbel et al., ‘The CBGM Applied to Variants from Acts: Methodological Background’; Hüffmeier, ‘The CBGM Applied to Variants from Acts’; Wachtel, ‘Constructing Local Stemmata for the ECM of Acts: Examples’.

[112] Carlson, ‘Comments on the Coherence-Based Genealogical Method’, 1.

[113] Ibid.

[114] See Carlson, ‘The Text of Galatians (Thesis)’; Carlson, *The Text of Galatians (Book)*.

Other Authors

Carlson's wider argument is that the CBGM is primarily useful for creating an initial text, but considerably less useful for understanding the history of the text (at least without much more interpretive work).[115]

Dirk Jongkind's SBL 2014 paper 'On the Nature and Limitations of the Coherence Based Genealogical Method' should be treated with care, as he declared it not ready for publication and so may not reflect his considered opinions. It is an updated version of his paper from the previous SBL, following which he visited the INTF for an "extended discussion" of it.[116] The crux of his argument (in continuity with his 2013 paper) is that the CBGM at least implies, by its name, assumptions and terminology, that it is modelling the historical development of the text in the manuscript tradition. Yet this is not what the CBGM claims, and instead the textual flow diagrams produced do not represent manuscript history. Rather the nodes represent disembodied texts and not the manuscripts from which they came (and Jongkind acknowledges this, although argues that the CBGM's documentation fails to make this point as strongly as required). Jongkind again (following on from 2013) sets out some imagined historical scenarios that the CBGM would misrepresent, in a section entitled "Why the CBGM does not lead to an adequate hypothesis of the overall development of the text."[117] He argues from situations where a small number of changes (in the collation, or local stemmata etc.) will change the result: a decision based on weak evidence that is then used as the basis for another decision is dubious science. The strength of textual flow must be considered, and perhaps should be shown on textual flow diagrams. This idea will be explored in Chapter 9. Carlson, in his comments on the SBL session, says that "Dirk's thought experiment needs to be extended or recrafted to find scenarios where it would adversely affect

---

[115] See Carlson, 'Comments on the Coherence-Based Genealogical Method', 2.
[116] Jongkind, 'On the Nature and Limitations of the Coherence Based Genealogical Method.', 1.
[117] Ibid., 6.

the reconstruction of the initial text. After all the CBGM is about the initial text."[118] If this comment represents the wider reception of Jongkind's paper then perhaps that is why he chose not to publish it. Jongkind, however, makes an important point, which is that the CBGM does claim to create a comprehensive hypothesis of the development of the text, and is not simply interested in the initial text. He does devote a section of his paper to explaining how the CBGM does a good job of establishing the initial text.

In 2015 Wasserman published a paper 'The Coherence Based Genealogical Method as a Tool for Explaining Textual Changes in the Greek New Testament' in *Novum Testamentum*. Unusually, Wasserman is a scholar who is not set against the CBGM, and indeed one who considers it useful "not only in the reconstruction of an initial text... corresponding to the classic goal of textual criticism, but also in order to survey the subsequent development of the text."[119] Wasserman points out that the CBGM can be seen as an improvement on the "long-established local-genealogical method", and that the CBGM serves to add control and consistency to this method.[120] He warns the reader that "It is important to note that there is no philological reasoning behind a textual flow diagram as such in a given variation-unit. The particular diagram simply displays possible sources of a given variant according to the underlying database of potential ancestors and on the condition defined by the connectivity option. This option, however, can be set at different tolerance levels by the text critic on philological grounds."[121] See Chapter 7 for a further discussion of connectivity. Wasserman's warnings are doubtless written in the context of (and to attempt to counter) the widespread misunderstandings of the CBGM. He continues with examples where he uses the CBGM to

---

[118] Carlson, 'Comments on the Coherence-Based Genealogical Method', 2.

[119] Wasserman, 'The Coherence Based Genealogical Method as a Tool for Explaining Textual Changes in the Greek New Testament', 209f.

[120] Ibid., 208.

[121] Ibid., 209.

refute some of Ehrman's famous arguments concerning orthodox corruption.[122] Wasserman's usage of the CBGM is correct, I believe, and his conclusions are persuasive.

Another rare case of a book that discusses the CBGM in some detail yet seems positively inclined towards it is Lin's 2016 work *The Erotic Life of Manuscripts*. This is an engaging book contrasting the work and history of biologists with that of textual scholars. As an appendix it includes an interesting transcript of an interview in 2011 between Lin, Wachtel and Mink. This meeting took place before the publication of Wachtel and Holmes' book *The Textual History of the Greek New Testament. Changing Views in Contemporary Research* which contains, of course, Mink's important contribution 'Contamination, coherence, and coincidence in textual transmission: The Coherence-Based Genealogical Method (CBGM) as a complement and corrective to existing approaches.' Lin's conclusion contains an interesting statement: "What makes CBGM so intriguing… is that it embraces the overarching trends in biology since the seventeenth century despite its practitioners' distancing themselves from biologists and their methods."[123] Her arguments showing the CBGM's similarity to phylogenetics are well made, and the interview in her appendix shows Wachtel's views clearly enough. Jongkind's review is worth quoting here giving, as it does, both a sense of Lin's tone and an insight into Jongkind's own views of the CBGM: "she seems to have accepted the foundation myth as promulgated in the early days of the CBGM as if the method is the best new thing since sliced bread. Despite this, warmly recommended to anyone with an interest in the field."[124]

In 2017 Stanley Porter, with Andrew Pitts, published a book entitled *Fundamentals of New Testament Textual Criticism*. Instead of describing its treatment of the CBGM in detail, it

---

[122] See Ehrman, *The Orthodox Corruption of Scripture: The Effect of Early Christological Controversies on the Text of the New Testament*.

[123] Lin, *The Erotic Life of Manuscripts*, 152.

[124] Jongkind, 'Review of "The Erotic Life of Manuscripts: New Testament Textual Criticism and the Biological Sciences", Yii-Jan Lin', 118.

is sufficient to quote from Carlson's review: "The major methodological advance of this edition, the Coherence-Based Genealogical Method, is misclassified as a stemmatic method and its workings are not explained but summarily attacked in a footnote (p. 90, n. 1). Since this method is the basis for the current and future critical editions of the New Testament, students are hardly well served."[125] It is evident that Porter's opinion of the CBGM has not improved since his 2013 publication.

2017 saw the publication of two key books on the CBGM, both involving Peter Gurry. Gurry's Cambridge PhD thesis was published in October 2017 as *A Critical Examination of the Coherence-Based Genealogical Method in New Testament Textual Criticism*, in Brill's New Testament Tools, Studies and Documents series. A month later, together with Tommy Wasserman, he published *A new approach to textual criticism: an introduction to the coherence-based genealogical method*. In order to preserve the independence of this thesis I am not going to engage with these two books throughout (although occasional references will be made where factual differences are at stake), but will consider them here as a fitting end to this chapter.

Gurry's monograph sets out to provide "a sustained study of the [CBGM]" based on the only data available; namely that for the Catholic Epistles in the ECM.[126] Gurry asserts that the only people using the CBGM for the New Testament are the ECM's editors.[127] This is true (at least the only major users), with the exception of the current thesis. His summary of the reception of the CBGM is consistent with that presented above, albeit shorter. He includes one author not included above, namely Timo Flink. Flink's doctoral dissertation on Atticism includes a relatively short section building on Wasserman's work on Jude in which he

---

[125] Carlson, 'Fundamentals of New Testament Textual Criticism. By Stanley E. Porter and Andrew W. Pitts.', 270.

[126] Gurry, *A Critical Examination of the Coherence-Based Genealogical Method*, 1.

[127] See ibid., 2.

engages with the CBGM.[128] Gurry sums up Flink's use of the CBGM as "idiosyncratic and flawed at a key point."[129] Gurry sets out a good introduction to the CBGM as used in the Catholic Epistles demonstrating his accurate grasp of the method. He then goes on, in part two, to examine methodological questions such as the nature of the initial text, and he broadly rejects CBGM's claim to be a "meta method" as it is not suitable for use by thoroughgoing eclectics, and unnecessary for those favouring Byzantine priority.[130] He uses the CBGM's ability to suggest coincidental agreement in variant units (i.e. multiple emergence) to explore scribal habits in James, and encourages others to do this as an alternative to using singular readings in this cause.

In part three of his book, Gurry turns to the more problematic claim that the CBGM can represent the history of the manuscript tradition – something that Jongkind and Carlson have also criticised.[131] First Gurry tests this using the Harklean Group and its relationship to the later Byzantine text, concluding that the CBGM "can be useful for answering historical questions", and shows how to do so with regard to these witnesses.[132] He then discusses the question of which variant readings should be included, considering "meaningful singular readings, orthographica, nonsense readings, and corrections", and makes various suggestions.[133] Gurry goes on to list a number of limitations and suggestions where the CBGM may be improved – overlapping in only one small place with the current thesis's Chapter 8 discussing the display of the strength of textual flow. Gurry concludes that "The CBGM is a valuable tool for reconstructing the text of the [N]ew Testament and for studying

---

[128] See Wasserman, *The Epistle of Jude*.

[129] Gurry, *A Critical Examination of the Coherence-Based Genealogical Method*, 25.

[130] See ibid., 101ff.

[131] See, for example, Jongkind, 'On the Weighing and Counting of Variants'; Jongkind, 'On the Nature and Limitations of the Coherence Based Genealogical Method.'; Carlson, *The Text of Galatians (Book)*; Carlson, 'Comments on the Coherence-Based Genealogical Method'.

[132] Gurry, *A Critical Examination of the Coherence-Based Genealogical Method*, 179.

[133] Ibid., 205.

its textual history."[134]

Gurry and Wasserman's book is a long-needed "accessible, student-friendly user's guide to the CBGM".[135] They successfully argue that "it is crucial that serious students of the New Testament understand the method's origins, procedures, application, and results."[136] Their book is to be highly recommended to anyone interesting in *using* the CBGM, and will also be useful to those seeking to engage critically with the method (although this book is explicitly not a "critical examination or test of the method").[137] They explain why someone should want to understand the method, and set out some of the more significant changes it has brought to the wider world of textual criticism – including some of the changes found in recent versions of the Nestle-Aland edition and the conclusive rejection of the concept of text-types (which had been central to the discipline for two centuries).[138] These changes illuminate the controversy surrounding the CBGM. They explain pre-genealogical coherence and how to use it and similarly genealogical coherence, textual flow diagrams and the global stemma (including its construction). The book is rooted entirely in the INTF's implementation of the CBGM, with instructions on which settings to choose in that interface.[139]

Gurry and Wasserman point out something that has long concerned me about the current use of the CBGM, namely that it still uses only a small portion of the available data – a selection of only Greek manuscripts, only as they left their scribes.[140] See Chapter 9 for a further discussion of the inclusion of correctors' texts. The problems of including more (all?) manuscripts are the practical question of transcribing them all, and also that including so

---

[134] Ibid., 221.

[135] Wasserman and Gurry, *A New Approach to Textual Criticism*, 15.

[136] Ibid., 121.

[137] Ibid., 15.

[138] Ibid., 10. Here Gurry and Wasserman state that the CBGM is itself a replacement for text types and they wonder whether the community will accept that.

[139] See INTF, '"Genealogical Queries" v1 (for the 1st Edition of ECM Vol IV)'; INTF, '"Genealogical Queries" v2 (for the 2nd Edition of ECM Vol IV)'; Mink, 'Guide to "Genealogical Queries"'.

[140] See the discussion above regarding Wasserman, 'Historical and Philological Correlations and the CBGM as Applied to Mark 1:1'.

many extremely similar Byzantine (and other) manuscripts would risk the editor missing important details amid the morass of similarity. See Chapter 7 for an experiment including all the extant manuscripts of John 18 in the CBGM. Wasserman and Gurry highlight a serious gap: “What scholars and student[s] most need now is a version of the CBGM that allows them to edit the local stemmata and create their own database of genealogical data.”[141] My Open Source implementation of the CBGM, created for this thesis, allows just that.[142]

---

[141] Wasserman and Gurry, *A New Approach to Textual Criticism*, 119.

[142] See https://doi.org/10.5281/zenodo.1296287 or https://github.com/edmondac/CBGM

Chapter 3: CBGM Example

John 6:21-24 (Papyri and Majuscules)

Introduction

The best way to understand the CBGM is to work through an example. What follows here is aimed, primarily, at introducing the method and not at creating a critical edition of the four verses in question: John 6:21-24. Beginning with the task of collating the witnesses, this example is presented in minute detail, allowing a transparent and thorough understanding of how to use the CBGM for such verses. Klaus Wachtel kindly reviewed a draft of this chapter and, along with many other helpful comments, argued for some different collation decisions. I have used his input to analyse the effect on the CBGM's results of making different collation decisions – including a resultant change to the initial text. There is also a brief excursus on the CBGM's decision to exclude non-firsthand correctors, considering the text of 01. I then compare the CBGM's results with those of phylogenetic software with positive results. The coherence tables and stemmata in this chapter were created using my own implementation of the CBGM.[1]

Witnesses

There are 30 extant papyri and majuscules witnessing to this passage included in the ECM edition of John[2]: 𝔓28, 𝔓75, 01, 02, 03, 05, 07, 09, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 039, 044, 045, 047, 063, 091, 0141 and 0211. 𝔓28 is

---

[1] See https://github.com/edmondac/CBGM or https://doi.org/10.5281/zenodo.1296287 . The analysis in this chapter was carried out in 2015 and 2016. As my implementation of the CBGM developed I encountered a number of bugs and problems that required the chapter be re-written. The final version of the software used was "beta1", see https://github.com/edmondac/CBGM/releases/tag/beta1 .

[2] Correct as of July 2015. The ECM edition of John is being worked on by the IGNTP.

fragmentary and only extant in parts of verses 21 and 22. 091 is only extant in nine folios but includes v23-24 and half of v22.

Collation

Before starting the CBGM proper, it is necessary to collate the witnesses. The following representation is based on the ECM's layout. The goal here is to define the variant units for use in the CBGM. It should be noted, at this point, that the CBGM is not interested in simple errors or orthographical readings but only those readings that could conceivably make sense in the context. Words without variation are not included in any variant unit because of Wachtel's rule of thumb for defining variant units: "As short as possible and as long as necessary"[3].

As a base text let us take a form of the *Textus Receptus (TR)* with each word given an even numbered address. In the ECM the running text is given even numbered addresses with the odd numbers being used for additions in other witnesses. Word numbers appear after a slash, e.g. John 6:21/2. The *TR* is a good choice, generally, as a base text as it tends to include the longer form of most texts. Hence additions to it will typically be small. If 03 were chosen, for example, as a base text instead, then v22 would contain a long addition as a single odd-numbered word which would better be represented as several variant units.

> [21]ηθελον ουν λαβειν αυτον εις το πλοιον και ευθεως το πλοιον εγενετο επι της γης εις ην υπηγον [22]τη επαυριον ο οχλος ο εστηκως περαν της θαλασσης ιδων οτι πλοιαριον αλλο ουκ ην εκει ει μη εν εκεινο εις ο ενεβησαν οι μαθηται αυτου και οτι ου συνεισηλθε τοις μαθηταις αυτου ο ιησους εις το πλοιαριον αλλα μονοι οι μαθηται αυτου απηλθον [23]αλλα ηλθεν πλοιαρια εκ τιβεριαδος εγγυς του τοπου

---

[3] Klaus Wachtel, personal email December 2014.

οπου εφαγον τον αρτον ευχαριστησαντος του κυριου [24]οτε ουν ειδεν ο οχλος οτι ιησους ουκ εστιν εκει ουδε οι μαθηται αυτου ενεβησαν αυτοι εις τα πλοια και ηλθον εις καπερναουμ ζητουντες τον ιησουν.

We will now regularise the text of all the witnesses and collate them. The output of this exercise is a list of variant units and attestations. First we consider John 6:21/2-14: "ηθελον ουν λαβειν αυτον εις το πλοιον". In John 6:21/2, 01 reads ηλθον where the rest of the witnesses attest ηθελον.[4] This is the first variant unit. In word 6, 01's reading λαβιν may be regularised as an orthographical variant of λαβειν, and 032's error reading βαλιν may be too. Three witnesses (05, 032 and 044) transpose words 6 and 8, giving a second variant unit. 063 spells πλοιον as πλυον, which can be regularised. Thus there are two variant units here, 21/2 and 21/6-8 – as shown in Table 1. All variant units will be highlighted in grey. Any words in white, therefore, are constant across all the manuscripts, post-regularisation.

ηθελον	ουν	λαβειν	αυτον	εις	το	πλοιον
2	4	6	8	10	12	14
b ηλθον		b αυτον λαβειν				

Table 1: Variant units John 6:21/2 and 21/6-8

Next we consider John 6:21/16-36: "και ευθεως το πλοιον εγενετο επι της γης εις ην υπηγον". In John 6:21/18 021's ευθεος may be regularised as orthographic. In word 22, 063 continues spelling πλοιον as πλυον, and also spells εγενετο as εγενετω – both of which are regularised. Words 20-24 are reordered by several witnesses creating the first variant unit here, with 05 introducing a further variation with εγενηθη instead of εγενετο. In words 28-30, 01 and 0211 read the accusative rather than the genitive. 𝔓75 here could support either

---

[4] 𝔓28 is lacunose after the first two letters, but that is enough to support the reading ηθελον. 091 is lacunose in this whole verse.

reading as only the first eta is extant – and the simplest way to treat this is to consider 𝔓75 as lacunose for this variant unit. 019 and 038 both read ιν for ην – which must be an error and so can be regularised. There are further orthographic readings for υπηγον, with 0211 reading υπιγον and 063 υπηγων. 01, however, has a genuine variant reading here: υπηντησεν. These variant units are shown in Table 2.

και	ευθεως	το	πλοιον	εγενετο	επι	της	γης	εις	ην	υπηγον
16	18	20	22	24	26	28	30	32	34	36
		b εγενετο το πλοιον				b την γην				b υπηντησεν
		c το πλοιον εγενηθη								

Table 2: Variant units John 6:21/20-24, 21/28-30 and 21/36

Now we turn to John 6:22/2-18: "τη επαυριον ο οχλος ο εστηκως περαν της θαλασσης". In John 6:22/2, 019 mistakenly wrote τι which can be regularised. 0211 inserted τε in word 3, which is the first variant unit in this verse. In word 10, 038 and 045 omit the article. In word 12, 01 reads εστως instead of εστηκως. Here 𝔓28 could support either reading – so will be considered lacunose (which it is in the following words). 03's περα is regularised to περαν in word 14. The resulting three variant units are shown in Table 3.

τη		επαυριον	ο	οχλος	ο	εστηκως	περαν	της	θαλασσης
2	3	4	6	8	10	12	14	16	18
	b τε				b *om*	b εστως			

Table 3: Variant units John 6:22/3, 22/10 and 22/12

Next is John 6:22/20-38: "ιδων οτι πλοιαριον αλλο ουκ ην εκει ει μη εν". John 6:22/20 presents several variations, all of which are possible and so cannot be regularised: 07, 09, 011, 013, 017, 021, 028, 030, 031, 034, 036, 039, 044, 045, 047, 063 and 0141 attest ιδων; 𝔓75, 02, 03 and 038 attest ειδον; 01 and 05 attest ειδεν; 𝔓28 reads ιδεν which is regularised to ειδεν;

and 0211 reads ειδως, yet another valid reading. Two readings present problems: 037 reads ειδων and 019, 022 and 032 attest ιδον. Now ειδων and ιδον are invalid in themselves – but it is not clear how to regularise them. Both words could be regularised to either ειδον or to ιδων. In an edition, the manuscripts could be recorded as providing support for either reading, but for the CBGM we need a single record of attestation. So they must be recorded as if they were lacunose in this location (i.e. unable to support any single reading). In word 30, 0211 reads ειν instead of ην, which can be regularised. Similarly, in 34, 063 reads η instead of ει. There are no other variant units in this section.

ιδων	οτι	πλοιαριον	αλλο	ουκ	ην	εκει	ει	μη	εν
20	22	24	26	28	30	32	34	36	38
b ειδον									
c ειδεν									
d ειδως									

Table 4: Variant unit John 6:22/20

We now consider John 6:22/40-52: "κεινο εις ο ενεβησαν οι μαθηται αυτου". 𝔓28 is lacunose from this point on and 091 has not yet begun. The passage John 6:22/40-52 is missing entirely from 𝔓75, 02, 03, 019, 022, 032, 044 and 063. This can simply be represented as an overlapping variant unit. In word 40, 01 misspells εκεινο as κεινο, which is therefore regularised.[5] 036, 039 and 045 have an error reading εκεινω which is also regularised. In addition to the witnesses missing the whole phrase, 05 and 0211 omit this word.[6] Similarly, word 42 is omitted by 034. 037 and 047 read ανεβησαν in word 46. 01 reads του ιησου in word 52, and 05 simply ιησου.

---

[5] This word is a simple variant unit here. However, if we consider other witnesses here (minuscules, lectionaries, correctors etc.) we would find that εκεινο is sometimes used to replace εν in word 38. So we would have to combine those units. But for our dataset, they are safe to remain separate.

[6] Note that in the attestations for this variant unit, only 063 and 0211 will support the omission *of the word from this phrase*. That is because the witnesses that omit the whole phrase cannot provide support for this word or its individual omission and are therefore recorded as lacunose.

Collation

εκεινο	εις	ο	ενεβησαν	οι	μαθηται	αυτου
40	42	44	46	48	50	52
b *om*	b *om*		b ανεβησαν			b του ιησου
						c ιησου

Table 5: Variant units John 6:22/40, 22/42, 22/46 and 22/52.

εκεινο	εις	ο	ενεβησαν	οι	μαθηται	αυτου
40	42	44	46	48	50	52
b *om*						

Table 6: Overlapping variant unit John 6:22/40-52

The next section is John 6:22/54-76: "και οτι ου συνεισηλθε τοις μαθηταις αυτου ο ιησους εις το πλοιαριον". 02 and 019 attest ο instead of ου for word 58, but this is an error and is regularised. 091 is now extant, beginning half way through the word συνειςηλθεν. Its movable nu is shared with other witnesses and can be regularised to συνεισηλθε. Similarly, 011, 019, 038 and 0211 spell word 60 συνησηλθεν, which is also regularised. 01, however, uniquely attests συνεληλυθι – which is a rare form (of συνερχομαι as opposed to συνεισερχομαι) and is thus the first variant unit here. 02 inserts a (clarifying) ο ιησους as word 61 despite there already being one at words 68-70, creating another variant unit. In words 62-66 01 simply reads αυτοις, making a further variant unit. 034 omits words 68-70. 05 has mistakenly missed the epsilon off word 72, but this can be regularised. Finally, 𝔓75, 01, 02, 03, 05, 017, 019, 022, 032, 044, 091, 0141 and 0211 read πλοιον for word 76 (and 031 is lacunose here). 063 is consistent, spelling the last word πλυαριον, which is regularised as before.

και	οτι	ου	συνεισηλθε		τοις	μαθηταις	αυτου	ο	ιησους	εις	το	πλοιαριον
54	56	58	60	61	62	64	66	68	70	72	74	76
			b συνεληλυθι	b ο ιησους	b αυτοις			b *om*				b πλοιον

Table 7: Variant units John 6:22/60, 22/61, 22/62-66, 22/68-70 and 22/76.

Now to John 6:22/78-88: “αλλα μονοι οι μαθηται αυτου απηλθον”. In word 80, 05 and 047 read μονον which is a valid form and is therefore included, even if it is grammatically incorrect. 01 misses out the final word, 038 attests εισηλθον, and 031 is once again lacunose.

αλλα	μονοι	οι	μαθηται	αυτου	απηλθον
78	80	82	84	86	88
	b μονον				b εισηλθον
					c *om*

Table 8: Variant units John 6:22/80 and 22/88.

Next is John 6:23/2-10: “αλλα ηλθεν πλοιαρια εκ τιβεριαδος”. 022 adds the word και at the start. 031 is lacunose until word 4. 01 and 05 omit word 2. All but 𝔓75, 01, 03, 05, 019 and 091 add δε as word 3. Words 4-6 have many readings: most witnesses attest ηλθεν πλοιαρια (including 030 and 034 which have no movable nus and are regularised, and 063 which continues to spell πλοιαρια with an upsilon); 𝔓75, 03 and 032 read ηλθεν πλοια; 021, 022, 036, 047 and 091 read ηλθον πλοιαρια; this is transposed in 017 as πλοιαρια ηλθον; 044 reads πλοια ηλθεν, 019 simply reads πλοιαρια and 0141 attests just πλοια. Then there are two very different readings: 01 reads επελθοντων ουν των πλοιων; and 05 reads αλλων πλοιαρειων ελθοντων. We will regularise πλοιαρειων to πλοιαριων. 03, 022 and 032 add της before τιβεριαδος and 031 is again lacunose. 017 and 030 spell word 10 as τηβεριαδος and 063 as τιβερηαδος – both of which are regularised as they clearly refer to the same place. Then 0141 adds ηλθεν at the end (presumably in place of word 4) and 019 similarly adds ηλθον. 01 and 05’s shared omission of word 2 and unusual readings for words 4-6 warrant an overlapping unit. The choice of location for ηλθον/ηλθεν as word 4, 6 or 11 means that the best solution is to have a single variant unit for words 4-10[7], with 01 and 05 overlapping words 2-10. These

---

[7] There is no need to continue to refer to “word 11” now, as the unit 4-10 will cover it.

variant units are shown in Table 9.

<table>
<tr><td></td><td>αλλα</td><td></td><td>ηλθεν</td><td>πλοιαρια</td><td>εκ</td><td>τιβεριαδος</td></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>6</td><td>8</td><td>10</td></tr>
<tr><td>b και</td><td></td><td>b δε</td><td colspan="4">b ηλθεν πλοια εκ τιβεριαδος</td></tr>
<tr><td></td><td></td><td></td><td colspan="4">c ηλθεν πλοια εκ της τιβεριαδος</td></tr>
<tr><td></td><td></td><td></td><td colspan="4">d ηλθον πλοιαρια εκ τιβεριαδος</td></tr>
<tr><td></td><td></td><td></td><td colspan="4">e ηλθον πλοιαρια εκ της τιβεριαδος</td></tr>
<tr><td></td><td></td><td></td><td colspan="4">f πλοιαρια ηλθον εκ τιβεριαδος</td></tr>
<tr><td></td><td></td><td></td><td colspan="4">g πλοια ηλθεν εκ τιβεριαδος</td></tr>
<tr><td></td><td></td><td></td><td colspan="4">h πλοιαρια εκ τιβεριαδος ηλθον</td></tr>
<tr><td></td><td></td><td></td><td colspan="4">i πλοια εκ τιβεριαδος ηλθεν</td></tr>
</table>

Table 9: Variant units John 6:23/1, 23/3, 23/4-10

<table>
<tr><td>αλλα</td><td>ηλθεν</td><td>πλοιαρια</td><td>εκ</td><td>τιβεριαδος</td></tr>
<tr><td>2</td><td>4</td><td>6</td><td>8</td><td>10</td></tr>
<tr><td colspan="5">b επελθοντων ουν των πλοιων εκ τιβεριαδος</td></tr>
<tr><td colspan="5">c αλλων πλοιαριων ελθοντων εκ τιβεριαδος</td></tr>
</table>

Table 10: Overlapping variant unit John 6:23/2-6

We turn to John 6:23/14-26: “εγγυς του τοπου οπου εφαγον τον αρτον ευχαριστησαντος του κυριου”. 032 omits words 12-16. 01 reads εγγυς ουσης here. 063 misspells εγγυς as εγγεις but this can be regularised. In word 16, 031 (which is extant once again) agrees with the majority of witnesses and reads τοπου. This results in a simple variant unit for words 12-16. 01 inserts και before εφαγον and omits the article in word 22, where 031 becomes lacunose once more. As might be expected, there is some orthographic variation in word 26, with 021 reading ευχαριστισαντος and 038 reading ευχαρηστησαντος – but both can be regularised. 031 appears halfway through the word offering no variation. 05 and 091, however, omit the entire phrase (26-30). 02 has a variant reading (θεου) for word 30, but this can all be represented as a single variant unit 26-30 as shown in Table 11.

εγγυς	του	τοπου	οπου	εφαγον	τον	αρτον	ευχαριστησαντος	του	κυριου
12	14	16	18	20	22	24	26	28	30
b εγγυς ουσης				b και εφαγον			b ευχαριστησαντος του θεου		
c *om*							c *om*		

Table 11: Variant units John 6:23/12-16, 23/20-22, 23/26-30

In verse 24, 01 has some unique readings. We consider the first half of the verse, John 16:24/2-28: "οτε ουν ειδεν ο οχλος οτι ιησους ουκ εστιν εκει ουδε οι μαθηται αυτου". For words 2-10, 01 reads και ιδοντες instead of οτε ουν ειδεν ο οχλος. For words 14-20 it reads ουκ ην εκει ο ιησους instead of ιησους ουκ εστιν εκει. These are best expressed as two overlapping variants. Word 6 exhibits not unexpected variation, with 013 reading ειπεν, 02, 017, 019, 022, 039, 044, 063, 091 and 0211 reading ιδεν (which can be regularised to ειδεν) and 030 reading εγνω. 063 is damaged in parts of verse 24 (12-14, 20-32, 38-42) and is largely unreadable – although is readable enough in places to provide some support, for example for the *a* reading in words 2-10 and 14-20. 038 adds the article before ιησους (a reading supported by 01), 013 omits ιησους and 031 is once again lacunose.[8] After word 18, 091 is lacunose. 01 omits word 28, having returned to the norm in the previous words.

οτε	ουν	ειδεν	ο	οχλος	οτι	ιησους	ουκ	εστιν	εκει	ουδε	οι	μαθηται	αυτου
2	4	6	8	10	12	14	16	18	20	22	24	26	28
		b ειπεν				b ο ιησους							b *om*
		c εγνω				c *om*							

Table 12: Variant units John 6:24/6, 24/14 and 24/28

---

[8] It would be possible to define two variant units here: one for the added article, and one for the omission of ιησους. Both approaches are valid collations, but would lead to subtly different coherence data later. The effect of collation decisions is explored later in this chapter.

Collation

οτε	ουν	ειδεν	ο	οχλος
2	4	6	8	10
b και ιδοντες				

Table 13: Overlapping variant unit John 6:24/2-10

ιησους	ουκ	εστιν	εκει
14	16	18	20
b ουκ ην εκει ο ιησους			

Table 14: Overlapping variant unit John 6:24/14-20

We must tackle the remaining half of verse 24 in two parts. First John 6:24/30-38: "ενεβησαν αυτοι εις τα πλοια". These few words prove surprisingly difficult to collate due to two witnesses: 05 and 0211. 05 here reads ελαβον εαυτοις πλοιαρια – which is best expressed as an overlapping unit for the whole phrase (30-38). 0211 here reads τοτε και αυτοι ενεβεισαν εις τα πλοια – uniquely adding τοτε and changing the word order of και αυτοι ενεβεισαν (some other witnesses also add και as word 31). This too works best as an overlapping variant (30-32). This is an excellent example of how overlapping units and normal units interact, as follows:

In word 30 (ενεβησαν) 𝔓75, 01 and 019 read ανεβησαν. 0211 reads ενεβεισαν (out of order) but this is regularised to ενεβησαν and thus 0211 provides support to the base text here. This can be counter intuitive as 0211 does not read either variant *in word 30*, but later. This is best explained by considering each variant unit as representing a question. Unit 30 is about variations in the word ενεβησαν while the overlapping unit 30-32 is really about word order: αυτοι ενεβησαν or ενεβησαν αυτοι, even though it also includes τοτε και at the start. τοτε is a unique addition in 0211 and so is easily subsumed into this unit. και is added here by 0211 and also by 030 and 036 as word 31. Therefore 0211 will support the addition of και in the variant unit for word 31 even though it has the word out of place (as this unit represents the question “is και included in this phrase?”) A further unit represents the omission of word 32 by 01 and 028.

Now 05 here (ελαβον εαυτοις πλοιαρια) is the only witness to omit εις and the article, and the only witness to attest this beginning to the phrase. This could be split into further variant units. In several of those cases the attestations would simply bisect the witnesses into 05 and everything else. But it is simpler and clearer to keep the whole phrase together. One problem with apparatuses is that they can make it difficult to determine what an individual witness reads over several words and this overlapping unit avoids that problem. However, splitting it up would create a variant unit for every word and therefore would change the pre-genealogical coherence. Such decisions taken at this stage will have consequences later, but they are hard to identify now.

01 has the singular το πλοιον in words 36 to 38, rather than the plural. So word 36 is a variant unit in which most witnesses attest τα, 01 το and 05 is considered lacunose because its reading is covered by an overlapping unit and it is unable to support either τα or το here. This is also true of units 24/30, 24/31, 24/32 and 24/30-32. In word 38, 𝔓75, 03, 019, 022, 032 and 044 read πλοιαρια (supported by 05 from the overlapping unit) instead of πλοια (or 01’s πλοιον). This support by 05 in word 38 is the reason that 24/36 and 24/38 exist as separate units, otherwise they could be combined into a single unit. 063 is briefly readable for words 34 and 36 and supports the base text in both cases. These all add up to the seven variant units shown in Table 15, Table 16 and Table 17.

Collation

ενεβησαν		αυτοι	εις	τα	πλοια
30	31	32	34	36	38
b αναβησαν	b και	b om		b το	b πλοιον
					c πλοιαρια

Table 15: Variant units John 6:24/30, 24/31, 24/32, 24/36 and 24/38

ενεβησαν	αυτοι	εις	τα	πλοια
30	32	34	36	38
b ελαβον εαυτοις πλοιαρια				

Table 16: Overlapping variant unit John 6:24/30-36

ενεβησαν	αυτοι
30	32
b τοτε και αυτοι ενεβησαν	

Table 17: Overlapping variant unit John 6:24/30-32

Finally we consider John 6:24/40-52: “και ηλθον εις καπερναουμ ζητουντες τον ιησουν”. Unexpectedly after the last few sections, these words exhibit little variation. The first variation is purely orthographic in nature with different spellings of καπερναουμ (καφαρναουμ and καπαρναουμ) which are regularised. The following word also has an orthographic regularisation, ζητουνταις for ζητουντες. The only real variant unit is in words 50-52 where 017 reads αυτον instead of τον ιησουν.

και	ηλθον	εις	καπερναουμ	ζητουντες	τον	ιησουν
40	42	44	46	48	50	52
					b αυτον	

Table 18: Variant unit John 6:24/50-52

The complete set of variant units and manuscript attestations is summarised in Table 19. The first letter in the second column represents the reading in the first variant unit, the second the second etc. A lacuna is represented by a question mark. As can be seen, there is a lot of variation, with only three witnesses attesting exactly the same (regularised) text throughout: 07, 09 and 039 – all of which read the base text in all but one variant unit (John 6:23/3, addition of δε). 031 agrees with the base text everywhere it is extant.

𝔓28	`a????aa?c????????????????????????????????`
𝔓75	`aab?aaaab?b???aaaabaaaaabaaaaaaaaabaaaaba`
01	`baabbaabcaaaabbababacab??bbab?bbbabaacbca`
02	`aabaaaaab?b???abaabaaaabaaabaaaaaaaaaaaaa`
03	`aabaaaaab?b???aaaabaaaaacaaaaaaaaaaaaaaba`
05	`abcaaaaacbaaacaaaabbaac??aacaaaaa??b???ba`
07, 09, 039	`aaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaa`
011	`aabaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaa`
013	`aaaaaaaaaaaaaaaaaaaaaaabaaaaabacaaaaaaaaa`
017	`aaaaaaaaaaaaaaaaaabaaaabfaaaaaaaaaaaaaaab`
019	`aabaaaaa??b???aaaabaaaaahaaaaaaaaabaaaaba`
021	`aaaaaaaaaaaaaaaaaaaaaaabdaaaaaaaaaaaaaaaa`
022	`aabaaaaa??b???aaaabaababeaaaaaaaaaaaaaaba`
028	`aaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaacaaa`
030	`aaaaaaaaaaaaaaaaaaaaaaabaaaaacaaaaaabaaaa`
031	`aaaaaaaaaaaaaaaaaa?a??a?aaaaaa??aaaaaaaaa`
032	`abbaaaaa??b???aaaabaaaabccaaaaaaaaaaaaaba`
034	`aaaaaaaaaaabaaaaabaaaaabaaaaaaaaaaaaaaaaa`
036	`aaaaaaaaaaaaaaaaaaaaaaabdaaaaaaaaaaabaaaa`
037	`aaaaaaaa?aaabaaaaaaaaaabaaaaaaaaaaaaaaaaa`
038	`aaaaaababaaaaaaaaaaabaabaaaaaaaabaaaaaaaa`
044	`abbaaaaaa?b???aaaabaaaabgaaaaaaaaaaaaaaba`
045	`aaaaaabaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaa`
047	`aaaaaaaaaaaabaaaaaabaaabdaaaaaaaaaaaaaaaa`
063	`aaaaaaaaa?b???aaaaaaaaabaaaaaaa???????a?a`
091	`??????????????aaaabaaaaadaacaaaa?????????`
0141	`aabaaaaaaaaaaaaaaabaaaabiaaaaaaaaaaaaaaaa`
0211	`aaababaadbaaaaaaaabaaaabaaaaaaaaabaabaaaa`

Table 19: Summary of attestations in John 6:21-24

𝔓28 is so fragmentary, and only differs from the base text in one variant unit (in which it agrees with 05), that it will be excluded from this point. Similarly, because the CBGM is interested in texts and not manuscripts, there is no benefit in keeping 09 and 039 (which read the same as 07). So the list of 27 witnesses to use from now on is: 𝔓75, 01, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 091, 0141 and 0211.

Pre-genealogical coherence

Pre-genealogical coherence data is now available. For example, pre-genealogical coherence for 047 is shown in Table 20.[9] Recall that a high value for pre-genealogical coherence argues in favour of a genealogical relationship between the two witnesses and consequently argues against multiple emergence of a variant. Conversely, a low value for pre-genealogical coherence argues in favour of multiple emergence.

W2	**NR**	**PERC1**	**EQ**	**PASS**
021	1	95.122	39	41
037	2	95.000	38	40
036	3	92.683	38	41
07		92.683	38	41
031	5	91.429	32	35
011	6	90.244	37	41
028		90.244	37	41
045		90.244	37	41
063	9	89.655	26	29
013	10	87.805	36	41
0141		87.805	36	41
017		87.805	36	41
030		87.805	36	41
034		87.805	36	41
038	15	82.927	34	41
044	16	81.081	30	37
022	17	80.556	29	36
02	18	78.378	29	37
03		78.378	29	37
019	20	77.778	28	36
032		77.778	28	36
091		77.778	14	18
0211	23	75.610	31	41
𝔓75	24	75.000	27	36
05	25	67.647	23	34
01	26	39.474	15	38

Table 20: Pre-genealogical coherence for W1=047

---

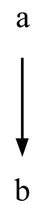
[9] The headings are defined as follows: **W2** – the witness to compare to 047 (W1); **NR** – the rank, or position in the table; **PERC1** – the pre-genealogical coherence percentage; **EQ** – the number of variant units in which the two witnesses agree; **PASS** – the number of variant units in which the two witnesses are both extant.

Local stemmata

We now consider the variant units in turn, trying first to create uncontroversial local stemmata and using pre-genealogical coherence data where applicable. Supporting witnesses are shown for each reading in the table following the discussion of the variant unit, with an ellipsis indicating support from all otherwise unmentioned witnesses. Only a selection of local stemmata are included here. For the rest please see Appendix 1.

<u>John 6:21/2</u>

Here 01 uniquely reads ηλθον which seems a straightforward copying error from ηθελον. The local stemma is shown in Table 21. This "*a* → *b*" local stemma will be typical of the simple ones that can be created at this stage of the process.

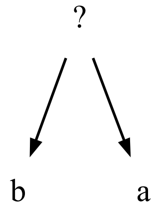


a	ηθελον	...
b	ηλθον	01
	lac	091

Table 21: Local stemma for John 6:21/2

<u>John 6:21/6-8</u>

John 6:21/6-8 is not straightforward and the creation of the local stemma will be deferred. The readings are shown in Table 22.

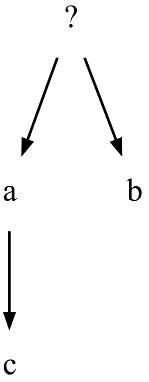


a	λαβειν αυτον	...
b	αυτον λαβειν	05, 032, 044
	lac	091

Table 22: Local stemma for John 6:21/6-8

John 6:21/20-24

Here we have three variant readings. The first two are purely about word order. In other situations considering other instances of the construction in question in John is valuable. However, there is little to be gained in this instance. For example, in John 7:43, witnesses are similarly divided between εγενετο εν τω οχλω and εν τω οχλω εγενετο. Again, this decision will be deferred until more evidence is available. However, the third reading here is easier to analyse. Codex Bezae has a tendency to change words to improve the text, and εγενετο to εγενηθη is such an improvement. Thus it seems clear that it descends from reading *a*. The partial local stemma is shown in Table 23.

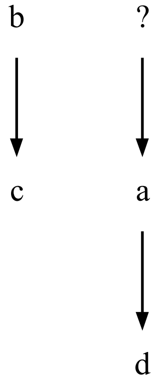


a	το πλοιον εγενετο	...
b	εγενετο το πλοιον	𝔓75, 02, 03, 011, 019, 022, 032, 044, 0141
c	το πλοιον εγενηθη	05
	lac	091

Table 23: Local stemma for John 6:21/20-24

John 6:22/20

The variations in John 6:22/20 present something of a challenge. Reading *b* (ειδον) provides a simple reading but the following clauses do not flow easily. This is quite a typical Johannine construction with different statements assembled next to one another but not well integrated into a convincing whole. Reading *a* (ιδων) appears to improve the text by removing a finite verb and making a large section parenthetical.[10] Reading *c* (ειδεν) is an improvement on ειδον as οχλος is a collective noun. Reading *d* (ειδως) seems like an improvement on ιδων. The only remaining question is whether *a* is derived from *b* or *c* but this decision must be deferred until genealogical coherence data is available. The partial local stemma is shown in Table 24.



a	ιδων	07, 011, 013, 017, 021, 028, 030, 031, 034, 036, 044, 045, 047, 063, 0141
b	ειδον	𝔓75, 02, 03, 038
c	ειδεν	01, 05
d	ειδως	0211
	lac	019, 022, 032, 037, 091

Table 24: Partial local stemma for John 6:22/20

<u>John 6:22/40-52</u>

The phrase "εκεινο εις ο ενεβησαν οι μαθηται αυτου" in John 6:22/40-52 seems likely to be a later addition: it clarifies the "one boat" and improves the text markedly. So this shorter version is the initial text. The local stemma is shown in Table 25. Note, this is the first complete local stemma where the initial text has deviated from the base text. In the ECM the readings would be relabelled so that the initial text was *a*. But for simplicity and clarity we

---

[10] This parenthetical section could potentially continue all the way to verse 24 when the crowd board the boats.

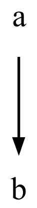
will leave the labels alone. There are a number of variations in the added phrase which are detailed in Appendix 1.

a	εκεινο εις ο ενεβησαν οι μαθηται αυτου ...	
b		𝔓75, 02, 03, 019, 022, 032, 044, 063
lac		091

Table 25: Local stemma for John 6:22/40-52

<u>John 6:22/60</u>

In John 22:60 all but 01 read συνεισηλθε (aorist of συνεισερχομαι) and 01 reads συνεληλυθι (perfect of συνερχομαι). This is a very rare form – indeed this is the only occurrence of the word in any manuscript of John included in the ECM. The ca corrector of 01 changed it to συνεισηλθεν.[11] It seems impossible to imagine this being the original, and so we create the local stemma in Table 26.[12]

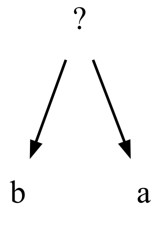


a	συνεισηλθε ...	
b	συνεληλυθι	01

Table 26: Local stemma for John 6:22/60

<u>John 6:22/62-66</u>

Once again, 01 here has a unique reading. This time, however, it is a simple reading that has much to recommend it as a forebear of the text attested by all the other witnesses:



a	τοις μαθηταις αυτου ...	
b	αυτοις	01

Table 27: Local stemma for John 6:22/62-66

---

[11] A correction originating in a revision of the text in the fifth to seventh centuries.

[12] This ties in with the CBGM maxim that singular readings are normally derivative. See the "Nine Guidelines" later in this chapter.

αυτοις is more primitive than τοις μαθηταις αυτου. Why would a scribe change to the more primitive form, especially when it adds some ambiguity (crowd or disciples)? On the face of the readings alone reading *b* therefore seems to be prior. But with only one witness supporting it (and a peculiar one in these verses) would that really be a legitimate choice? NA27 does not even mention the variant but simply reads τοις μαθηταις αυτου. The decision will be deferred for now[13]. The local stemma is shown in Table 27.

John 6:22/80

This variant unit affords us an opportunity to use pre-genealogical coherence. Here 05 and 047 both read the grammatically incorrect form μονον (accusative singular) instead of μονοι (nominative plural). Pre-genealogical coherence allows us to explore the question of whether this is a case of multiple emergence. Pre-genealogical coherence data is shown for 05 in Table 28 and for 047 in Table 20 on page 92. 047 is 05's joint tenth closest relative (but with only 67.647% coherence) yet in the other direction 05 is 047's *twenty-fifth* closest (with the same 67.647% coherence). This is an example of the importance of considering the percentage and not just the rank in these CBGM tables. See Chapter 7 for a discussion of this issue. These coherence values suggest that it is unlikely that these witnesses are related.

W2	**NR**	**PERC1**	**EQ**	**PASS**
091	1	87.500	14	16
044	2	76.667	23	30
019	3	75.862	22	29
032		75.862	22	29
03	5	73.333	22	30
022	6	72.414	21	29
𝔓75		72.414	21	29
0141	8	70.588	24	34
063	9	69.231	18	26
011	10	67.647	23	34
017		67.647	23	34
021		67.647	23	34
0211		67.647	23	34
028		67.647	23	34
036		67.647	23	34
047		67.647	23	34
07		67.647	23	34
037	18	66.667	22	33
02		66.667	20	30
031	20	65.517	19	29
030	21	64.706	22	34
045		64.706	22	34
013	23	61.765	21	34
034		61.765	21	34
038	25	58.824	20	34
01	26	33.333	11	33

Table 28: Pre-genealogical coherence for W1=05

---

[13] This would be a place to consider non-Greek evidence. If there were early Latin or Syriac witnesses attesting this shorter form then it would have more credence than a single Greek witness.

Local stemmata

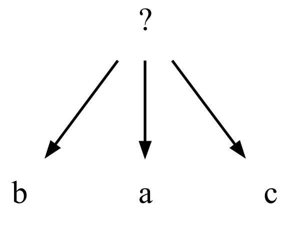
This in turn implies that this variant reading is a case of multiple emergence. Both manuscripts are uncials with no spaces, so the scribe would have had to write ΑΛΛΑΜΟΝΟ**<u>Ν</u>**ΟΙΜΑΘΗΤΑΙ changing only one letter from ΑΛΛΑΜΟΝΟ**<u>Ι</u>**ΟΙΜΑΘΗΤΑΙ – which does seem like an easy enough mistake to happen more than once. The local stemma is shown in Table 29.

a μονοι ...
b μονον 05, 047

Table 29: Local stemma for John 6:22/80

<u>John 6:23/2-10</u>

The overlapping variant unit John 6:23/2-10 contains the unique genitive absolute readings attested by 01 and 05. εκ τιβεριαδος is included in this overlapping unit in order to make the next unit (6:23/4-10) simpler. There is no variation in these two words and 01 and 05 have little or nothing to say on the relevant questions in that unit, such as “where should ηλθον go?” 05’s reading is also attested by some early Latin and Syriac witnesses, implying that it is a very early form.[14] Nevertheless, no local stemma is yet apparent. The readings are shown in Table 30.



a αλλα ηλθεν πλοιαρια εκ τιβεριαδος ...
b επελθοντων ουν των πλοιων εκ τιβεριαδος 01
c αλλων πλοιαριων ελθοντων εκ τιβεριαδος 05

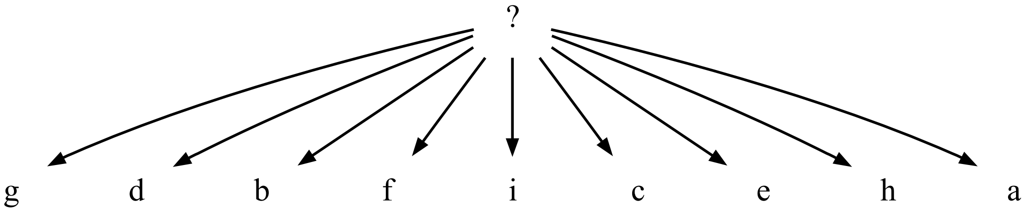
Table 30: Local stemma for overlapping unit John 6:23/2-10

---

[14] NA27 lists support for this reading as “D (b j r^1) sy^c”

John 6:23/4-10

This variant unit has more variation than any other in this dataset. All witnesses in this variant unit contain a variant on ηλθον (in one of three locations) and a variant on πλοιαρια and some include the article before τιβεριαδος.[15] At this stage it is difficult to create a local stemma, so this will be deferred until genealogical coherence data is available. The readings are shown in Table 31.



a	ηλθεν πλοιαρια εκ τιβεριαδος	...
b	ηλθεν πλοια εκ τιβεριαδος	𝔓75
c	ηλθεν πλοια εκ της τιβεριαδος	03, 032
d	ηλθον πλοιαρια εκ τιβεριαδος	021, 036, 047, 091
e	ηλθον πλοιαρια εκ της τιβεριαδος	022
f	πλοιαρια ηλθον εκ τιβεριαδος	017
g	πλοια ηλθεν εκ τιβεριαδος	044
h	πλοιαρια εκ τιβεριαδος ηλθον	019
i	πλοια εκ τιβεριαδος ηλθεν	0141
	lac	01, 05

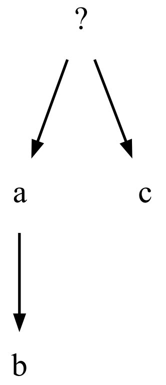
Table 31: Local stemma for John 6:23/4-10

---

[15] All witnesses appear here apart from 01 and 05, which are covered by the overlapping unit John 6:23/2-10.

John 6:23/26-30

02’s reading ευχαριστησαντος του θεου is evidently a mistake in John 6:23/26-30 and must derive from the *a* reading. 05 and 091 omit the phrase and this omission has a claim to being original. It is certainly very early, being attested by some old Latin and Syriac witnesses – the text making sense and being simpler.[16] Table 28 above shows the pre-genealogical coherence for 05. Notably, 091 is 05’s closest relative with 87.5% coherence. This strongly suggests that this omission is *not* a case of multiple emergence. But whether the omission is prior or posterior to reading *a* cannot be conclusively decided at this point. The partial local stemma is shown in Table 32.

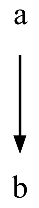


a	ευχαριστησαντος του κυριου	...
b	ευχαριστησαντος του θεου	02
c		05, 091

Table 32: Partial local stemma for John 6:23/26-30

John 6:24/30-32

John 6:24/30-32 is an overlapping unit created for 0211’s unique reading. This is primarily a change in word order, with an extra τοτε at the start. If we were to have constructed a single variant unit for words 30-32 (instead of this overlapping unit and individual variant units for words 30, 31 and 32) then 0211’s reading here would derive from 030 and 036’s ενεβησαν και αυτοι. However, in this case it is derived from the *a* reading of



a	ενεβησαν αυτοι	...
b	τοτε και αυτοι ενεβησαν	0211
lac		05, 063, 091

Table 33: Local stemma for John 6:24/30-32

---

[16] NA27 lists support for this reading as “D 091 *pc* a e $\text{sy}^{s,c}$”.

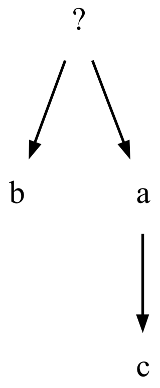
this overlapping unit. It should be noted that either way of dividing the units would have resulted in the genealogical relationship being legitimately encoded, although yielding subtly different statistical data. The local stemma is shown in Table 33.

John 6:24/36 and 38

In John 6:24/36 and 24/38 01 has the crowd getting into a single boat which is problematic and seems to be a mistake originating from a somewhat confused scribe.[17] These few verses are unusually complex, and many scribes seem to have wrestled with the text trying to make more sense of it. The rest of the witnesses disagree about whether the plural boats are small or just "boats". The first local stemma is shown in Table 34 but the second cannot be completed yet. Its partial local stemma is shown in Table 35.

a	τα	...
b	το	01
	lac	05, 091

Table 34: Local stemma for John 6:24/36



a	πλοια	...
b	πλοιαρια	𝔓75, 03, 05, 019, 022, 032, 044
c	πλοιον	01
	lac	063, 091

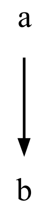
Table 35: Incomplete local stemma for John 6:24/38

---

[17] While 01 has omitted the word οχλος at the start of the verse. οχλος is still the subject at this point, carried over from earlier.

John 6:24/50-52

The final variant unit is John 6:24/50-52. Here most witnesses read *τον ιησον* but 017 alone reads the stylistically better (there is no need for the second mention of Jesus in the verse, as it is unambiguous that the disciples were seeking him and not someone else) and simpler, and thus posterior, *αυτον*.[18] As such the local stemma can be constructed and is shown in Table 36.



a	τον ιησουν ...	
b	αυτον	017
	lac	091

Table 36: Local stemma for John 6:24/50-52

Genealogical Coherence and Textual Flow

Now that all readings have been collated and regularised, and local stemmata created, a prototype initial text “A” becomes available (defined in twenty-eight variant units at this stage). Also genealogical coherence data can be calculated and the local stemmata can be re-evaluated as a result. It is important to note that the differences between the number of prior and posterior variants will be small in this example, as there are only 41 variant units. Differences of only one or two between the values would normally cause some doubt but for the sake of this example we will consider any difference as indicating a meaningful and stable textual flow. Genealogical coherence data can be presented in the form of a table of potential ancestors. For example see Table 37.[19]

---

[18] 017 alone attests it in this dataset. Actually this reading is shared with a number of minuscules and at least one lectionary manuscript.

[19] The headings are defined as follows: **W2** – the witness to compare to 02 (W1); **NR** – the rank, or position in the table; **D** – Whether the relationship is directed or not; **PERC1** – the pre-genealogical coherence percentage; **EQ** – the number of variant units in which the two witnesses agree; **PASS** – the number of variant units in which the two witnesses are both extant; **W1<W2** and **W2>W2** – the number of prior and posterior readings in W2 respectively; **UNCL** – the number of variant units where the relationship between the readings attested by the two witnesses is undefined; **NOREL** – the number of variant units where there is no ancestral relationship between the readings attested by the two witnesses. Prior and posterior counts are recorded in the two columns labelled “W1<W2” and “W1>W2”. To remember which column is which, consider the greater than or less than sign as an arrow head in a stemma. Thus “W1<W2” is where W2 is

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
A	1		92.000	23	25	2			
03	2		86.486	32	37	3		2	
0141			86.486	32	37	2	1	2	
011		-	86.486	32	37	2	2	1	
037		-	86.111	31	36	2	2	1	
022	4		86.111	31	36	2	1	2	
031	5		83.871	26	31	2	1	2	
07		-	83.784	31	37	2	2	2	
044	6		83.784	31	37	2		4	
𝔓75	7		83.333	30	36	3	1	2	
032			83.333	30	36	2	1	3	
019			83.333	30	36	3	1	2	
063	10		82.759	24	29	2	1	2	
021		-	81.081	30	37	2	2	3	
017		-	81.081	30	37	2	2	3	
091	11		77.778	14	18	2		2	

Table 37: Potential ancestors with W1=02

A simple algorithm taking the highest ranked potential ancestor (within a set connectivity / ranking) sharing a witness' reading if such an ancestor exists, or otherwise the highest ranked potential ancestor(s) attesting the parent reading(s), allows the creation of textual flow diagrams. Since the number of witnesses is small, the reading to which each attests can be included in its label and the whole tradition represented in a single diagram for each variant unit. This would not be possible with larger datasets, but it simplifies visualisation here. The witnesses are also colour-coded by reading, again to aid visualisation. As with the CBGM's standard textual flow diagrams, the rank of the potential ancestor is shown after a slash in the descendent where it is not the first-ranked ancestor. For example see Figure 6 for the textual flow diagram of John 6:24/30 with absolute connectivity. "Absolute connectivity" means any potential ancestor is acceptable in the tree – so even a twentieth ranked ancestor would be allowed. Note that all witnesses in this diagram are connected to either their best or second best potential ancestor (depicted by /2 in the tree) apart from 01

---

before W1 in the stemma, pointing down to its descendent – and it is therefore the count of prior readings in W2.

which descends from its 11th best potential ancestor 𝔓75. So this tree would look the same if connectivity were set to eleven, but different if it were set to ten.

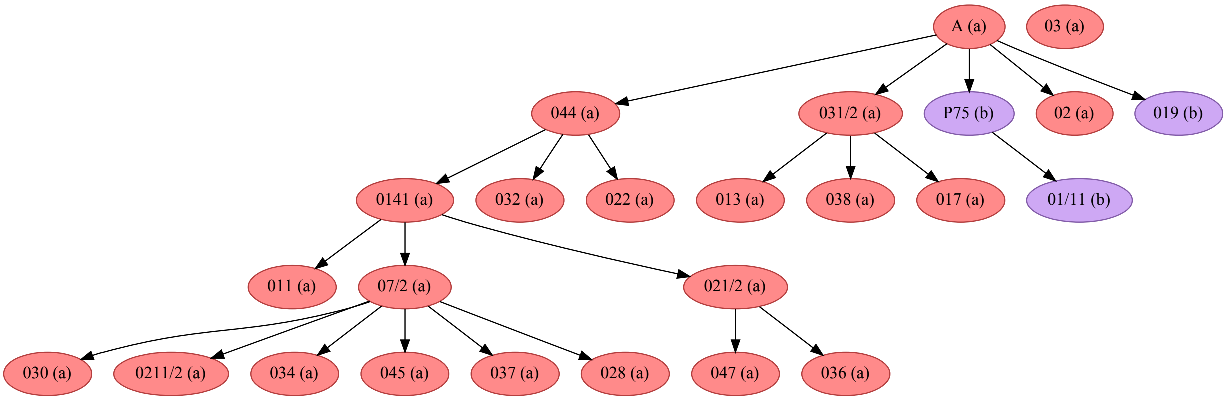


Figure 6: Textual flow diagram for John 6:24/30 with absolute connectivity

This textual flow diagram highlights three things: first variant *b* is a case of multiple emergence (which is perfectly believable as it is a change of ενεβησαν to ανεβησαν); and secondly witness 03 is disconnected from the rest of the flow.[20] The second point might normally be a cause for concern, but at this point 03 has no potential ancestors, and so textual flow will always show it separately. 091 also has no potential ancestors at this stage and neither does A. A is not expected to have ancestors, but 03 and 091 should have. This is likely to be resolved by defining further local stemmata. Note that 05, 063 and 091 are lacunose in this variant unit and so do not appear in this diagram.

Now genealogical coherence (potential ancestors) and textual flow can be considered for all unresolved variant units, of which there are fourteen. Again, only a subset of decisions is recorded here with the remainder included in Appendix 1.

John 6:23/4-10

The readings for this unit are shown in Table 38. Above we deferred any decisions about this local stemma until later – but now we must grasp the nettle.

---

[20] Thus the textual flow diagram is not a "tree", but is technically called a "forest". This normally shows that the coherence is *not* perfect and must be investigated.

a	ηλθεν πλοιαρια εκ τιβεριαδος	...
b	ηλθεν πλοια εκ τιβεριαδος	𝔓75
c	ηλθεν πλοια εκ της τιβεριαδος	03, 032
d	ηλθον πλοιαρια εκ τιβεριαδος	021, 036, 047, 091
e	ηλθον πλοιαρια εκ της τιβεριαδος	022
f	πλοιαρια ηλθον εκ τιβεριαδος	017
g	πλοια ηλθεν εκ τιβεριαδος	044
h	πλοιαρια εκ τιβεριαδος ηλθον	019
i	πλοια εκ τιβεριαδος ηλθεν	0141
	lac	01, 05

Table 38: Readings for John 6:23/4-10

First we consider πλοια versus πλοιαρια. This word, in one form or another, appears six times in this passage. The first three have no variation, reading πλοιον (21/14), πλοιον (21/22) and πλοιαριον (22/24). Then in 22/76 the attestation is split between πλοιαριον and πλοιον. The boat in question here is the same one as in 22/24 with πλοιον being the initial text, having been changed in some witnesses by assimilation (see above). Now the boats in 23/6 are the same boats that the crowd board in 24/38 – but there is no consistency as to which word is used; indeed most witnesses say "small boats" in one place and just "boats" in the other. The strongest argument in 23/6, however, is the contrary of 22/24. Here it seems likely that the initial text reads πλοιαρια (adding to the drive to assimilate in 22/24) and a handful of witnesses that did not make that assimilation then make a separate assimilation to πλοια by following their reading of πλοιον in 22/24.[21]

Now considering those readings in 23/4-10 attesting πλοια, it is then a simple matter to surmise that *b* comes from *a*, and *c* comes from *b* (the addition of the article). Reading *i* clearly comes from *g* (this is consistent with the potential ancestors of 0141 in Table 39). Reading *g* could come from *b* or *h*, but 044's potential ancestors do not contain 019 (see

---

[21] 𝔓75, 01, 02, 03, 05, 017, 019, 022, 032, 044, 091, 0141 and 0211 resisted the urge to assimilate in 22/24. Only 𝔓75, 03, 032, 044 and 0141 then attest the assimilation in 23/6.

Table 40) and even though 𝔓75 is an undirected relative this is the only remaining possibility (and this will now introduce a direction) – and so *g* is derived from *b*.

Turning to the readings attesting πλοιαρια, we hypothesise that reading *d* comes from *a*. Thus 036 and 047 both follow their first potential ancestor 021. 091 is more troublesome as it has no potential ancestors, but this is simply resolved by the hypothesis that *a* is the initial text. If this proves impossible, then we will reconsider 091. It might seem obvious that reading *e* is derived from *d*, but the potential ancestors of 022 (Table 41) strongly suggest that it comes from *a* or *c* instead – and *a* is perfectly acceptable from a text-critical perspective. The same argument for 017 implies that *f* comes from *a* via 017's highest ranked potential ancestor 031. In exactly the same way *h* can be seen to be derived from *a*.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
031		-	94.286	33	35			2	
044	1		89.189	33	37	1		3	
022		-	88.889	32	36	1	1	2	
A	2		88.000	22	25	2		1	
03	3		86.486	32	37	2		3	
063		-	86.207	25	29	1	1	2	
032		-	86.111	31	36	1	1	3	
019	4		86.111	31	36	2	1	2	
𝔓75	5		83.333	30	36	2	1	3	
091			83.333	15	18	1		2	

Table 39: Potential ancestors with W1=0141

After defining this local stemma 03 still has no potential ancestors (see Table 42). But it does have undirected genealogical relationships with A, 𝔓75 and 091. The unclear and "no relation" variant units referred to in the table are as follows: NOREL with A in John 6:23/4-10; UNCL with 091 in John 6:23/26-30; and NOREL with 091 in John 6:23/4-10. So solving this variant unit is key.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
A	1		92.000	23	25	1		1	
03	2		89.189	33	37	1		3	
019		-	88.889	32	36	1	1	2	
𝔓75		-	86.111	31	36	1	1	3	
091	3		83.333	15	18	1		2	

Table 40: Potential ancestors with W1=044

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
044	1		91.667	33	36	1		2	
03			91.667	33	36	2		1	
A			91.667	22	24	2			
032		-	88.889	32	36	1	1	2	
019	4		88.889	32	36	2	1	1	
0141		-	88.889	32	36	1	1	2	
𝔓75	5		88.571	31	35	2	1	1	
063		-	85.714	24	28	1	1	2	
091	6		77.778	14	18	2		2	

Table 41: Potential ancestors with W1=022

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
A		-	96.154	25	26				1
𝔓75		-	94.444	34	36	1	1		
091		-	88.889	16	18			1	1

Table 42: Potential ancestors with W1=03

Above we surmised that reading *c* (03) came from *b* (𝔓75) by a simple addition of the article, and that *b* came from *a* (changing πλοιαρια to πλοια). It is not unreasonable to allow *c* to be derived directly from *a* in this case, since πλοιαρια and πλοια have been so confused here. And thus 03 gains a potential ancestor in A. All these hypotheses will be checked for consistency later. The new textual flow diagram is shown in Figure 7 and the local stemma in Table 43.

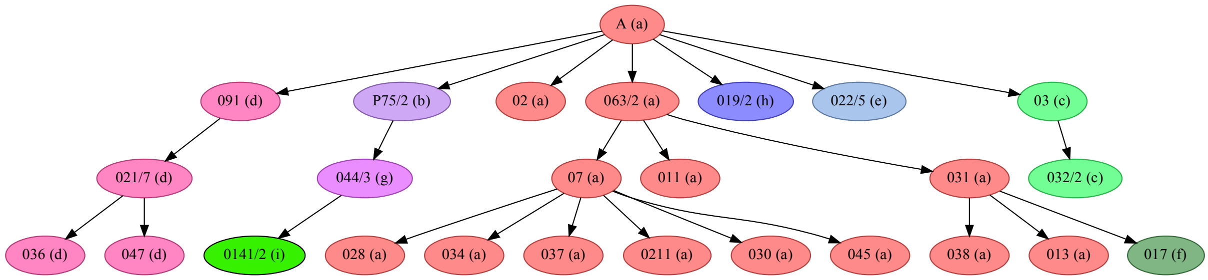
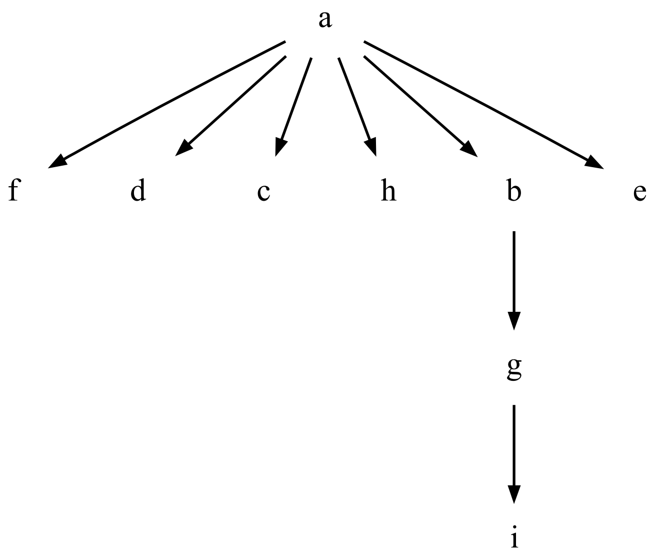


Figure 7: Textual flow diagram for John 6:23/4-10 with reading c derived from a



a	ηλθεν πλοιαρια εκ τιβεριαδος	...
b	ηλθεν πλοια εκ τιβεριαδος	𝔓75
c	ηλθεν πλοια εκ της τιβεριαδος	03, 032
d	ηλθον πλοιαρια εκ τιβεριαδος	021, 036, 047, 091
e	ηλθον πλοιαρια εκ της τιβεριαδος	022
f	πλοιαρια ηλθον εκ τιβεριαδος	017
g	πλοια ηλθεν εκ τιβεριαδος	044
h	πλοιαρια εκ τιβεριαδος ηλθον	019
i	πλοια εκ τιβεριαδος ηλθεν	0141
	lac	01, 05

Table 43: Local stemma for John 6:23/4-10

John 6:22/20

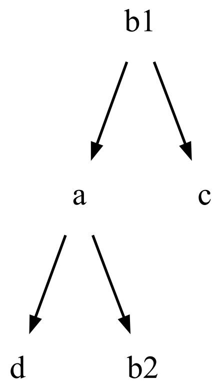
This variant unit provides an interesting challenge: relating ιδων (*a*), ειδον (*b*), ειδεν (*c*) and ειδως (*d*). We have previously identified that reading *b* is the initial text, with *c* coming from *b* and *d* coming from *a*. The attestation for *b* is problematic, however, as 038 is not closely related to any other witness attesting reading *b*. Potential ancestors for 038 are shown in Table 44.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
031	1		94.286	33	35	1		1	
045	2		92.683	38	41	2		1	
07	3		90.244	37	41	3		1	
037	4		90.000	36	40	3	1		
028	5		87.805	36	41	3	1	1	
021			87.805	36	41	3	1	1	
013			87.805	36	41	2	1	1	1
011			87.805	36	41	3		2	
063	9		86.207	25	29	3		1	
036	10		85.366	35	41	3	2	1	
034			85.366	35	41	3	2	1	
030			85.366	35	41	3	1	2	
047		-	82.927	34	41	3	3	1	
017	13		82.927	34	41	4	2	1	
0141			82.927	34	41	4		2	1
02	15		78.378	29	37	5	2	1	
A	16		76.923	20	26	6			
03	17		75.676	28	37	6	1	2	
0211	18		75.610	31	41	4	3	2	1
022	19		75.000	27	36	5	2	2	
044	20		72.973	27	37	5		4	1
𝔓75	21		72.222	26	36	6	2	2	
032			72.222	26	36	5	2	3	
019			72.222	26	36	6	2	2	
091	24		66.667	12	18	4	1	1	
05		-	58.824	20	34	4	4	5	1

Table 44: Potential ancestors with W1=038

The first witness not to attest *a* is ranked 16 (A) with a much lower percentage agreement than most of the witnesses above it. We are forced to conclude that reading *b* is a case of multiple emergence and we must split it into two readings: *b1* and *b2*. Reading *b2* must be derived from *a* – a case of a reversed variation. Also witnesses 01 and 05 are simply not related in this dataset (05 is 01’s twenty-fifth highest ranked potential ancestor out of twenty-seven, with an agreement of only 33.3%) and so reading *c* must also be a case of multiple emergence but there is no need to split the reading here as it is always derived from *b1*.

The only question remaining is whether *a* is derived from *c* or *b1*. This can be answered simply by considering textual flow diagrams for the two hypotheses. Figure 8 shows the textual flow diagram for John 6:22/20 with *b1* → *a* and connectivity=3.[22] This is consistent with the analyses above and all witnesses are connected (albeit with multiple emergence for *a*, *b1* and *b2*, and *c*). Figure 9 shows the same data but with *c* → *a* and absolute connectivity (i.e. more lenient). In this case the relationships cannot be resolved and we have a forest not a tree, showing that this hypothesis is false. Thus the local stemma can be created – see Table 45.



a	ιδων	...
b1	ειδον	𝔓75, 02, 03
b2	ειδον	038
c	ειδεν	01, 05
d	ειδως	0211
	lac	019, 022, 032, 037, 091

Table 45: Local stemma for John 6:22/20

---

[22] This diagram is the same for any connectivity setting between 2 and 24. At connectivity 25 01 descends from 05. But multiple-emergence is perfectly believable for reading *c*, so connectivity 2 is fine.

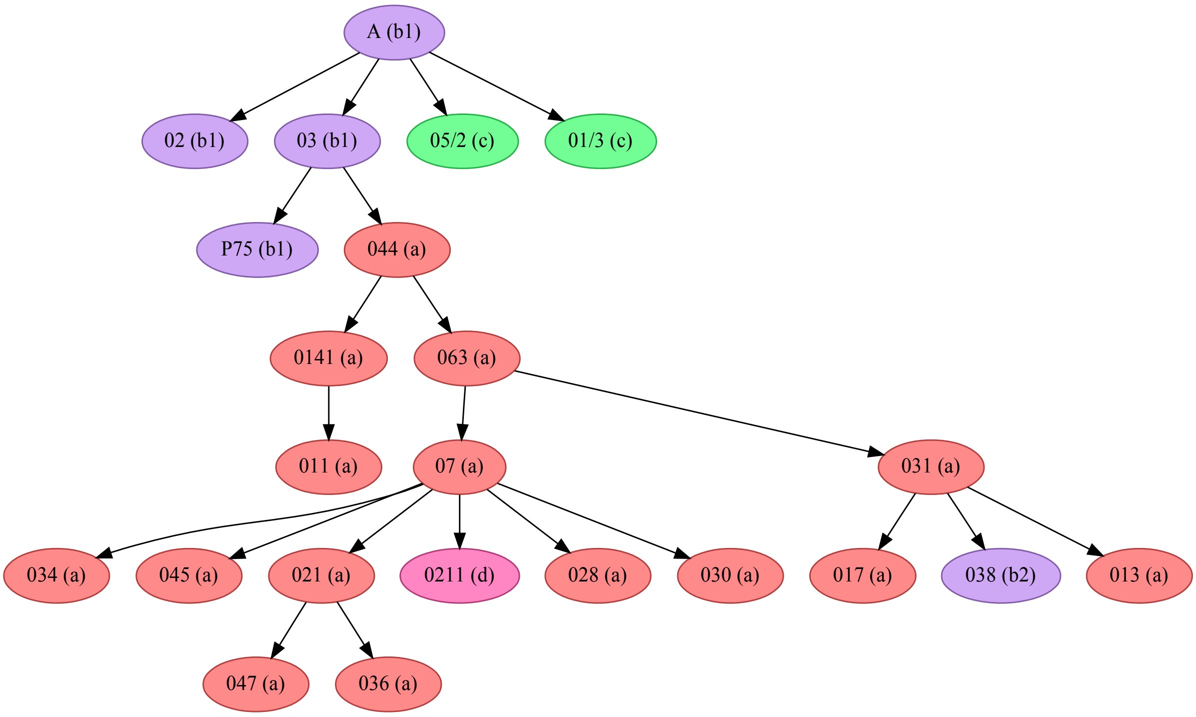


Figure 8: Textual flow diagram for John 6:22/20 with b1 → a and connectivity=3

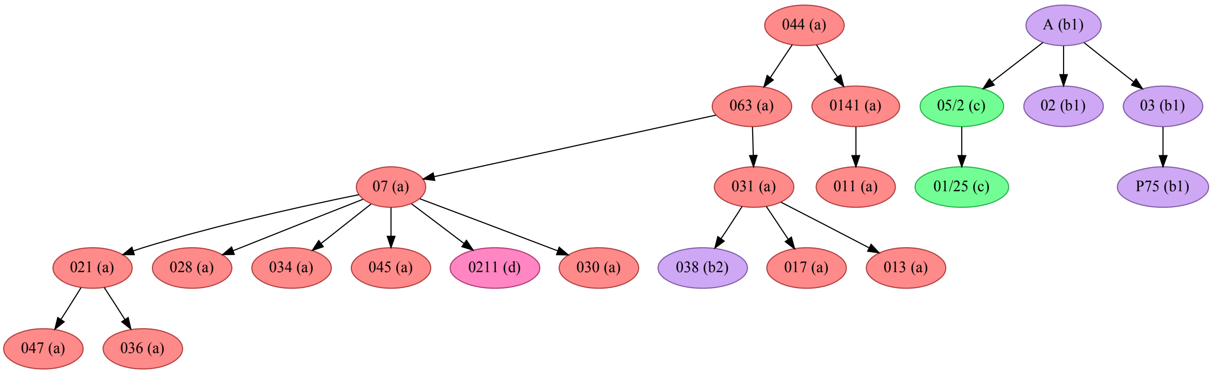


Figure 9: Textual flow diagram for John 6:22/20 with c → a and connectivity=499

<u>John 6:23/26-30</u>

This variant unit can also be solved by considering hypothetical textual flow diagrams, this time all with absolute (=499) connectivity (in other words any potential ancestor is acceptable). A hypothesis that cannot create a single tree with absolute connectivity is almost certainly false. We already know that *a* → *b*. So is reading *c* (om) the initial text, or does it derive from *a* or *b*? Figure 10 shows that *c* cannot derive from *b* – as the textual flow diagram is a forest. So all that remains is to decide which of *a* or *c* is the initial text. The two hypotheses are shown in Figure 11 and Figure 12, and the latter is clearly the only viable solution. The completed local stemma is shown in Table 46.

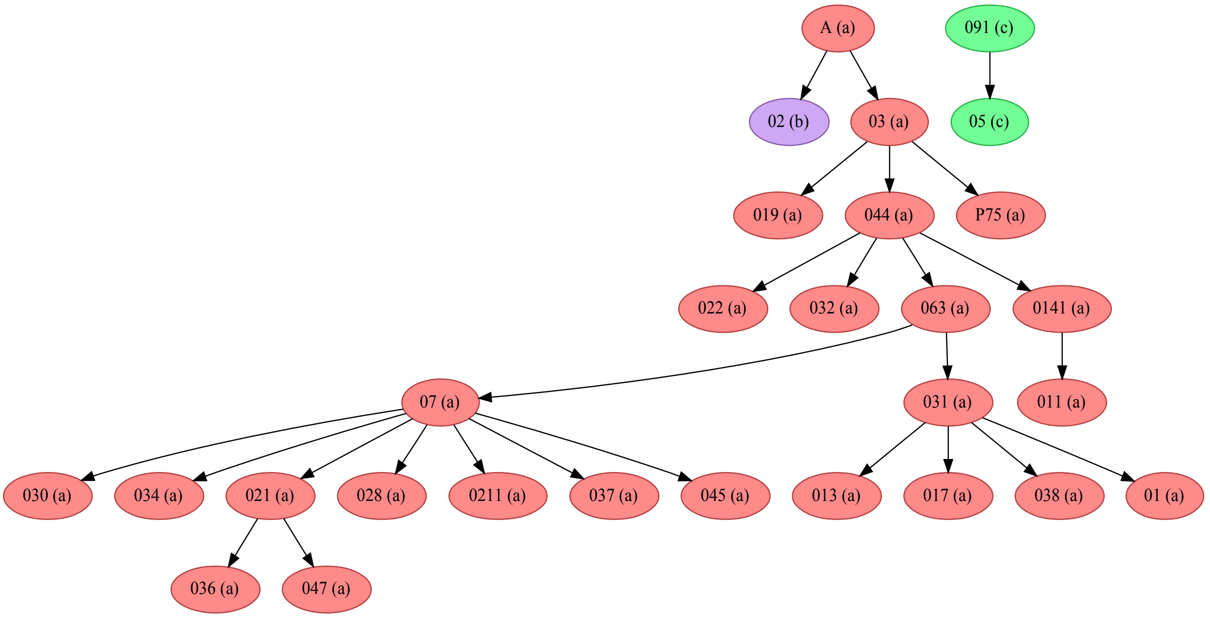


Figure 10: Textual flow diagram for John 6:23/26-30 with initial text a and b → c

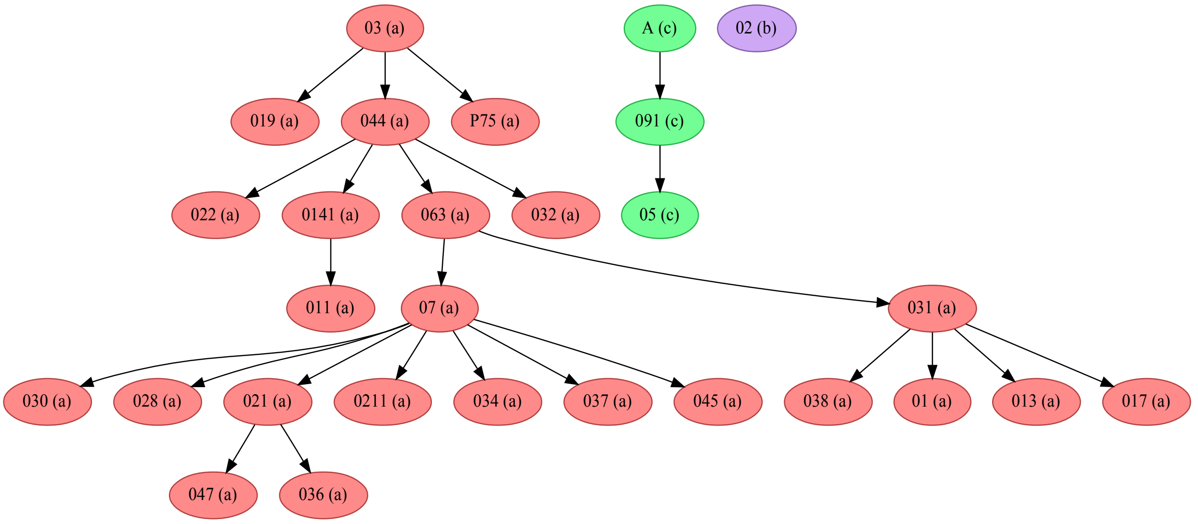


Figure 11: Textual flow diagram for John 6:23/26-30 with initial text c and c → a

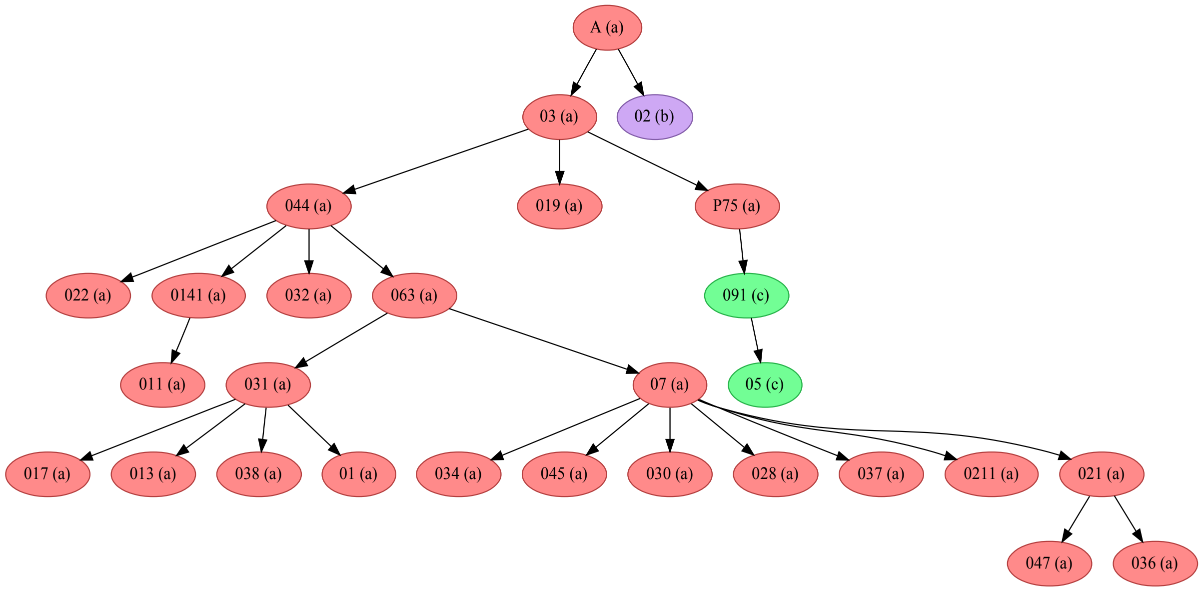
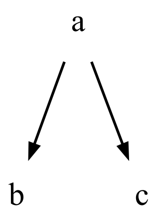


Figure 12: Textual flow diagram for John 6:23/26-30 with initial text a and a → c

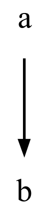


a ευχαριστησαντος του κυριου ...	
b ευχαριστησαντος του θεου	02
c	05, 091

Table 46: Local stemma for John 6:23/26-30

John 6:24/2-10

For John 6:24/2-10 the situation is more difficult as the CBGM cannot help. The choice is which of *a* and *b* is the initial text, yet the textual flow is perfect in either scenario. However, it is hard to believe that the single witness 01 (which is an outlier and usually posterior in this dataset) attests the initial text against uniformity in all others.[23] Table



a οτε ουν ειδεν ο οχλος ...	
b και ιδοντες	01

Table 47: Local stemma for John 6:24/2-10

---

[23] Again, see the "Nine Guidelines" later which state that a singular reading is normally derivative.

47 shows the local stemma.

John 6:24/14-20

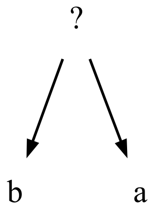
Exactly the same argument holds for this variant unit – again where 01 has a unique reading. See Table 48 for the local stemma.

a	ιησους ουκ εστιν εκει	...
b	ουκ ην εκει ο ιησους	01
	lac	031

Table 48: Local stemma for John 6:24/14-20

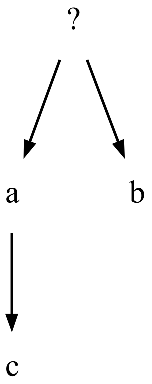
Unresolved variant units

The following five variant units remain either partially or fully unresolved. The remaining question in two is simply one of word order where multiple emergence would not pose a problem (John 6:21/6-8, Table 49 and John 6:21/20-24, Table 50):



a	λαβειν αυτον	...
b	αυτον λαβειν	05, 032, 044
	lac	091

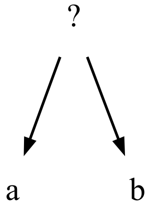
Table 49: Incomplete local stemma for John 6:21/6-8



a	το πλοιον εγενετο	...
b	εγενετο το πλοιον	𝔓75, 02, 03, 011, 019, 022, 032, 044, 0141
c	το πλοιον εγενηθη	05
	lac	091

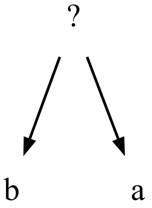
Table 50: Incomplete local stemma for John 6:21/20-24

The final three unresolved variant units present equally believable hypothetical textual flow diagrams in either possible resolution; see Table 51, Table 52 and Table 53.



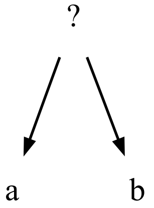
a	της γης	...
b	την γην	01, 0211
	lac	𝔓75, 091

Table 51: Incomplete local stemma for John 6:21/28-30



a	τοις μαθηταις αυτου	...
b	αυτοις	01

Table 52: Incomplete local stemma for John 6:22/62-66



a	εφαγον τον	...
b	και εφαγον	01

Table 53: Incomplete local stemma for John 6:23/20-22

Consistency of Results

Now that the local stemmata have been defined (as far as possible) it is worth checking whether there are any contradictions in the decisions. We exclude the unresolved variant units above from this process, as they cannot create connected textual flow diagrams. There are thirty-six resolved variant units. Coherent textual flow diagrams can be created for the majority of these with connectivity 1 or 2. John 6:22/20 requires connectivity 10 for witness 01. John 6:22/40 requires connectivity 9 for 05. John 6:22/76 requires connectivity 7 for 0211. 23/4-10 requires connectivity 8 for 021.

One variant unit, however, creates a disconnected textual flow diagram even at absolute connectivity: John 6:22/52 (see Figure 13), with 05 separate from the rest of the witnesses. The label "OL_PARENT" in this diagram is a device that acts like the initial text but is not the same thing as the initial text (which here is lacunose).[24] This indicates that the reading directly descending from it is the ancestral reading for this overlapping unit. This is purely an algorithmic device of my own making, and will make no appearance in the final edition. We must revisit the decisions made in that unit now.

---

[24] This is the only way to represent, in this dataset, a kind of initial text for this variant unit without implying that the actual initial text attests a reading here. The ranking of "-1" is there to emphasise that this is a hypothetical witness. To exclude this possibility would place a restriction on such variant units that does not exist for those where the initial text is not lacunose. No extant witness can serve in this role here since the two witnesses that occur beneath it (0141 and 031) have undirected genealogical coherence – and thus neither can serve as ancestor to the other.

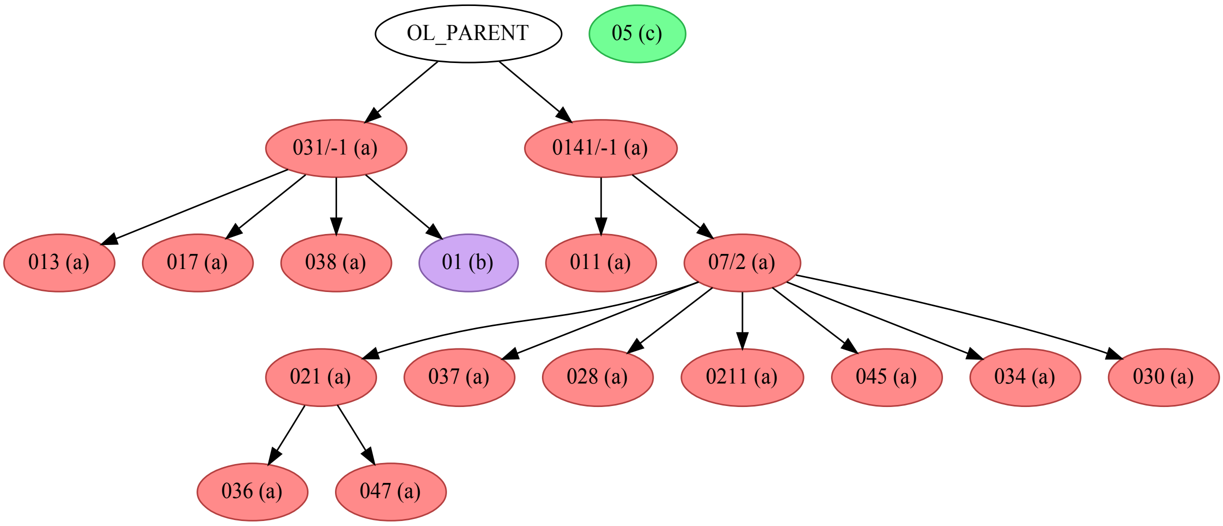


Figure 13: Textual flow diagram for John 6:22/52

Earlier we decided that 05’s reading *c* (αυτου ιησου) was derived from both the majority reading *a* (αυτου) and 01’s *b* (του ιησου). 05’s potential ancestors are shown in Table 54. 0141 is the highest ranked potential ancestor (ninth) to read anything in this variant unit. 01 is not a potential ancestor of 05 – which is the cause of the problem here. So, if the previous decision is now shown to be incompatible with the textual flow, then we must consider alternatives. 05 is 01’s twenty-fifth highest (third lowest) ranked potential ancestor, suggesting that any relationship between the two manuscripts (and their readings) is very unlikely. Therefore we are forced to conclude that *b* and *c* must derive from *a*. This allows a textual flow diagram with perfect coherence at connectivity 9 (see Figure 14). The new local stemma is shown in Table 55.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
091	1		87.500	14	16	2			
044	2		76.667	23	30	5		1	1
032	3		75.862	22	29	5	1	1	
019			75.862	22	29	5		2	
03	5		73.333	22	30	6		2	
𝔓75	6		72.414	21	29	6		2	
022			72.414	21	29	5	1	2	
A	8		72.000	18	25	7			
0141	9		70.588	24	34	6		2	2
063	10		69.231	18	26	5	1	1	1
07	11		67.647	23	34	7	1	1	2
047			67.647	23	34	6	2	1	2
036			67.647	23	34	7	1	1	2
028			67.647	23	34	7	1	1	2
0211			67.647	23	34	6	1	2	2
021			67.647	23	34	7	1	1	2
017			67.647	23	34	7	1	1	2
011			67.647	23	34	6	1	2	2
037	19		66.667	22	33	7	2	1	1
02			66.667	20	30	6	1	2	1
031	21		65.517	19	29	7		1	2
045	22		64.706	22	34	7	2	1	2
030			64.706	22	34	7	2	1	2
034	24		61.765	21	34	7	3	1	2
013			61.765	21	34	7	3	1	2
038	26		58.824	20	34	7	4	1	2

Table 54: Potential ancestors with W1=05

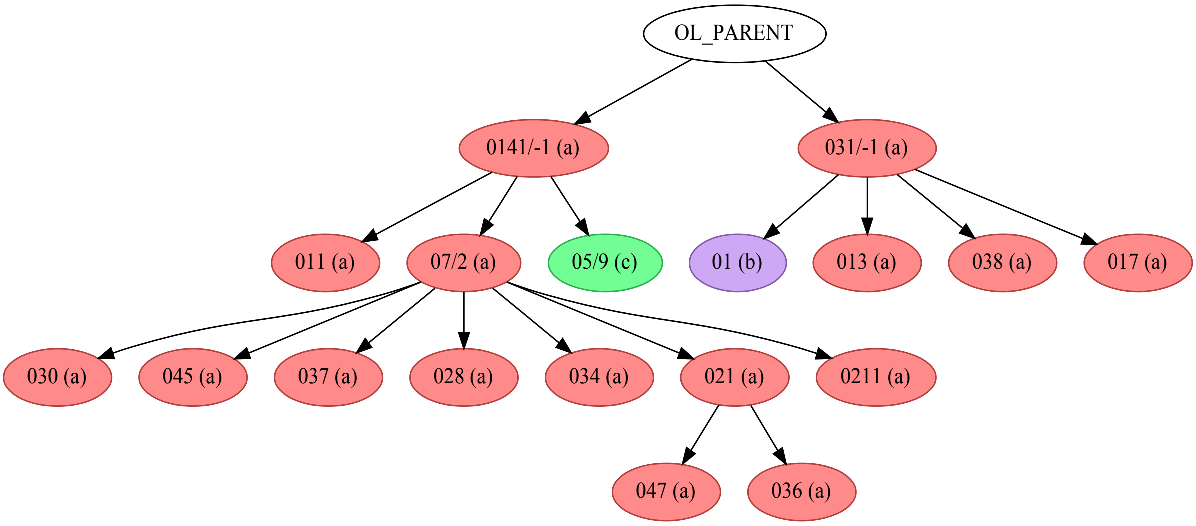
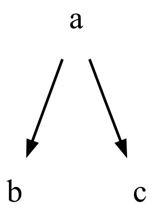


Figure 14: Textual flow diagram for John 6:22/52



a	αυτου	...
b	του ιησου	01
c	αυτου ιησου	05
	lac	𝔓75, 02, 03, 019, 022, 032, 044, 063, 091

Table 55: Local stemma for John 6:22/52

Patterns and important witnesses

Rather like the CBGM of the catholic epistles, a few witnesses emerge as often holding key places in the textual flow. In this case they are: 03 and under it 044; 063 (where it is extant) and under that 07 and 031; and sometimes 0141. As an example consider the textual flow diagrams below:

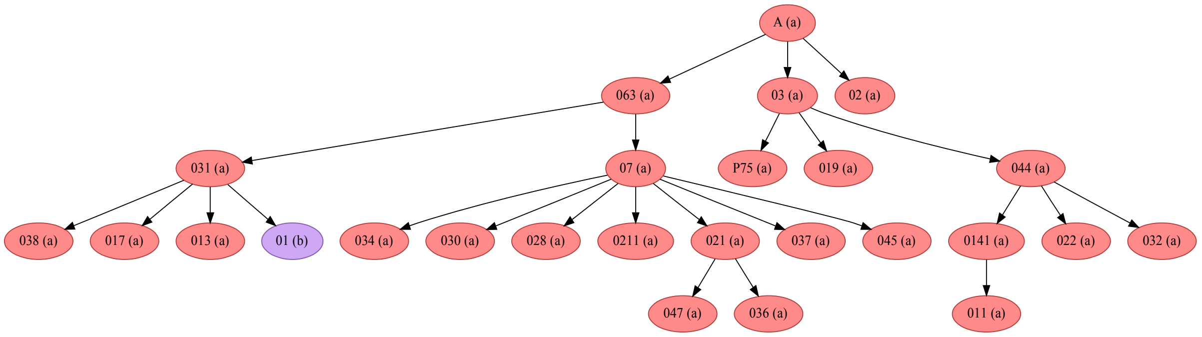


Figure 15: Textual flow diagram for John 6:24/36

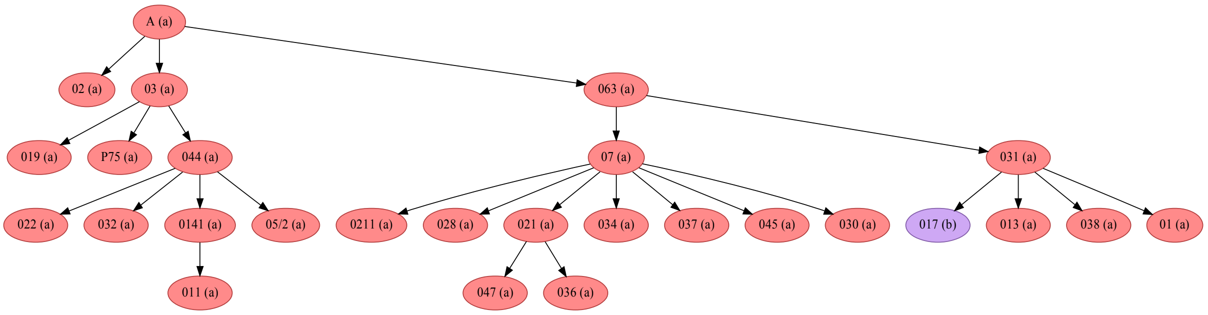


Figure 16: Textual flow diagram for John 6:24/50-52

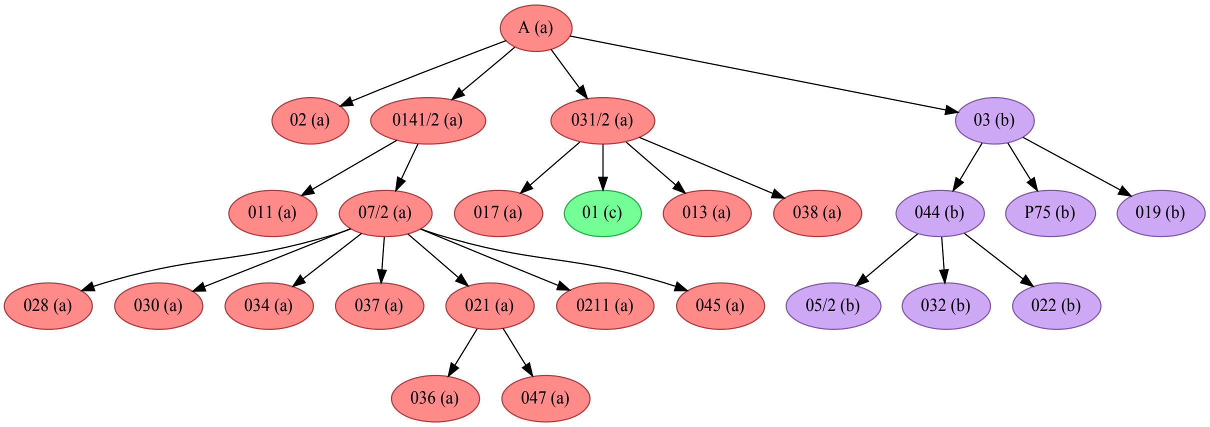


Figure 17: Textual flow diagram for John 6:24/38

The Final Apparatus

We have now (as far as is possible) established our initial text. It is as follows, with both possible readings shown in places of uncertainty. Underneath the initial text is a full positive apparatus, similar in style to that used by the ECM. For the only time in this chapter I have relabelled the readings since the ECM would always have reading *a* as the initial text in its apparatus. Some variant units required changing for the purposes of the apparatus, for example 23/1, 23/3, 23/2-10 and 23/4-10 are represented here simply as 23/2 and 23/4-10.

Unlike the ECM, however, I have not modified the word numbers (for example 22/38 is followed by 22/54 since the chosen base text contained a later insertion).

21 ηθελον ουν	λαβειν αυτον αυτον λαβειν	εις το πλοιον και ευθεως	το πλοιον εγενετο εγενετο το πλοιον	επι	της γης την γην	εις ην υπηγον
2 4	6 8	10 12 14 16 18	20 22 24	26	28 30	32 34 36
2 b ηλθον			20-24 c το πλοιον εγενηθη			36 b υπηντησεν

21/2	a	ηθελον 𝔓75, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 0141, 0211
	b	ηλθον 01
	-	091
21/6-8	a	λαβειν αυτον 𝔓75, 01, 02, 03, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 034, 036, 037, 038, 045, 047, 063, 0141, 0211
	b	αυτον λαβειν 05, 032, 044
	-	091
21/20-24	a	το πλοιον εγενετο 01, 07, 013, 017, 021, 028, 030, 031, 034, 036, 037, 038, 045, 047, 063, 0211
	b	εγενετο το πλοιον 𝔓75, 02, 03, 011, 019, 022, 032, 044, 0141
	c	το πλοιον εγενηθη 05
	-	091
21/28-30	a	της γης 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 0141
	b	την γην 01, 0211
	-	𝔓75, 091
21/36	a	υπηγον 𝔓75, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 0141, 0211
	b	υπηντησεν 01

22	τη	επαυριον	ο	οχλος	ο	εστηκως	περαν	της	θαλασσης	ειδον	οτι	πλοιαριον	αλλο	ουκ	ην	εκει	ει	μη	εν	και
	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34	36	38	54

3 b τε

10 b om.

12 b εστως

20 b ιδων
c ειδεν
d ειδως

40-52 b εκεινο εις ο ενεβησαν οι μαθηται αυτου
c εκεινο εις ο ενεβησαν οι μαθηται του ιησου
d εκεινο ο ενεβησαν οι μαθηται αυτου
e εκεινο εις ο ανεβησαν οι μαθηται αυτου
f εις ο ενεβησαν οι μαθηται αυτου
g εις ο ενεβησαν οι μαθηται αυτου ιησου

22/3 a om. 𝔓75, 01, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 0141
b τε 0211
- 091

22/10 a ο 𝔓75, 01, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 044, 047, 063, 0141, 0211
b om. 038, 045
- 091

22/12 a εστηκως 𝔓75, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 0141, 0211
b εστως 01
- 091

22/20 a ειδον 𝔓75, 02, 03, 038
b ιδων 07, 011, 013, 017, 021, 028, 030, 031, 034, 036, 044, 045, 047, 063, 0141
c ειδεν 01, 05
d ειδως 0211
- 019, 022, 032, 037, 091

22/40-52 a om. 𝔓75, 02, 03, 019, 022, 032, 044, 063
b εκεινο εις ο ενεβησαν οι μαθηται αυτου 07, 011, 013, 017, 021, 028, 030, 031, 036, 038, 045, 0141
c εκεινο εις ο ενεβησαν οι μαθηται του ιησου 01
d εκεινο ο ενεβησαν οι μαθηται αυτου 034
e εκεινο εις ο ανεβησαν οι μαθηται αυτου 037, 047
f εις ο ενεβησαν οι μαθηται αυτου 0211
g εις ο ενεβησαν οι μαθηται αυτου ιησου 05
- 091

οτι ου συνεισηλθε | τοις μαθηταις αυτου αυτοις | ο ιησους εις το πλοιον αλλα μονοι οι μαθηται αυτου απηλθον

56 58 60 62 64 66 68 70 72 74 76 78 80 82 84 86 88

60 b συνεληλυθι
68-70 b om.
76 a πλοιαριον
88 b εισηλθον
c om.

61 b ο ιησους
80 b μονον

22/60 a συνεισηλθε 𝔓75, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 091, 0141, 0211
b συνεληλυθι 01

22/61 a om. 𝔓75, 01, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 091, 0141, 0211
b ο ιησους 02

22/62-66 a τοις μαθηταις αυτου 𝔓75, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 091, 0141, 0211
b αυτοις 01

22/68-70 a ο ιησους 𝔓75, 01, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 036, 037, 038, 044, 045, 047, 063, 091, 0141, 0211
b om. 034

22/76 a πλοιον 𝔓75, 01, 02, 03, 05, 017, 019, 022, 032, 044, 091, 0141, 0211
b πλοιαριον 07, 011, 013, 021, 028, 030, 034, 036, 037, 038, 045, 047, 063
\- 031

22/80 a μονοι 𝔓75, 01, 02, 03, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 063, 091, 0141, 0211
b μονον 05, 047

22/88 a απηλθον 𝔓75, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 032, 034, 036, 037, 044, 045, 047, 063, 091, 0141, 0211
b εισηλθον 038
c om. 01
\- 031

23 αλλα ηλθεν πλοιαρια εκ τιβεριαδος εγγυς του τοπου οπου | εφαγον τον
και εφαγον | αρτον

2 4 6 8 10 12 14 16 18 20 22 24

2 b αλλα δε
c και αλλα δε
d αλλων
e om.

4-10 b ηλθεν πλοια εκ τιβεριαδος
c ηλθεν πλοια εκ της τιβεριαδος
d ηλθον πλοιαρια εκ τιβεριαδος
e ηλθον πλοιαρια εκ της τιβεριαδος
f πλοιαρια ηλθον εκ τιβεριαδος
g πλοια ηλθεν εκ τιβεριαδος
h πλοιαρια εκ τιβεριαδος ηλθον
i πλοια εκ τιβεριαδος ηλθεν
j επελθοντων ουν των πλοιων εκ τιβεριαδος 01
k πλοιαριων ελθοντων εκ τιβεριαδος 05

12-16 b εγγυς ουσης
c om.

23/2 a αλλα 𝔓75, 03, 019, 091
b αλλα δε 02, 07, 011, 013, 017, 021, 028, 030, 032, 034, 036, 037, 038, 044, 045, 047, 063, 0141, 0211
c και αλλα δε 022
d αλλων 05
e om. 01
- 031

23/4-10 a ηλθεν πλοιαρια εκ τιβεριαδος 02, 07, 011, 013, 028, 030, 031, 034, 037, 038, 045, 063, 0211
b ηλθεν πλοια εκ τιβεριαδος 𝔓75
c ηλθεν πλοια εκ της τιβεριαδος 03, 032
d ηλθον πλοιαρια εκ τιβεριαδος 021, 036, 047, 091
e ηλθον πλοιαρια εκ της τιβεριαδος 022
f πλοιαρια ηλθον εκ τιβεριαδος 017
g πλοια ηλθεν εκ τιβεριαδος 044
h πλοιαρια εκ τιβεριαδος ηλθον 019
i πλοια εκ τιβεριαδος ηλθεν 0141
j επελθοντων ουν των πλοιων εκ τιβεριαδος 01
k πλοιαριων ελθοντων εκ τιβεριαδος 05

23/12-16 a εγγυς του τοπου 𝔓75, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 034, 036, 037, 038, 044, 045, 047, 063, 091, 0141, 0211
b εγγυς ουσης 01
c om. 032

23/20-22 a εφαγον τον 𝔓75, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 091, 0141, 0211
b και εφαγον 01

ευχαριστησαντος του κυριου **24** οτε ουν ειδεν ο οχλος οτι ιησους ουκ εστιν εκει ουδε οι μαθηται αυτου

26 28 30 2 4 6 8 10 12 14 16 18 20 22 24 26 28

26-30 b ευχαριστησαντος του θεου
c om.

2-10 b οτε ουν ειπεν ο οχλος
c οτε ουν εγνω ο οχλος
d και ιδοντες

14-20 b ουκ εστιν εκει
c ο ιησους ουκ εστιν εκει
d ουκ ην εκει ο ιησους

28 b om.

23/26-30 a ευχαριστησαντος του κυριου 𝔓75, 01, 03, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 0141, 0211

b ευχαριστησαντος του θεου 02

c om. 05, 091

24/2-10 a οτε ουν ειδεν ο οχλος 𝔓75, 02, 03, 05, 07, 011, 017, 019, 021, 022, 028, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 091, 0141, 0211

b οτε ουν ειπεν ο οχλος 013

c οτε ουν εγνω ο οχλος 030

d και ιδοντες 01

24/14-20 a ιησους ουκ εστιν εκει 𝔓75, 02, 03, 05, 07, 011, 017, 019, 021, 022, 028, 030, 032, 034, 036, 037, 044, 045, 047, 063, 091, 0141, 0211

b ουκ εστιν εκει 013

c ο ιησους ουκ εστιν εκει 038

d ουκ ην εκει ο ιησους 01

\- 031

24/28 a αυτου 𝔓75, 02, 03, 05, 07, 011, 013, 017, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 0141, 0211

b om. 01

\- 063, 091

ενεβησαν αυτοι εις τα πλοια και ηλθον εις καπερναουμ ζητουντες τον ιησουν

30 32 34 36 38 40 42 44 46 48 50 52

30-38 b ενεβησαν αυτοι εις τα πλοιαρια
c ανεβησαν αυτοι εις τα πλοιαρια
d ενεβησαν και αυτοι εις τα πλοια
e τοτε και αυτοι ενεβησαν εις τα πλοια
f ανεβησαν εις το ηπλοιον
g ενεβησαν εις τα πλοια
h ελαβον εαυτοις πλοιαρια

50-52 b αυτον

30-38 a ενεβησαν αυτοι εις τα πλοια 02, 07, 011, 013, 017, 021, 028, 030, 031, 034, 036, 037, 038, 045, 047, 0141

b ενεβησαν αυτοι εις τα πλοιαρια 03, 022, 032, 044

c ανεβησαν αυτοι εις τα πλοιαρια 𝔓75, 019

d ενεβησαν και αυτοι εις τα πλοια 030, 036

e τοτε και αυτοι ενεβησαν εις τα πλοια 0211

f ανεβησαν εις το ηπλοιον 01

g ενεβησαν εις τα πλοια 028

h ελαβον εαυτοις πλοιαρια 05

\- 063, 091

24/50-52 a τον ιησουν 𝔓75, 01, 02, 03, 05, 07, 011, 013, 019, 021, 022, 028, 030, 031, 032, 034, 036, 037, 038, 044, 045, 047, 063, 0141, 0211

b αυτον 017

\- 091

Optimal Substemmata

Now we move on to the creation of optimal substemmata for each witness. It is worth recalling Mink's note about these ancestors: "Stemmatic ancestors are those which fit the rule for constructing optimal substemmata. They must not be confused with the historical exemplars of the descendant which normally are lost."[25] "The rule...", here, is that stemmatic ancestors are the necessary subset of all potential ancestors (those other witnesses that have more prior readings than posterior) required to explain all the readings in our witness. There can be many possible combinations of potential ancestors for a given witness.[26] The *optimal* substemma for that witness will contain the smallest number of necessary ancestors that explain all its readings – either by agreement or by the ancestor's reading being prior.

Table 56 shows the potential ancestors of 063 and their possible combinations are shown in Table 57 using the same table structure as in the CBGM presentation. Each combination where "Offen" is zero fully explains the readings in 063 and so we must choose the "best" combination. Mink states: "... a larger number of variants explained by agreement seems to be preferable. Yet, the larger number may result from more coincidental agreements or, more importantly, a larger combination which may contain less closely related ancestors that cause such agreements."[27] He also states: "If at a place of variation a descendant corresponds only to a witness of lesser coherence, the nature of the variant must be used to verify whether the variant in question actually links the two witnesses, or whether the correspondence of variants is only a matter of coincidence."[28] In other words, direct agreement is better, unless it requires the inclusion of a prohibitively distant relation.

---

[25] Mink, 'The Coherence-Based Genealogical Method (CBGM) – Introductory Presentation, 1.0', 133.

[26] The number of combinations is the size of the power set of potential ancestors (minus 1, the empty set). So if a witness has 15 potential ancestors, then the number of possible combinations is $|P(15)| - 1 = 2^{15} - 1 =$ 32,767. Yet for 2 potential ancestors the number of combinations is just 3.

[27] Mink, 'The Coherence-Based Genealogical Method (CBGM) – Introductory Presentation, 1.0', 509.

[28] Mink, 'Problems of a Highly Contaminated Tradition', 30.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
A	1		87.500	21	24	3			
044	2		86.207	25	29	1		2	1
0141		-	86.207	25	29	1	1	1	1
019	3		85.714	24	28	2	1	1	
03	4		82.759	24	29	3	1	1	
02		-	82.759	24	29	2	2	1	
𝔓75	5		82.143	23	28	3	1	1	
091		-	76.471	13	17	2	2		

Table 56: Potential ancestors with W1=063

In general, the smallest combination is best and so {A} alone is the first combination from Table 57 to consider because we want minimum "Offen" (unexplained by this combination) and the smallest number of witnesses (in "Vorf"/"Vorfanz"). {A} explains 21 variants by direct agreement, and three by posteriority – namely 22/20 (variants on ιδων), 22/76 (πλοιαριον versus πλοιον) and 23/3 (addition of δε). When combined with 019, 03 or 𝔓75 it explains the variants in the same way. However, combination {A, 044} (the first row in the table) explains all but 22/76 by direct agreement. A is 063's first ranked potential ancestor, with 87.5% coherence, and 044 is second with 86.207%. Now 22/20 and 23/3 are not highly connective variants, and so it could be argued that the contribution made by 044 is not valuable enough to include. However, since its coherence is nearly as high as A it seems good to include it – and minimise the number of variants needing to be explained by posteriority. Figure 18 shows the optimal substemma for 063.

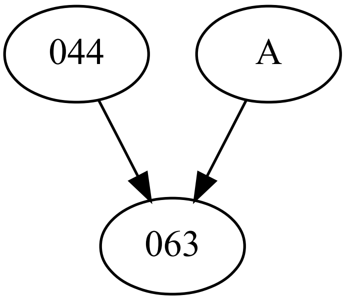


Figure 18: Optimal substemma for 063

Vorf	**Vorfanz**	**Stellen**	**Post**	**Fragl**	**Offen**	**Hinweis**
A, 044	2	23	1	5	0	<<
A, 044, 019	3	23	1	5	0	<<
A, 044, 03	3	23	1	5	0	<<
A, 044, 𝔓75	3	23	1	5	0	<<
A, 044, 019, 03	4	23	1	5	0	<<
A, 044, 019, 𝔓75	4	23	1	5	0	<<
A, 044, 03, 𝔓75	4	23	1	5	0	<<
A, 044, 019, 03, 𝔓75	5	23	1	5	0	<<
044	1	22	1	5	1	
044, 019	2	22	1	5	1	
044, 03	2	22	1	5	1	
044, 𝔓75	2	22	1	5	1	
044, 019, 03	3	22	1	5	1	
044, 019, 𝔓75	3	22	1	5	1	
044, 03, 𝔓75	3	22	1	5	1	
044, 019, 03, 𝔓75	4	22	1	5	1	
A	1	21	3	5	0	<<
A, 019	2	21	3	5	0	
A, 03	2	21	3	5	0	
A, 𝔓75	2	21	3	5	0	
A, 019, 03	3	21	3	5	0	
A, 019, 𝔓75	3	21	3	5	0	
A, 03, 𝔓75	3	21	3	5	0	
A, 019, 03, 𝔓75	4	21	3	5	0	
03	1	20	3	5	1	
𝔓75	1	20	3	5	1	
019, 03	2	20	3	5	1	
019, 𝔓75	2	20	3	5	1	
03, 𝔓75	2	20	3	5	1	
019, 03, 𝔓75	3	20	3	5	1	
019	1	20	2	5	2	

Table 57: Combinations of ancestors for 063[29]

---

[29] Definition of headings: **Vorf** – combinations; **Vorfanz** – the number of ancestors in combination; **Stellen** – the number of variants explained by agreement with an ancestor (table sorted by this); **Post** – the number of variants explained by direct posterity (i.e. one ancestor has the parent reading); **Fragl** – the number of cases of unknown source of variant (i.e. incomplete local stemma); **Offen** – the number of cases not explained by this combination; **Hinweis** – pointer "<<", indicating the best combination of ancestors compared with other combinations which are equal in number of ancestors. See Mink, 'The Coherence-Based Genealogical Method (CBGM) – Introductory Presentation, 1.0', 504.

Vorf	**Vorfanz**	**Stellen**	**Post**	**Fragl**	**Offen**	**Hinweis**
A	1	30	2	5	0	<<

Table 58: Combinations of ancestors for 03

Working out the optimal substemmata can be very easy, or more difficult. For example, Table 58 shows that there is only one possible combination of ancestors for 03, namely {A} (for A is 03's only potential ancestor) which explains all 03's readings.

Vorf	**Vorfanz**	**Stellen**	**Post**	**Fragl**	**Offen**	**Hinweis**
063, A	2	25	1	5	4	<<
A	1	24	2	5	4	<<
063	1	18	1	5	11	

Table 59: Combinations of ancestors for 031

Others are only slightly more complex, like 031 whose combinations of ancestors are shown in Table 59. 063 is 031's closest ancestor, but does not explain eleven of 031's variant units (either by agreement or direct posteriority). Unfortunately, no combination explains all 031's readings, as none of its ancestors contain the large addition John 6:22/42-52 (and its constituent four smaller units), which is attested by 031. {A} explains twenty-four by agreement and two by posteriority. Here all that is needed is to determine whether 063 should be added to A in order to explain John 6:22/20 by agreement rather than posterity. 031's potential ancestors are shown in Table 60. 063 is 031's closest ancestor and so it seems prudent (switching the order for clarity) to require 063 and include A to explain the variant units that 063 cannot. Therefore {063, A} is the best combination of ancestors – even though it cannot explain everything. The hypothesis, then, is that 031 represents the genesis of the addition in John 6:22/40-52 although in reality this almost certainly occurred in a manuscript now lost.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
07		-	100.000	35	35				
011		-	97.143	34	35			1	
063	1		95.833	23	24	1			
0141		-	94.286	33	35			1	1
A	2		92.308	24	26	2			
044		-	83.871	26	31	1	1	2	1
03		-	83.871	26	31	2	2	1	
02		-	83.871	26	31	2	2	1	

Table 60: Potential ancestors with W1=031

Some witnesses have very large lists of possible combinations. 01, for example, has 134,217,728. Nevertheless, an algorithm that compares the shortest combinations first means that we can be confident that we will find the best combination of ancestors even if we only generate the first, say, 100,000 for analysis. 01 has two combinations of ancestors that best explain its readings: {028, 038, 𝔓75, 05}, {028, 038, 019, 05}. But these both require that 05 be included – which is 01's twenty-fifth closest ancestor with only 33.3% coherence. However, there is a single combination of size three {028, 038, 𝔓75} that explains only one more reading by posteriority, namely John 6:22/20.[30] This is a change from ειδον to ειδεν – which is far better than having to include such a distant relative as 05. We could repeat this step and remove 038 thus explaining one further reading by posteriority. 038 is 01's third ranked potential ancestor and normally this might be enough to keep it – but the percentage agreement is only 44.737. The variant in question is the addition of the article before ιησους in John 6:24/14 and as such we will accept an explanation by posteriority here too. So 01's optimal substemma consists of just {028, 𝔓75}.

---

[30] 028, 038, 𝔓75 are the second, joint third and joint tenth closest potential ancestors respectively.

Global Stemma

Including those in Appendix 1, the optimal substemmata have all been created, and the global stemma can be drawn. It is shown in Figure 19. The dashed lines represent those places where it was impossible to choose between different optimal substemmata.

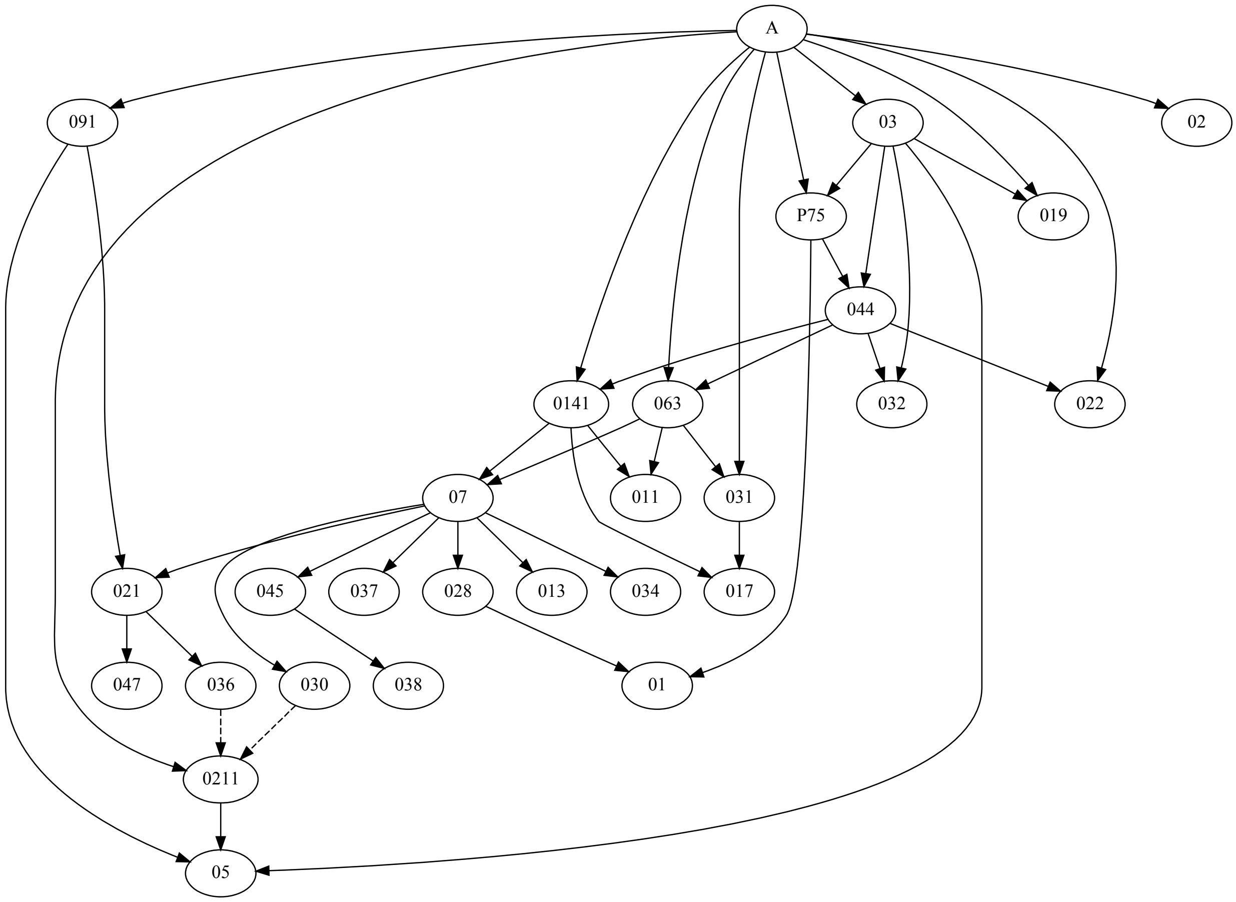


Figure 19: Global Stemma

The Effects of Collation Decisions

Klaus Wachtel kindly reviewed a draft of this chapter and, along with many other helpful comments, argued for some different collation decisions. In this section I will explore how the various stages of the CBGM would have been different if I had made those decisions in the first place. Specifically, Wachtel's suggestions involve my decisions regarding some of the overlapping variant units in v23 and v24. These changes make no difference to the witnesses that are considered too fragmentary or were previously identical (with respect to their readings in the defined variant units) and so 𝔓28, 09 and 039 are still excluded.

John 6:23/2-10

In v23 I defined variant units for words 2-10 and also 4-10. Wachtel argues for a single variant unit for words 2-10 with no overlaps. This converts my four variant units 23/1, 23/3, 23/2-10 and 23/4-10 into one variant unit, as shown in Table 61. 031 is partially lacunose in such a way as it could support several readings – and therefore 031 must also be considered as totally lacunose for this variant unit.

a	αλλα ηλθεν πλοιαρια εκ τιβεριαδος	No witnesses
b	αλλα ηλθεν πλοια εκ της τιβεριαδος	03
c	αλλα ηλθεν πλοια εκ τιβεριαδος	𝔓75
d	αλλα πλοιαρια εκ τιβεριαδος	019
e	αλλα ηλθον πλοιαρια εκ τιβεριαδος	091
f	αλλα δε ηλθεν πλοιαρια εκ τιβεριαδος	...
g	αλλα δε ηλθεν πλοια εκ της τιβεριαδος	032
h	αλλα δε πλοια εκ τιβεριαδος	0141
i	αλλα δε ηλθον πλοιαρια εκ τιβεριαδος	021, 036, 047
j	αλλα δε πλοιαρια ηλθον εκ τιβεριαδος	017
k	αλλα δε πλοια ηλθεν εκ τιβεριαδος	044
l	και αλλα δε ηλθον πλοιαρια εκ της τιβεριαδος	022
m	αλλων πλοιαριων ελθοντων εκ τιβεριαδος	05
n	επελθοντων ουν των πλοιων εκ τιβεριαδος	01
	lac	031

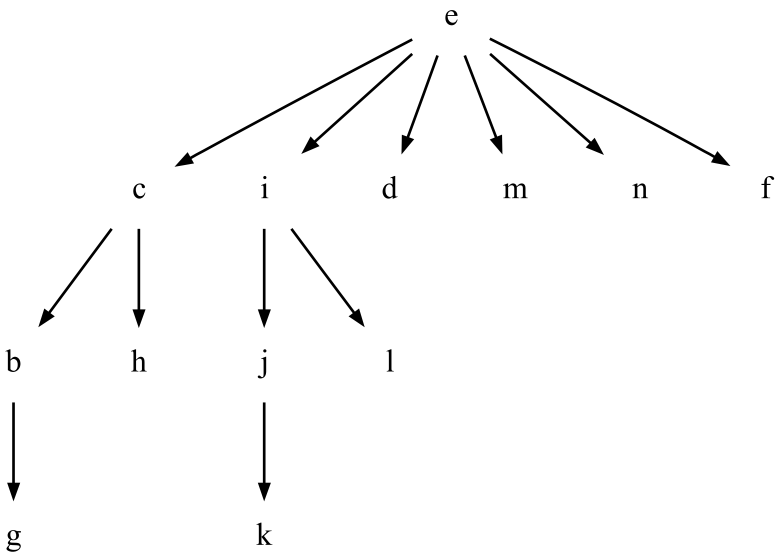
Table 61: Readings for the new variant unit John 6:23/2-10

The first notable feature of Table 61 is the "No witnesses" for reading *a*. This highlights an intriguing fact about critical editions: the text they show as their running line may never have been contained in a real manuscript. The running line here is the *Textus Receptus*, and it is not difficult to imagine no papyri or majuscules attesting it. However, when it is broken into smaller parts (as in my collation earlier) each part is attested to by one witness or another. So considering small variant units hides this fact about critical texts.

The numerical differences of note here are that we now have a single variant unit instead of four; and this unit has fourteen different readings as opposed to a maximum of nine in my collation. This section already contained the most difficult variant units to collate and the most difficult local stemmata to create – to some extent due to the large number of different readings in John 6:23/4-10. The task is now even harder. It is necessary now to consider the changes between ηλθεν and ηλθον. Since both are legitimate forms and differ in only one letter, it seems reasonable to give changes in this word less weight than others and thus allow it to change between otherwise dependant readings. Otherwise, applying the same arguments as above for constructing the local stemma we see the following:

- The addition of δε: *a* → *f*, *b* → *g*, *c* → *h* (including the omission of ηλθεν), *e* → *i*;
- The change from πλοιαρια to πλοια: *a* → *c*, *j* → *k* (including ηλθεν → ηλθον);
- The addition of της: *c* → *b*, *i* → *l* (including the addition of και);
- The change from ηλθον πλοιαρια → πλοιαρια ηλθον: *i* → *j*;
- Omission of ηλθον: *a* or *e* → *d*;
- The genitive absolutes attested to by 01 and 05 (which we noted above are surely unrelated due to their very low pre-genealogical coherence) are so different from the rest of the readings that it is difficult to define a relationship. The only realistic choice is to say their readings here stem from the initial text (*a*), diverging very early from the rest of the tradition.

Now, since reading *a* is not attested by any of our witnesses, it is logical that we try to avoid a conjecture and thus avoid it being the initial text. Even including it as an intermediate stage might affect genealogical relationships, and so we will now collapse it out of the local stemma and let reading *e* take its place: for reading *e* differs from *a* only inasmuch as it reads ηλθον and not ηλθεν. The local stemma is shown in Table 62.



b	αλλα ηλθεν πλοια εκ της τιβεριαδος	03
c	αλλα ηλθεν πλοια εκ τιβεριαδος	𝔓75
d	αλλα πλοιαρια εκ τιβεριαδος	019
e	αλλα ηλθον πλοιαρια εκ τιβεριαδος	091
f	αλλα δε ηλθεν πλοιαρια εκ τιβεριαδος	...
g	αλλα δε ηλθεν πλοια εκ της τιβεριαδος	032
h	αλλα δε πλοια εκ τιβεριαδος	0141
i	αλλα δε ηλθον πλοιαρια εκ τιβεριαδος	021, 036, 047
j	αλλα δε πλοιαρια ηλθον εκ τιβεριαδος	017
k	αλλα δε πλοια ηλθεν εκ τιβεριαδος	044
l	και αλλα δε ηλθον πλοιαρια εκ της τιβεριαδος	022
m	αλλων πλοιαριων ελθοντων εκ τιβεριαδος	05
n	επελθοντων ουν των πλοιων εκ τιβεριαδος	01
	lac	031

Table 62: Local stemma for the new variant unit John 6:23/2-10

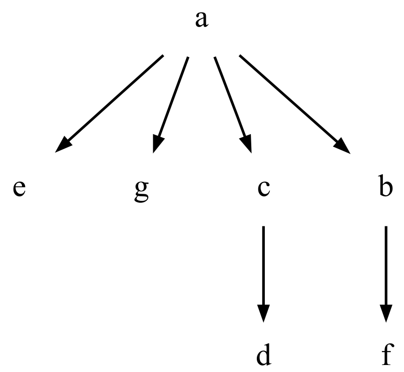
<u>John 6:24/30-32</u>

Here Wachtel suggests creating a single variant unit of 24/30-32. This has the effect of eliminating my variant units 24/30, 24/31, 24/32 and 24/30-38. Again, this reduces the number of variant units by introducing a significantly more complex unit. Here 091 is lacunose and 063 is sufficiently lacunose to offer no help.

Once again using the arguments from when we constructed the original local stemmata (see the tables in Appendix 1 and Table 33 above), we can say that:

- ενεβησαν → ανεβησαν: *a* → *g*, *b* → *f*;
- Addition of και: *a* → *c*;
- Omission of αυτοι: *a* → *b*;
- Addition of τοτε and change of word order: *c* → *d*;
- 05's unusual reading *e* is best derived from *a*.

The resulting local stemma is shown in Table 63.

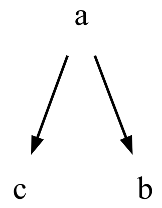


a	ενεβησαν αυτοι	...
b	ενεβησαν	028
c	ενεβησαν και αυτοι	030, 036
d	τοτε και αυτοι ενεβησαν	0211
e	ελαβον εαυτοις	05
f	ανεβησαν	01
g	ανεβησαν αυτοι	𝔓75, 019
	lac	063, 091

Table 63: Local stemma for the new variant unit John 6:24/30-32

<u>John 6:24/34-36</u>

Removal of the previous overlapping unit 24/30-38 means that we must replace 24/36 with a new unit 24/34-36 to accommodate 05's omission. Here it can simply be said that readings *b* and *c* both derive from *a* (since 05 and 01 are unrelated). The local stemma is shown in Table 64. The total number of variant units is now 34 as opposed to 41.



a	εις τα ...	
b	εις το	01
c	om	05
	lac	091

Table 64: Local stemma for the new variant unit John 6:24/34-36

The Effects of Collation Decisions

Pre-genealogical Coherence

Pre-genealogical coherence data will be different with this new collation. The question is whether this will make a *significant* difference. Consider the pre-genealogical coherence for 02: old collation in Table 65 and new collation in Table 66. There are a number of notable changes. For example 022 has been promoted from joint fifth to third place, and 𝔓75 from joint tenth to sixth. The very top and bottom of the tables is essentially unchanged, in this case.

We used pre-genealogical coherence to help determine the local stemmata of John 6:22/80, where 05 and 047 were being considered to determine whether a reading was a case of multiple emergence. Previously 047 was 05's joint tenth closest relative (with 67.647% coherence) and 05 was 047's twenty-fifth closest (with the same 67.647% coherence). Now 047 is 05's joint twelfth closest relative (with 62.857% coherence) and 05 is 047's twenty-sixth closest (with the same 62.857% coherence). So the decision is unchanged.

Pre-genealogical coherence data was also used in John 6:23/26-30, where we decided that 05 and 091's close relationship implied that their shared reading was not a case of multiple emergence. 05 and 091 are still very closely related (with 87.500% coherence) and again, this decision is unchanged.

In John 6:24/30 pre-genealogical coherence was considered briefly, before the variant was determined as having little connectivity – i.e. multiple emergence of that unit was easily conceivable. This unit no longer exists in our new collation.

So in this dataset, the change in pre-genealogical coherence has not created any difficulties. But it is easy to imagine a scenario in which decisions that were justified by one collation could be challenged by another. This highlights the importance of collating carefully, and also adds weight to the CBGM's claim that it will show what the user has critically defined rather than merely creating a mechanistic result.

W2	NR	PERC1	EQ	PASS
A	1	90.625	29	32
03	2	86.486	32	37
0141		86.486	32	37
011		86.486	32	37
037	5	86.111	31	36
022		86.111	31	36
031	7	83.871	26	31
07	8	83.784	31	37
044		83.784	31	37
𝔓75	10	83.333	30	36
032		83.333	30	36
019		83.333	30	36
063	13	82.759	24	29
045	14	81.081	30	37
034		81.081	30	37
028		81.081	30	37
021		81.081	30	37
017		81.081	30	37
047	19	78.378	29	37
036		78.378	29	37
030		78.378	29	37
013		78.378	29	37
091	23	77.778	14	18
038	24	75.676	28	37
0211		75.676	28	37
05	26	66.667	20	30
01	27	32.353	11	34

*Table 65: Pre-genealogical coherence with W1=02 (**old** collation)*

W2	NR	PERC1	EQ	PASS
A	1	88.462	23	26
03	2	87.097	27	31
022	3	86.667	26	30
0141	4	83.871	26	31
011		83.871	26	31
𝔓75	6	83.333	25	30
037		83.333	25	30
019		83.333	25	30
063	9	80.769	21	26
031		80.769	21	26
07	11	80.645	25	31
044		80.645	25	31
032	13	80.000	24	30
091		80.000	12	15
045	15	77.419	24	31
034		77.419	24	31
028		77.419	24	31
021		77.419	24	31
017		77.419	24	31
047	20	74.194	23	31
036		74.194	23	31
030		74.194	23	31
0211		74.194	23	31
013		74.194	23	31
038	25	70.968	22	31
05	26	61.290	19	31
01	27	26.667	8	30

*Table 66: Pre-genealogical coherence with W1=02 (**new** collation)*

Textual Flow and Consistency

Genealogical coherence data will have changed just as pre-genealogical coherence data did. When checking the consistency of all our decisions above we found that all our variant units had perfect coherence with connectivity of 8 or less. Now, two units have potential problems:

First, in John 6:24/38 reading *b* is now a case of multiple emergence (see Figure 20). This is not hard to believe however, in context, since πλοια (*a*) and πλοιαρια (*b*) are so confused in this passage.

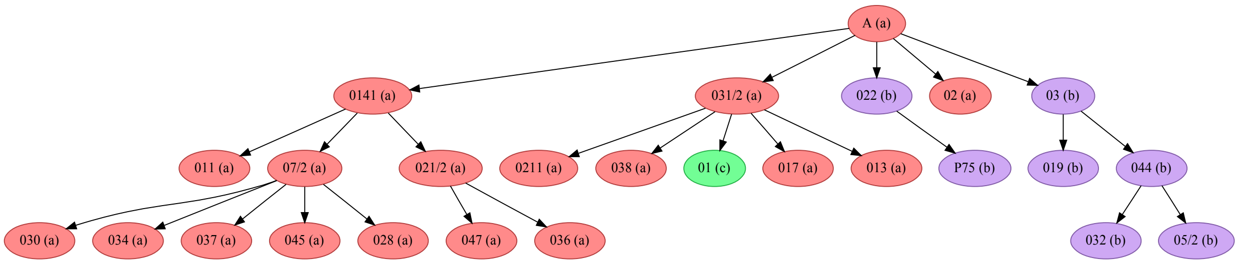


Figure 20: Textual flow diagram for John 6:24/38

Second, and more problematic, is John 6:23/2-10 (our new complicated variant unit). Its textual flow diagram (with absolute connectivity) is shown in Figure 21. This shows disconnected textual flow, and therefore suggests the local stemma is wrong. The diagram shows that reading *f* is a case of multiple emergence, and readings *b*, *k* and *l* have no viable parents.

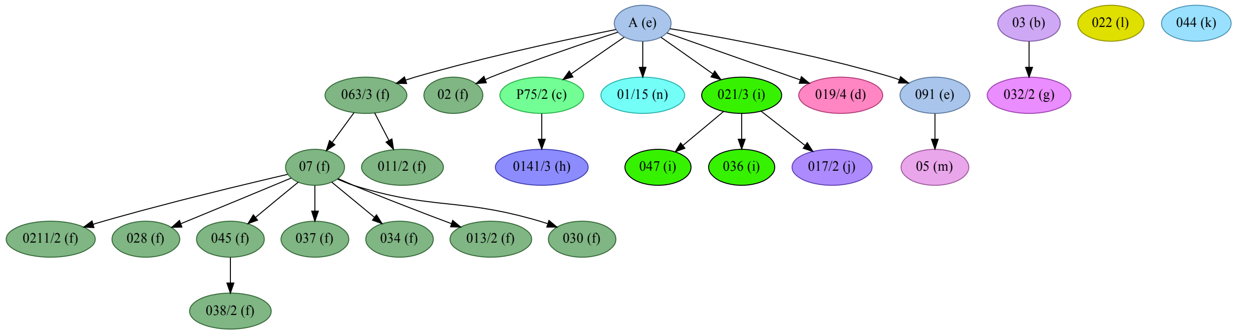


Figure 21: Textual flow diagram for John 6:23/2-10

Table 62 on page 134 shows the readings and local stemma we constructed for John 6:23/2-10. We will consider the readings individually. First we consider reading *f*, one of several readings including an added δε. While αλλα δε is unusual in John's Gospel this does not seem like a connective variant and so multiple emergence is acceptable.[31]

Next we turn to reading *b* – 03's addition of της to reading *c*. Table 67 shows the

---

[31] See discussion of John 6:23/3 above.

potential ancestors for 03 – namely A alone (reading *e*). 𝔓75 attests reading *c*, but has an undirected genealogical relationship with 03. All that would be required for this relationship to become directed (recall our criteria for this example of allowing weak textual flow to be meaningful) is for one more variant unit to show flow from 𝔓75 to 03; making 𝔓75 a potential ancestor of 03 and thus supporting the argument that *b* descends from *c*. In the table, the W1>W2 entry is for this variant unit, where the W1<W2 entry is for John 6:24/30-32 where 03 reads ενεβησαν and 𝔓75 the posterior form ανεβησαν (see Table 63 above). But what is to be done? These are the only two places (in the new collation) where these manuscripts disagree. In a larger corpus it would be hoped that another manuscript attesting the desired reading might come to the rescue with a stronger textual flow. But in this case should we change the decision on this unit or the other? The easiest answer would be if reading *b* here could derive from reading *e* – and therefore from 03's only available potential ancestor – but that would require three separate modifications to the reading: ηλθον to ηλθεν; πλοιαρια to πλοια; and the addition of της. That is too hard to accept, and so we turn to Mink's advice on what to do in these circumstances, when prior variants are not found in potential ancestors.[32]

W2	NR	D	PERC1	EQ	PASS	READING	W1<W2	W1>W2	UNCL	NOREL
022		-	96.667	29	30	l				1
𝔓75		-	93.333	28	30	c	1	1		
A	1		92.308	24	26	e	2			
091		-	86.667	13	15	e	1	1		

Table 67: Potential ancestors with W1=03 showing their reading in John 6:23/2-10

In Mink's seminal paper of 2004 'Problems of a highly contaminated tradition' he addresses the problem of what to do if the source variant of a reading is only found in non-ancestors. He describes exactly our problem: "If the tradition is dense and every witness has a

---

[32] See Mink, 'Problems of a Highly Contaminated Tradition', 59–63.

series of closely related potential ancestors, it will be easy to find the prior variants among them which explain the posterior variants in a descendant. Generally speaking, this is also true for a less dense tradition, but in such a tradition the prior variant which corresponds to a posterior one may only occur in one witness, which is not a potential ancestor."[33] The solution is to introduce an "intermediary node" which contains only this variant unit.[34] Note that this is distinctly not a hyparchetype, as it is not a hypothetical witness but instead simply represents the fact that 𝔓75 contains a reading which is the source of a variant in 03 – even though 𝔓75 is not a potential ancestor of 03. This is a way of encoding the presence of contamination where no extant witnesses can show the development of the text in this place. It is important to note that this new witness (called "03<𝔓75") highlights the existence of this contamination without implying that it must have happened in this exact place in the stemma.

In tables of potential ancestors, this intermediary node shows up in one of three places. For 𝔓75 it has undirected genealogical coherence with 100% pre-genealogical coherence. For all other witnesses (apart from 03) it has undirected genealogical coherence with 0% pre-genealogical coherence. For 03 it appears as a potential ancestor but with 0% pre-genealogical coherence (see Table 68). This basically makes it a potential ancestor of last resort for 03. To highlight how this works, the optimal substemma of 03 is shown in Figure 22.

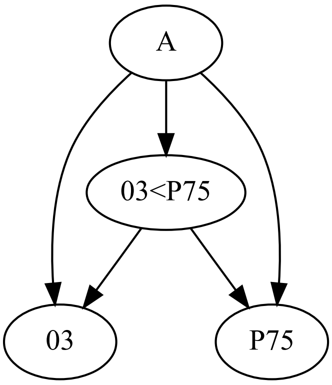


Figure 22: Optimal substemma for 03

---

[33] Ibid., 59.
[34] Ibid., 60.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
022		-	96.667	29	30				1
𝔓75		-	93.333	28	30	1	1		
A	1		92.308	24	26	1			
03<𝔓75	2		0.000		1	1			

Table 68: Potential ancestors with W1=03 including 03<𝔓75

John 6:23/2-10's reading *k* is apparently straightforward from a text-critical perspective. 044 reads αλλα δε πλοια ηλθεν εκ τιβεριαδος, which seemingly descends from 017's reading αλλα δε πλοιαρια ηλθον εκ τιβεριαδος. However, as can be seen from Table 69, not only do 044 and 017 have undirected genealogical coherence, they are also relatively distantly related. If 044 were to descend from readings *b* or *e* then the coherence data would be supportive. It is plausible that *k* could descend from *e* – with two words being swapped and the addition of δε – and so we will accept this hypothesis. If we had more witnesses available then perhaps reading *j* would be a potential ancestor. That would seem better, but our hypothesis is acceptable for now. Following this decision, 017 disappears from the table of potential ancestors for 044.

W2	NR	D	PERC1	EQ	PASS	READING	W1<W2	W1>W2	UNCL	NOREL
022		-	93.333	28	30	l			1	1
03	1		90.323	28	31	b	1		1	1
A	2		88.462	23	26	e	2			1
0141		-	87.097	27	31	h	1	1	1	1
𝔓75		-	86.667	26	30	c	1	1	1	1
031		-	84.615	22	26		1	1	2	
02		-	80.645	25	31	f	2	2	1	1
017		-	80.645	25	31	j	2	2	2	
03<𝔓75		-	0.000		1	c				1

Table 69: Potential ancestors with W1=044 showing readings from John 6:23/2-10

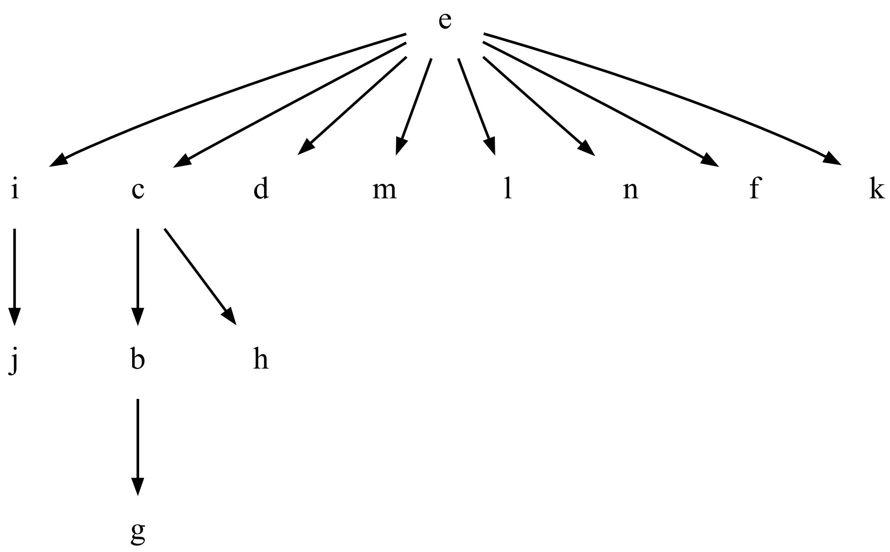
Finally, we must consider variant *l*, attested to by 022. Table 70 shows that 022 has only one potential ancestor: A. So could *l* be descended from the initial text *e*, witnessed to by 091? Reading *e* is αλλα ηλθον πλοιαρια εκ τιβεριαδος, and *l* is και αλλα δε ηλθον πλοιαρια εκ της τιβεριαδος. In other words, *l* would be a trio of additions: ϰαι, δε and της. This is the only variant to include ϰαι, and so that part is not genealogically informative. We have already accepted that δε was added independently in several witnesses, and the addition of the article could likewise have emerged multiple times. So, while not perhaps ideal, it is believable that *l* emerged from *e* as far as the CBGM goes.[35] If this scenario was less plausible, then we would have had to resort to another intermediary node. The new local stemma for John 6:23/2-10 is shown in Table 71.

W2	**NR**	**D**	**PERC1**	**EQ**	**PASS**	**W1<W2**	**W1>W2**	**UNCL**	**NOREL**
03		-	96.667	29	30				1
044		-	93.333	28	30			1	1
A	1		92.000	23	25	1			1
0141		-	90.000	27	30	1	1		1
031		-	88.000	22	25	1	1	1	
021		-	83.333	25	30	2	2	1	
03<𝔓75		-	0.000		1				1

Table 70: Potential ancestors with W1=022

---

[35] This is one situation where it seems highly likely that the CBGM does not represent reality. It is probable that witnesses have been lost that would have better explained the evolution of this variant unit – but without them this is the best we can do.



b	αλλα ηλθεν πλοια εκ της τιβεριαδος	03
c	αλλα ηλθεν πλοια εκ τιβεριαδος	𝔓75
d	αλλα πλοιαρια εκ τιβεριαδος	019
e	αλλα ηλθον πλοιαρια εκ τιβεριαδος	091
f	αλλα δε ηλθεν πλοιαρια εκ τιβεριαδος	…
g	αλλα δε ηλθεν πλοια εκ της τιβεριαδος	032
h	αλλα δε πλοια εκ τιβεριαδος	0141
i	αλλα δε ηλθον πλοιαρια εκ τιβεριαδος	021, 036, 047
j	αλλα δε πλοιαρια ηλθον εκ τιβεριαδος	017
k	αλλα δε πλοια ηλθεν εκ τιβεριαδος	044
l	και αλλα δε ηλθον πλοιαρια εκ της τιβεριαδος	022
m	αλλων πλοιαριων ελθοντων εκ τιβεριαδος	05
n	επελθοντων ουν των πλοιων εκ τιβεριαδος	01
	lac	031

Table 71: Updated local stemma for the new variant unit John 6:23/2-10

Optimal Substemmata and the Global Stemma

The potential ancestors have changed for a number of witnesses with this new collation, and as such the optimal substemmata are also likely to change. Consider 𝔓75: Table 72 shows the potential ancestors in the old collation, and Table 73 in the new collation. Its optimal substemma previously consisted of {03, A} but now 03 is not even a potential ancestor. Table 74 shows the new possible combinations of ancestors for 𝔓75. The decision comes down to {A} or {022, A} – and is effectively a question of how connective is the variant unit John 6:24/38. There is probably *some* genealogical meaning in this unit (πλοια versus πλοιαρια) and so it seems best to choose {022, A}.

Following precisely the same arguments as in the main section above, the optimal substemmata can be defined. The best combinations of ancestors are the same as before, with the following exceptions: 01 {028, A}; 03 {A, 03<𝔓75}; 07 {063, 0141}; 017 {021, 0141}; 019 {03, A} or {022, A}; 021 {063, 0141, A}; 022 {A}; 030 {07, 036}; 032 {03}; 044 {03, A}; 0141 {A, 𝔓75}; and 0211 {A, 030}. The new global stemma is shown next to the old one for comparison in Figure 23. Perhaps the only significant change is the promotion of 022 from a leaf node in the old global stemma to an important ancestor of 𝔓75 (and then through 0141 it is an ancestor of most of the extant witnesses). The other changes mostly fall into the category of an edge moving up or down one level in the hierarchy.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
019		-	97.143	34	35				1
03	1		94.444	34	36	1			1
A	2		90.625	29	32	3			
02		-	83.333	30	36	3	3		
037		-	80.000	28	35	3	3	1	
031	3		80.000	24	30	3	2	1	

*Table 72: Potential ancestors with W1=𝔓75 in the (**old** collation)*

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
03<𝔓75		-	100.000	1	1				
019		-	96.552	28	29				1
03		-	93.333	28	30	1	1		
022	1		93.103	27	29	1			1
A	2		88.462	23	26	3			
044		-	86.667	26	30	1	1	1	1
091		-	86.667	13	15	1	1		
032		-	86.207	25	29	1	1	1	1
02		-	83.333	25	30	2	2		1
031		-	80.000	20	25	2	2	1	
037		-	79.310	23	29	2	2	1	1

*Table 73: Potential ancestors with W1=𝔓75 in the (**new** collation)*

Vorf	Vorfanz	Stellen	Post	Fragl	Offen	Hinweis
022, A	2	24	2	4	0	<<
A	1	23	3	4	0	<<
022	1	23	1	4	2	<<

Table 74: Combinations of ancestors for 𝔓75

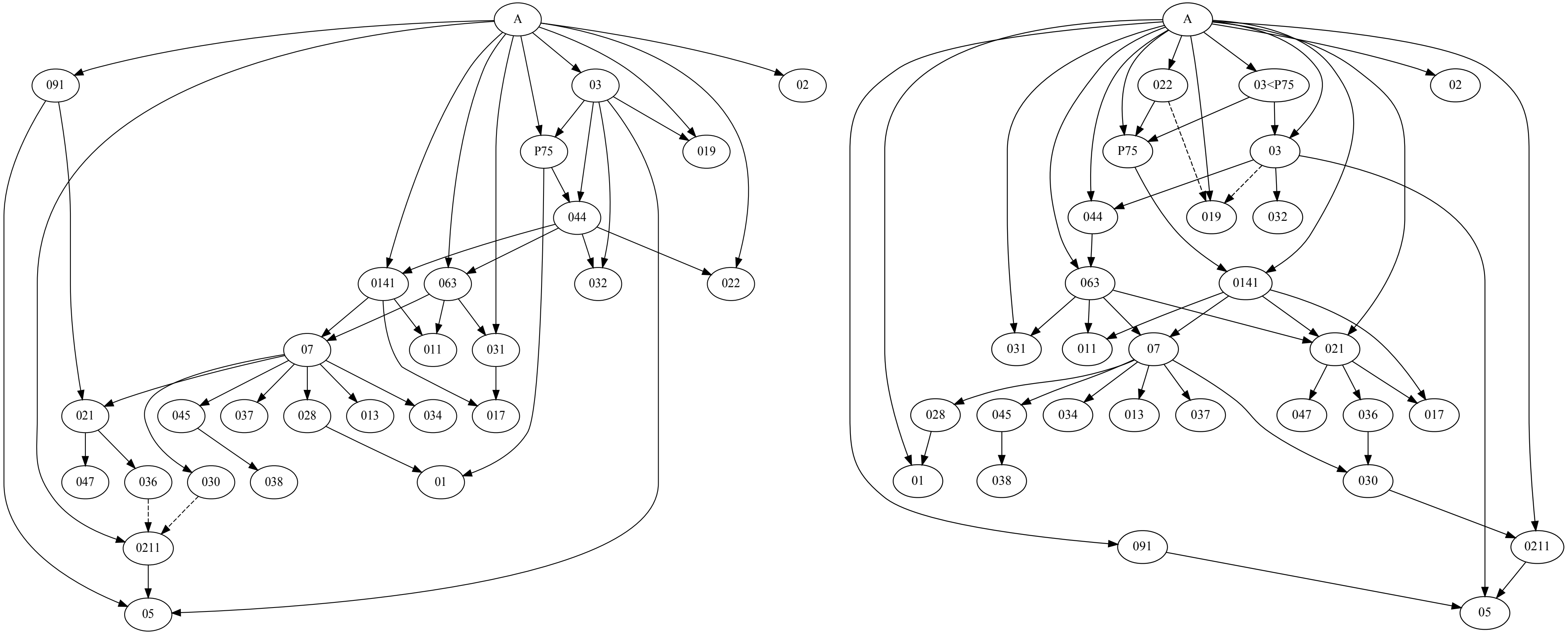


Figure 23: Old and new global stemmata

Initial Text

Finally we will take the initial text as derived from the original decisions and compare it with that derived from Wachtel's changes. The only change to the initial text is John 6:23/4 where *ηλθεν* now becomes *ηλθον*. While this is singularly undramatic, it is nevertheless important to note that changing the collation has led to a change in the initial text.

Correctors

The INTF's standard way of applying the CBGM is to take the text of the manuscript as it left the firsthand – including any corrections the firsthand made. The inclusion of firsthand corrections is logical, in the sense that this was the form of text as it left the scribe, even if it is not always as it left the scriptorium if a supervisor made corrections.[36] Yet some manuscripts have a rich history of correction, and in some cases these corrections would have significant genealogical value. Consider a heavily corrected manuscript being used as an exemplar for another. If these corrections were not to be included in the CBGM then valuable source readings for local stemmata might not be present. The contrary argument to this would be that another manuscript probably (hopefully?) carries any important readings in its firsthand.

As an example, take the fourth-century manuscript Codex Sinaiticus (01) which is famous as being the most heavily corrected New Testament manuscript.[37] In the original collation (not the one with Wachtel's changes) we have seen that 01 has a number of strange and unique readings. In fifteen variant units it has a unique reading, and in several others it shares a reading with only one or two other witnesses. As such it has low pre-genealogical

---

[36] This was true in 2015. As of 2017 Wachtel has adopted Parker's policy of including selected scriptorium corrections.

[37] See Parker, *Codex Sinaiticus*, 79; 'Codex Sinaiticus: Experience the Oldest Bible. The Transcription.'

coherence with all other witnesses.[38] Yet only three of the unique readings remain in the ultimate form of text attested to by 01, after its correctors have done their work. John 6:22/88 was modified by corrector "S1" in the scriptorium.[39] A further nine variant units were corrected to read our initial text by corrector "ca" who operated somewhere between the fifth and seventh centuries.[40] In two more units "ca" corrected the text towards but not exactly to our initial text.[41] In fact, "ca" worked so thoroughly that if we consider the text as it left his hand (thus incorporating corrections from S1 and ca) as a witness in its own right then a remarkable picture emerges.

Table 76 shows that 01ca is most closely related to A (86.667%), then 03 (85.714%) – and most distantly related to 01 (52.941%). Having completed the local stemmata above and thus defined the initial text where possible, we can also consider pre-genealogical coherence data for A (see Table 77) - which shows that 01ca is A's fifteenth closest relation (compared with 01 the twenty-eighth) and is only 7% behind its closest (03 with 93.75% coherence).

Genealogical coherence data also has some surprises. Table 78 shows that 01ca is now 01's highest ranked potential ancestor – a situation that is definitely wrong from a manuscript perspective, but intriguing from a text perspective. Was corrector "ca" making corrections based on a manuscript that was a real ancestor of 01? This "backwards" textual flow is commonplace in the CBGM. See Chapter 9 for a discussion of this phenomenon.

---

[38] Its closest relative is 031 with 48.485% coherence.

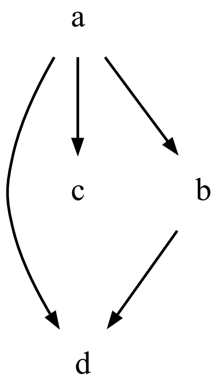
[39] S1's corrections were part of the manuscripts production process, but it is not always possible to tell who made the correction – the firsthand or another scriptorium corrector.

[40] John 6:21/36, 22/52 (actually he deleted the large addition 22/40-52), 22/60, 22/62-66, 24/2-10, 24/14-20, 24/28, 24/36 and 24/38.

[41] John 6:23/12-16 and 23/20-22

Correctors

John 6:23/12-16 presents a problem. We can say, for certain, that the local stemma should be as shown in Table 75 (from a manuscript copying perspective). But this is unable to create a coherent textual flow – because 01 is not a potential ancestor of 01ca. Indeed, we have just shown that 01ca is the best potential ancestor of 01. In fact, if we had a copy of the corrected text of 01 (but did not know its history) then we might think it was an ancestor of 01 rather than the other way around. It might seem that we are therefore left to choose between an impossible textual flow (as far as the CBGM is concerned) and a false local stemma (as far as manuscript history is concerned). However, this can be resolved by the introduction of another intermediary node “01ca<01” that is lacunose apart from John 6:23/12-16 where it agrees with 01.[42]



a	εγγυς του τοπου	...
b	εγγυς ουσης	01
c	om	032
d	εγγυς του τοπου ουσης	01ca

Table 75: Local stemma for John 6:23/12-16

The new optimal substemma for 01 needs only {01ca, 028} (explaining two more readings by posteriority compared to {01ca, 028, 038, 𝔓75}, namely the addition of an article and changing *ανεβησαν* to *ενεβησαν*). The new global stemma is shown in Figure 24. See Chapter 9 for a further discussion of the inclusion or exclusion of correctors in the CBGM.

---

[42] Mink, ‘Problems of a Highly Contaminated Tradition’, 60.

W2	NR	PERC1	EQ	PASS
A	1	86.667	26	30
03	2	85.714	30	35
032	3	85.294	29	34
022		85.294	29	34
019		85.294	29	34
044	6	82.857	29	35
𝔓75	7	82.353	28	34
037		82.353	28	34
063	9	81.481	22	27
07	10	80.000	28	35
021		80.000	28	35
017		80.000	28	35
0141		80.000	28	35
031		80.000	24	30
047	15	77.143	27	35
045		77.143	27	35
036		77.143	27	35
034		77.143	27	35
028		77.143	27	35
02		77.143	27	35
011		77.143	27	35
091	22	75.000	12	16
030	23	74.286	26	35
013		74.286	26	35
038	25	71.429	25	35
0211		71.429	25	35
05	27	70.000	21	30
01	28	52.941	18	34

Table 76: Pre-genealogical coherence with W1=01ca

W2	NR	PERC1	EQ	PASS
03	1	93.750	30	32
031	2	92.308	24	26
𝔓75	3	90.625	29	32
02		90.625	29	32
037	5	90.323	28	31
019		90.323	28	31
07	7	87.500	28	32
044		87.500	28	32
0141		87.500	28	32
011		87.500	28	32
063		87.500	21	24
091		87.500	14	16
032	13	87.097	27	31
022		87.097	27	31
01ca	15	86.667	26	30
045	16	84.375	27	32
034		84.375	27	32
028		84.375	27	32
021		84.375	27	32
017		84.375	27	32
047	21	81.250	26	32
036		81.250	26	32
030		81.250	26	32
0211		81.250	26	32
013		81.250	26	32
038	26	78.125	25	32
05	27	72.000	18	25
01	28	41.379	12	29

Table 77: Pre-genealogical coherence with W1=A

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
01ca	1		52.941	18	34	13		2	1
031	2		48.485	16	33	12		3	2
028	3		47.368	18	38	14	1	3	2
07	4		44.737	17	38	15	1	3	2
038			44.737	17	38	13	2	3	3
021			44.737	17	38	15	1	3	2
017			44.737	17	38	15	1	3	2
0141			44.737	17	38	15		4	2
013			44.737	17	38	14	1	3	3
037	10		43.243	16	37	15	2	3	1
𝔓75	11		42.424	14	33	14		3	2
019			42.424	14	33	13		4	2
045	13		42.105	16	38	15	2	3	2
036			42.105	16	38	15	2	3	2
030			42.105	16	38	15	2	3	2
011			42.105	16	38	15	1	4	2
A	17		41.379	12	29	16			1
047	18		39.474	15	38	15	3	3	2
034			39.474	15	38	15	3	3	2
0211			39.474	15	38	15	4	2	2
063	21		38.462	10	26	10	1	3	2
03	22		38.235	13	34	15		4	2
032	23		36.364	12	33	14		5	2
022			36.364	12	33	14	1	4	2
044	25		35.294	12	34	14		5	3
05	26		33.333	11	33	9	5	4	4
091			33.333	5	15	6	1	2	1
02	28		32.353	11	34	16	2	4	1

Table 78: Potential ancestors with W1=01

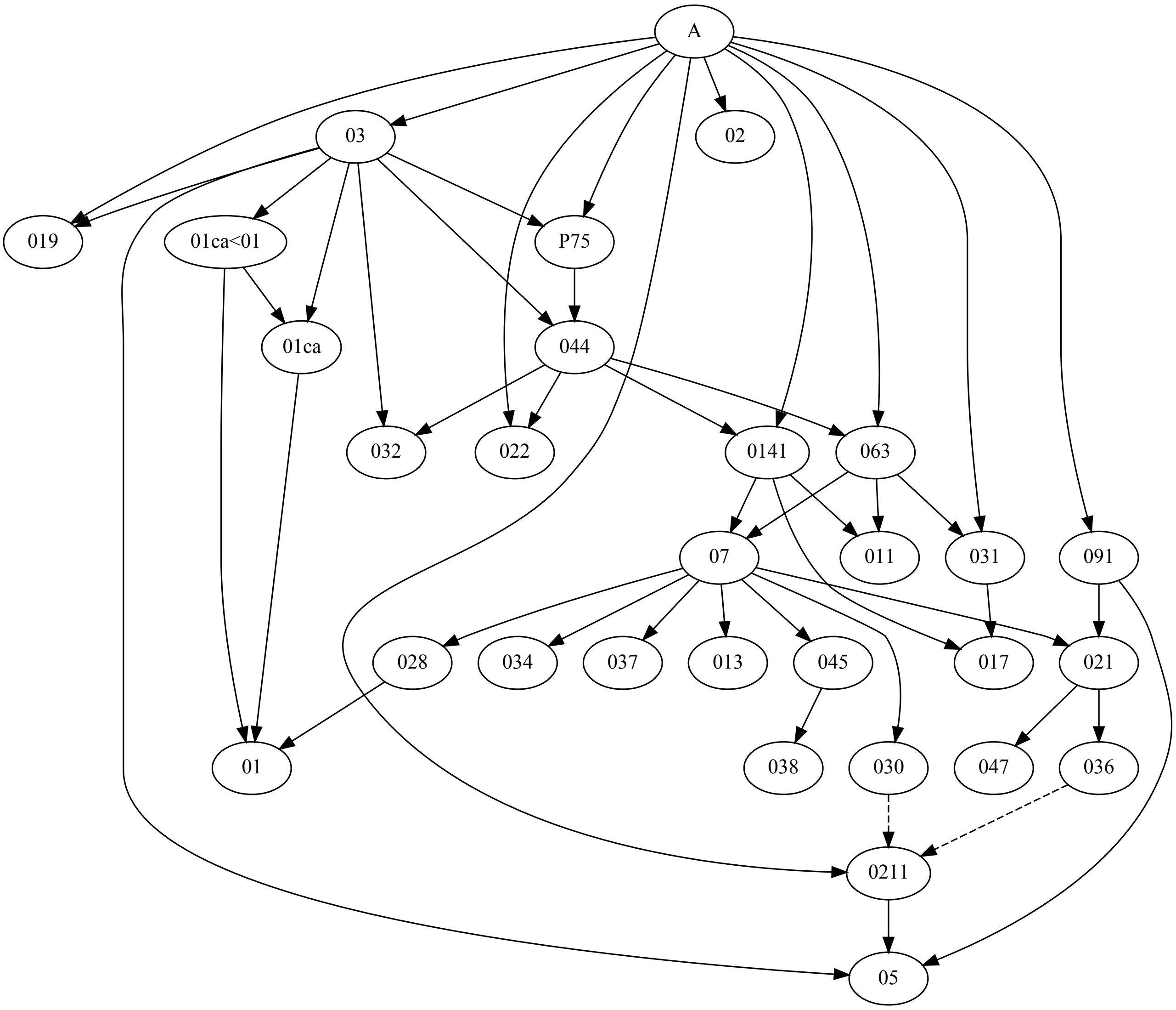


Figure 24: Global stemma including 01ca

Phylogenetic Comparison

Now that the global stemma has been created let us briefly compare it to the output from two different phylogenetic programs.

Median-Joining Network

This median-joining network was created for this data set at the point where we began to create local stemmata.[43] The median-joining network is shown in Figure 25. Note that any variant unit where a witness has a gap or lacuna is excluded – and hence several witnesses appear as identical when they are actually slightly different. Also 091 has been excluded as it is too lacunose and would dramatically reduce the data used to create the network.

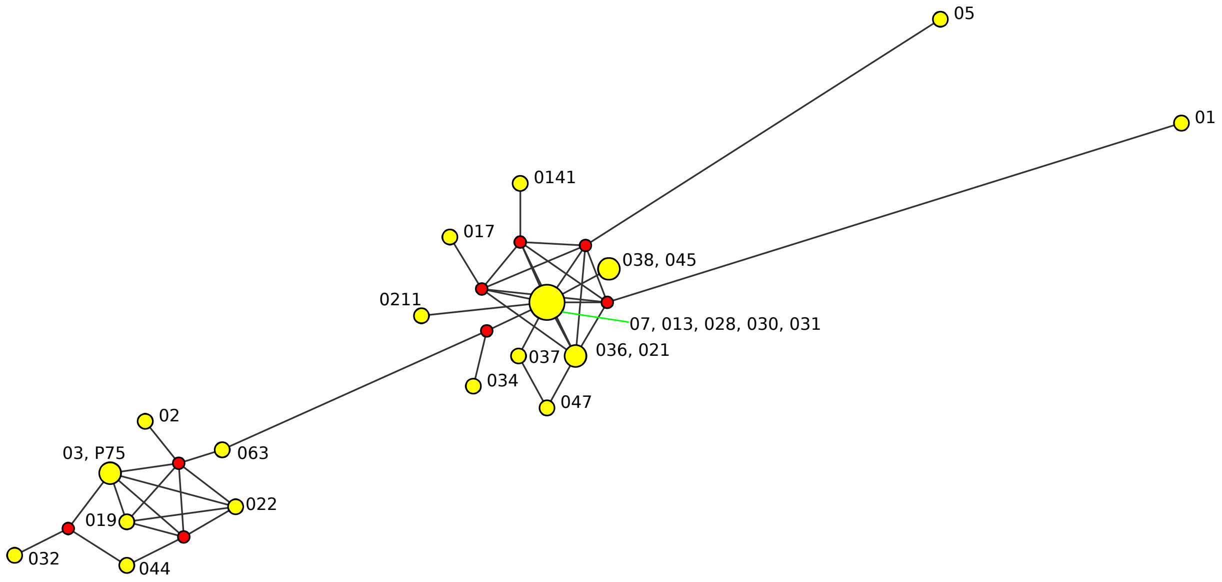


Figure 25: Median-joining network for John 6:21-24

It is interesting to note the two distinct clusters of witnesses. These groups are also present in the global stemma (Figure 19 on page 131) – where they are also linked by 063 and 07 (although in the network 07, 013, 030, 031 and 028 are all identical). The distance between

---

[43] For a detailed description of the creation of a median-joining network from this kind of data see Chapter 5.

the groups is not discernible in the global stemma. This is an example of the different strengths and weaknesses of the two approaches.

Phylogenetic Tree

Figure 26 shows the resulting tree from applying *MrBayes* (Metropolis-coupled Markov chain Monte Carlo analysis on BlueBEAR) to this dataset – again at the point where we began to create local stemmata.[44]

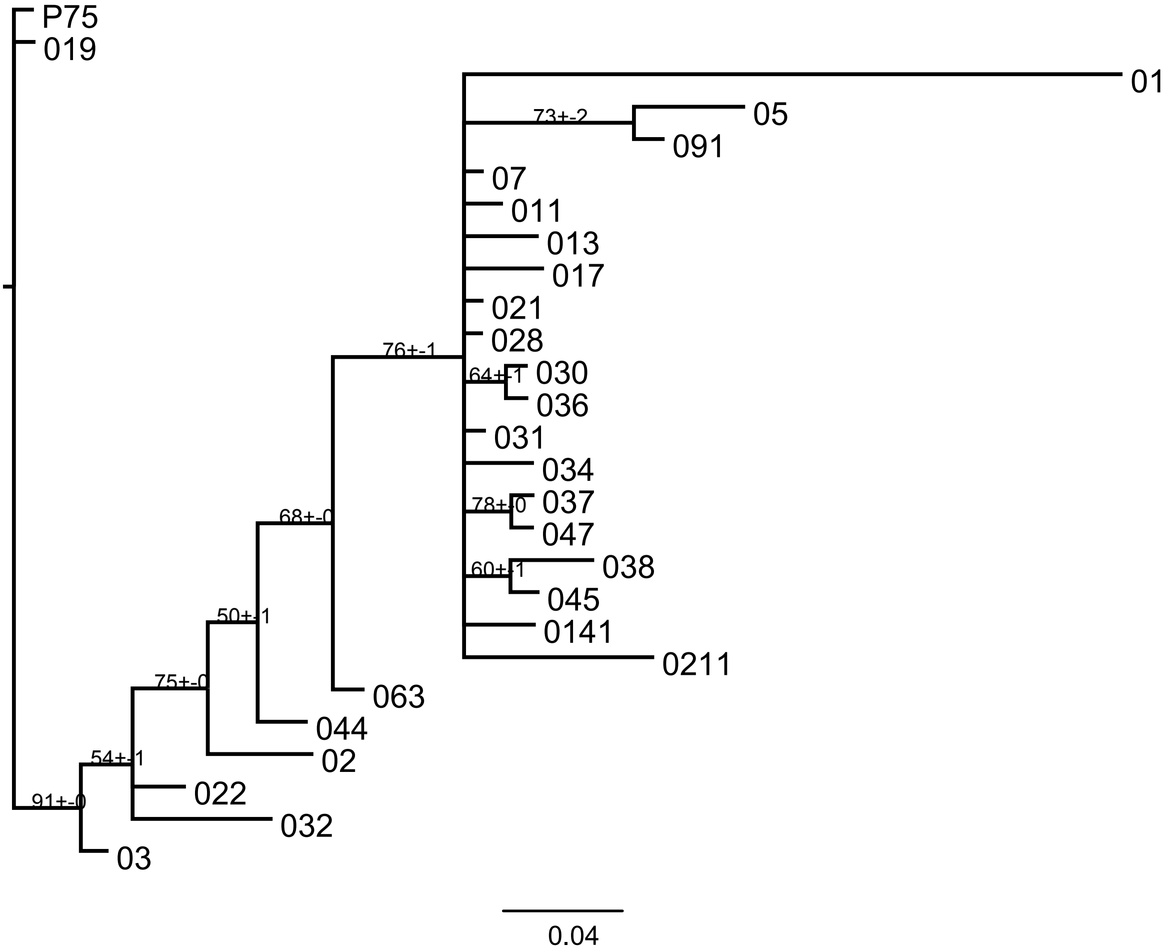


Figure 26: MrBayes tree for John 6:21-24

Again, many similar features are discernible from this tree and the global stemma. But different ones stand out when compared to the median-joining network. For example, here the

---

[44] For detailed information about *MrBayes* and this kind of tree see Chapter 5.

relationships between 05 and 091, or 038 and 045, can be seen in both this tree and the global stemma. However, this tree suggests a similar relationship between 037 and 047, which is not implied by the global stemma. As with the median-joining network, the same two distinct groups can be seen (although not as clearly as in the network).

In all three diagrams we find 01 at the end of a branch, if branch is the right word in the global stemma. The phylogenetic diagrams are able to highlight the enormous distance between 01 and the rest of the corpus, whereas this concept does not exist for the global stemma. We have already said above (see "Correctors") that 01 is famous for being the most corrected manuscript. But how should we explain the distance from the rest of the corpus? Were 01's scribes of low quality, or perhaps unscrupulous, missing bits out often or inventing new readings? Discussing an example from the Old Testament of 01 Parker states that "the prophetic books were not very well written in the first place, and therefore needed more [corrections] to get them up to standard."[45] Is this the case throughout 01, or is it possible that 01 represents a part of the tradition that is now mostly lost? Phylogenetic (and the CBGM's) diagrams frequently raise such questions.

As a final note, it is intriguing how well *MrBayes* has defined the top of the tree (although this could equally well be the bottom as far as *MrBayes* is concerned) and one could even draw "A" in at the root.

---

[45] Parker, *Codex Sinaiticus*, 89.

INTF's Nine Guidelines

At SBL in 2014 the editorial team of ECM Acts presented a paper based on their experience from Acts and expanding the original assumptions of the method. This paper included nine guidelines for creating local stemmata. Now that the above CBGM walkthrough is complete, it is valuable to reflect on the Acts team's guidelines. In the following section the guidelines are in italics, followed by my discussion.[46] Note that TP stands for transcriptional probability (traditional, internal, text-critical criteria) and GC for genealogical coherence.[47]

> *1. Singular readings and unique readings of small groups which differ from the mainstream of transmission are secondary. Exceptions to this rule require strong support from internal criteria. (Ex. 1, 4, 7) However, as was the case with the Catholic Letters, such variants are systematically subjected to text critical analysis if they are supported by witnesses closely related to A.*

At face value this guideline is worrying. It feels like a re-worded inversion of the maxim that variants must be weighed and not counted. If a reading has sufficiently little support (a very low count) then it can be discarded. But the caveat of the second sentence mitigates against this – and the Acts team are saying that, in their experience, readings with few attestations can normally be safely discarded. The third sentence brings in the idea of "good witnesses" of editions past, which in the ECM are the very same witnesses that appear as closely related to A. This example highlights the practical nature of the guidelines: they are designed to help the editorial team progress through the task of creating local stemmata

---

[46] See Gäbel et al., 'The CBGM Applied to Variants from Acts: Methodological Background', 3; For the examples mentioned here see Wachtel, 'Constructing Local Stemmata for the ECM of Acts: Examples'.

[47] It should be noted that TP is actually defined in the guidelines as "transcriptional priority", presumably by mistake. But in the same paper, and many other CBGM papers, it means "transcriptional probability". See, for example, Hüffmeier, 'The CBGM Applied to Variants from Acts' presented in the same session at the same conference.

quickly. Such optimisations are necessary, but require a careful hand to avoid simply producing an edition based on 03 and its related witnesses.

> *2. An attestation lacking coherence is a sign of multiple emergence i.e. posteriority of a variant. (Ex. 1, 2, 3, 5) Multiple emergence weakens the force of internal criteria which might be used to account for the priority of the variant.*

Multiple emergence is a difficult issue. Yes, the guideline is correct that multiple emergence is a strong sign of posteriority. However, as can be seen earlier in this chapter the easiest way to avoid multiple emergence of a reading is to declare that it is the initial text. Then such readings nearly always have perfect coherence. Of course, standard text-critical criteria will prevent this practice, but the corollary to this point is more important: multiple emergence of the initial text (i.e. that the *a* reading emerged on multiple occasions) is nearly always hidden because, in most cases, A is a potential ancestor of all witnesses in the attestation.

> *3. Good coherence of an attestation is primarily a sign of unfractured transmission. Good coherence is a valid argument for the priority of a variant only if supported by internal criteria. (Ex. 6, 8, 11)*

This is somewhat the converse argument to the previous point. There poor coherence (where visible) was a sign of posteriority. Here good coherence can strengthen an existing text-critical argument for priority – but in itself cannot make that claim. This is a useful guideline, as otherwise the scholar could fall into the trap of using the CBGM's external criterion of coherence in favour of text-critical internal criteria. It should also be noted that coherence (or otherwise) in a connective variant (broadly equivalent to Lachmann's "indicative error") should be weighted differently from coherence in a reading that could easily arise multiple times.

> *4. A strong argument for assessing a variant as initial text is provided by an attestation which combines coherence and a broad range of witnesses closely related to A. (Ex. 9, 10, 14, 15) In such cases strong coherence only materializes if A is part of the attestation.*

This guideline returns to "good witnesses" but now requiring a broad range of them. It is hard to argue against a reading that has support from, for example, 𝔓60, 𝔓66, 𝔓75, 01, 03, 04, 019, 032, and so on. The final point is true because these witnesses have few potential ancestors, where later witnesses typically have many.

> *5. The priority of a majority reading is indicated if it is linguistically more difficult or contextually less suitable and thus atypical of the majority text. This may be valid even if the competing variant has a broad range of witnesses closely related to A. (Ex. 10)*

This guideline highlights that the editorial team of Acts unsurprisingly have now mixed their own text-critical criteria with the definition of the CBGM. The CBGM can, in theory, be used by different approaches. This guideline may need to be discarded by scholars preferring different criteria. But the ECM editors follow the NA tradition and thus this guideline is a straightforward corollary to *lectio difficilior potior* – but a necessary counter to the previous guideline.

> *6. The source of a variant is likely to be a similar variant. If the attestation of a variant indicates that two or more other variants need to be considered as possible sources then TP suggests that the one which requires the least change to be transformed into the variant in question is preferable.*

This guideline is common sense, and a straightforward application of Occam's Razor. Nevertheless, it highlights a potential problem with the CBGM – namely that it effectively encourages the scholar to consider the text outside of its manuscript context. Now this is a guideline, and like the CBGM's assumptions we should consider it as being generally applicable but should allow exceptions. And the word "likely" in the first sentence reinforces that standpoint. The scholar should be on the lookout for exceptions. If circumstances and time permit (say for a variant attested by only a few witnesses) then the attesting manuscripts should be examined to see if any other factors need to be considered for the origin of the variant – for example the manuscript could be a catena including the source text in the commentary.

7. The source of a variant is questionable:

> *a) if GC and TP point to different potential source variants or cannot be aligned with each other for other reasons; (Ex. 8, 13)*
>
> *b) if we cannot decide which of two or more variants is the prior one because neither GC nor TP provides a convincing argument. (Ex. 18)*

This guideline can be seen as an indication to slow down. An editorial team must proceed quickly through the local stemmata, as there are thousands to be defined in the course of applying the CBGM to a typical New Testament book. Here the CBGM and traditional textual criticism disagree, and the local stemmata must be analysed in greater detail. This situation, we must assume, will occasionally or even regularly lead to an undefined initial text. It is important to note that the instinct of many of the CBGM's critics might be to ignore GC here and accept TP's solution – but that will not do. The CBGM provides checks and balances that have a genuine meaning and weight, and cannot easily be discarded.

> *8. Consciously introduced editorial variants are exceptional. If possible, variation should be explained with reference to the process of copying itself and to known causes of error. (Ex. 16, 17)*

What is a "consciously introduced editorial variant"? It seems likely that this refers to an editorial conjecture. In that case this guideline is another example of Acts' editors text-critical approach mixing with the CBGM itself. That conjectures are to be avoided is obvious to some, and clearly wrong to others. In the ECM they will, evidently, be rare.

> *9. If the witnesses of a variant have to be assigned to different source variants, then the attestation should split accordingly.*

Splitting an attestation (i.e. *b* becomes *b1* and *b2*) is common in the INTF's application of the CBGM – perhaps emphasised by the omission of a list of such examples following this guideline. Seven of the eighteen examples in Wachtel's *Constructing Local Stemmata* include split attestations. Being willing to split a reading – because the attestation clearly forms two or more groups – is important. But doing so feels at the time like adding unnecessary complexity. Perhaps due to this unconscious desire to keep things simple, there is only a single split attestation in my CBGM walkthrough above, namely John 6:22/20. Similarly, to date, Parker (ITSEE) has only a single split attestation in his full set of local stemmata for John's Gospel. Does this indicate that splitting attestations is actually unnecessary? We cannot say for sure, but the editors of Acts would argue that it should be done.

In summary, these guidelines are a useful aid for those who wish to apply the CBGM. They are particularly valuable if the scholar wishes to emulate the Acts team's text-critical methodology of reasoned eclecticism and a preference for older witnesses.

Conclusions

Perhaps the first conclusion to draw is that carrying out the full CBGM process on even a small dataset is very time consuming. It is also something like a pyramid, with the collation at the bottom, then upwards through the initial local stemmata, their updating from genealogical coherence, optimal substemmata and finally to the global stemma at the top. Thus any layer is only as valid as the decisions made in the layer below – and, as has happened several times while writing this chapter, any mistake identified in a lower layer can lead to many hours of re-working those decisions that depended on it. Similarly, a change made at a low layer will affect the higher ones – for example altering the collation can yield a different initial text.

The positive side of the first conclusion is that the CBGM is powerful, and does what it sets out to do – namely genuinely capture the textual decisions made and allow the user to see their consequences. The emergent initial text and the insights into the development of the text through the majuscule period are valuable results.

A final (tentative) conclusion is that applying independent phylogenetic techniques to this data has produced results that share many features with the CBGM's global stemma. It seems reasonable to infer that the CBGM is therefore producing valid results and that it could be considered just as trustworthy as are these phylogenetic techniques. This conclusion will be tested in much more detail in Chapter 10.

PART TWO

PHYLOGENETICS

Chapter 4: Stemmatics and Phylogenetics

Introduction

This chapter begins with an introduction to stemmatics and a discussion about why phylogenetics can be applied to manuscript studies. This is followed by a short review of projects that have successfully applied phylogenetics to manuscript traditions. After this, I describe my exploration of the IGNTP's transcriptions of John's Gospel and initial experiments using phylogenetic software. The following quotation from the special "Thematic Section on Studia Stemmatologica" from the *Digital Scholarship in the Humanities* (DSH) journal (2016, 31.3) neatly summarises the convergence of disciplines and techniques that underpin my thesis:

> *While having originated in textual scholarship, modern stemmatology lies at the intersection of several scientific disciplines. On one hand, it is associated with humanities where texts are used as sources and both their contents and the relationships between the individual textual witnesses are valuable as historical evidence. On the other hand, evolutionary ideas are equally present and methodologically interesting in disciplines like mathematics, statistics, computer science, and obviously, evolutionary biology and cladistics, the study of biological evolution and speciation. Academics working throughout the arts and humanities are borrowing the latest techniques employed in the sciences as tools that can help us understand the richness of human knowledge and creativity all the way from its roots, throughout the history, until our present society.*[1]

---

[1] Heikkilä and Roos, 'Thematic Section on Studia Stemmatologica', 520.

Stemmatics

In 1831 Karl Lachmann published the first critical edition of the Greek New Testament.[2] He defined a methodology that still persists today at the heart of textual criticism. A fundamental principle he set out is that of the *Leitfehler*, or indicative error, which Parker describes clearly as “a distinctive reading representing an alteration of the text which is found in two or more manuscripts and cannot have been made on two separate occasions.”[3] Such an indicative error might be a large omission or addition, or a distinctive alteration to the text, for example. Lachmann used these indicative errors to identify how manuscripts are related to one another, in particular how they are descended from one another through the scribal copying process. The principle is that manuscripts which share sufficient common errors will have a shared ancestry – and this ancestry can be reconstructed using hypothetical witnesses even though most of the ancestors themselves have been lost. This information is represented in family tree form, and is known as a stemma, the construction of which is called *recensio*/recension. Lachmann’s “ultimate goal was the establishment of texts that would constitute the nearest possible approximation to the authors' originals.”[4] In that, he shares a goal with many textual scholars who followed him. This method is widely known as “Lachmannian Stemmatics” and is the basis of much twentieth and twenty-first century textual scholarship.[5]

Lachmann also introduced the method of documenting the differences between a manuscript and a base text, which is still used today by transcribers to create data that can be

---

[2] See Lachmann, *Novum Testamentum Graece*. Lachmann published critical editions of many other works throughout the nineteenth century, but his *Novum Testamentum Graece* is of particular interest to this thesis for obvious reasons.

[3] Parker, *An Introduction to New Testament Manuscripts and Their Texts*, 162.

[4] Caruso, ‘Review: The Genesis of Lachmann’s Method by Sebastiano Timpanaro; Glenn W. Most’, 469.

[5] It is notable that Timpanaro has shown that the use of such a stemma actually predates Lachmann, despite him being popularly associated with its invention. See Timpanaro, *The Genesis of Lachmann’s Method*, 65.

collated. The related process of changing the base text into what the scholar thinks is an appropriate critical text is called *emendatio*/emendation. However, it is *recensio* or stemmatics that concerns us here. The following brief review is intended to bridge the gap between the nineteenth and twenty-first centuries, and will be necessarily abridged, focusing on authors and points of particular relevance to this thesis.

Lachmann's theories, applied by him to a wide range of texts that included classical authors, the New Testament, and medieval German writings, became the most common methodology for textual editing. The way in which it came to be applied is probably best shown in the work of Paul Maas, who in 1927 wrote what became the standard handbook on textual criticism, and which was translated into English in 1958. Importantly, he highlights that it is impractical to apply stemmatics thoroughly to works where a large number of witnesses exist as "this would involve an enormous amount of labour [and] be far too expensive to publish".[6] The result is that "we shall often have to make shift for the time being – establishing the stemmatic position only roughly, eliminating witnesses after testing samples."[7] He highlights the problem of contamination, and states that "where contamination exists the science of stemmatics in the strict sense breaks down."[8] These points are crucial to understanding the value that modern, computer-aided methods can bring to this field.

In 1963 Timpanaro first published his seminal book *La genesi del metodo del Lachmann.* Most's English translation was published in 2005 as *The Genesis of Lachmann's Method*.[9] Timpanaro points out that half of Lachmann's method, namely *emendatio*, "had

---

[6] Maas, *Textual Criticism*, 27.

[7] Ibid., 27f.

[8] Ibid., 48.

[9] It is pleasing for a textual critic to see that Glen Most's 2005 translation of Timpanaro includes a "general critical apparatus" listing the development of Timpanaro's work through its various Italian and German editions. Most's primary source was Timpanaro's personal copy of his final 1985 edition. See Timpanaro, *The Genesis of Lachmann's Method*, 26.

been practiced since antiquity."[10] For example, Timpanaro shows Erasmus doing just this in the sixteenth century.[11] Regarding the other half, *recensio*, he states that "the great novelty of nineteenth-century textual criticism was the scientific foundation of *recensio*."[12] He describes the widespread reluctance in the eighteenth and nineteenth centuries to allow a text other than the *textus receptus* of the New Testament to be printed, while printing variants at the bottom of the page was admitted. The fault had arisen because the popularity of the *textus receptus* had made it become the *de facto* tradition, and deviation from the tradition (even towards older manuscripts) could not be allowed. This then led, in turn, to sceptics attacking the authenticity of the biblical text *per se* due to the number of variants amassed in these other editions.

Timpanaro describes how Bentley, from a desire to "defend the authority of the biblical text... sought to establish the text more solidly and thereby to defeat skepticism."[13] Bentley's plan sounds very familiar to a modern textual critic: he would create an edition based on the "most ancient and venerable manuscripts" in Greek and Latin and therefore identify a text from the time of the council of Nicea.[14] He would use versions and patristic citations to confirm his decisions. He also makes the mistaken statement that, concerning thirty thousand variant readings, "there will scarce be two hundred out of so many thousands that can deserve the least consideration."[15] Bentley sadly did not complete this project, in part due to opposition (from those defending the *textus receptus*) and in part the size and complexity of the task. Timpanaro points out that Lachmann, in his critical editions (1831-1850), was finally able to achieve Bentley's goal, of creating "an edition founded solely on

---

[10] Ibid., 43.
[11] See ibid., 49.
[12] Ibid., 43.
[13] Ibid., 63; See, for example, Bentley, *Remarks upon a Late Discourse of Free-Thinking*.
[14] Bentley, *Dr. Bentley's Proposals for Printing a New Edition*, 3.
[15] Ibid., 4.

the ancient manuscripts and [the Vulgate]” after which the *textus receptus* never reasserted itself.[16]

In 1973 West effectively replaced Maas’ book as the standard English handbook by publishing his *Textual Criticism and Editorial Technique.* He criticises Maas for treating contamination as "a regrettable deviation about which nothing can be done, instead of as a normal state of affairs."[17] He introduces the problems of contamination and multiple parenthood of manuscripts, gives examples of how scribal errors can arise, and discusses some examples of classical authors who are known to have produced a new edition of their own work. He then walks the reader through the process of creating a critical edition, from organising a corpus of manuscripts through to publishing a text and critical apparatus. He highlights the importance of regularising the data: “... we shall surely come nearer the truth by regularizing the spelling than by committing ourselves to the vagaries of the tradition.”[18] West’s point about regularisation turns out to be particularly important when preparing data for use with computer-aided methods such as the CBGM or phylogenetic software, since they are not otherwise able to tell, for example, that Ἰεροσόλυμα and Ἱερουσαλήμ should be treated as the same.

Lachmannian Stemmatics can be applied successfully to certain parts of the New Testament tradition, for example Anderson on Family 1 in Matthew and Welsby on Family 1 in John.[19] The reason for this success is that the effects of contamination are greatly reduced within the manuscript members of Family 1. Contamination makes applying traditional Lachmannian Stemmatics impossible for the New Testament manuscript tradition as a whole.

---

[16] Timpanaro, *The Genesis of Lachmann’s Method*, 84.

[17] West, *Textual Criticism and Editorial Technique*, 5. Timpanaro also criticises Maas, to the extent to which Courtney describes Maas as Timpanaro’s “bugbear”. Courtney, ‘Review: The Genesis of Lachmann’s Method by Sebastiano Timpanaro; Glenn W. Most’, 313.

[18] West, *Textual Criticism and Editorial Technique*, 69.

[19] See Anderson, *The Textual Tradition of the Gospels*; Welsby, ‘A Textual Study of Family 1 in The Gospel of John.’

Stemmatics

As has been seen in Chapter 1, the CBGM claims to be the answer to contamination – for example Parker's assertion in his 2008 introduction to textual criticism states that "We are at last able to make Lachmannian stemmatics workable in complex textual traditions."[20]

Phylogenetics for manuscript studies

The similarity of Phylogenetics and Manuscript Studies

While Lachmann was working on manuscript stemmatics, Charles Darwin was exploring the idea of how species relate to one another. In 1859 he published his famous work *On the Origin of Species*, in which he introduces the idea of the tree of life.[21] The fields of textual criticism and evolutionary biology have developed alongside each other, and have frequently interacted. For example consider Streeter's 1924 book *The Four Gospels: A Study of Origins* in which he sets out to provide an introduction "to Textual Criticism, to the Synoptic Problem and the Johannine question."[22] Parker asks "Is Streeter drawing on Darwin's observations...?"[23] Lin describes him taking a "distinctly Darwinian approach to the study of texts and their geographical origin."[24]

The evolution of manuscripts (or texts) and the evolution of species have many similarities. Chapter 2 cites some key New Testament scholars' views on the matter which will be briefly revisited here. Lin's book *The Erotic Life of Manuscripts* is subtitled "New Testament Textual Criticism and the Biological Sciences", and contains a detailed history of the origins of the fields and their similarities.[25] She states that "From the organization of diversity by classification to the hierarchization of groups and finally to the association of

---

[20] Parker, *Textual Scholarship and the Making of the New Testament*, 84.

[21] See Darwin, *On the Origin of Species*.

[22] Streeter, *The Four Gospels: A Study of Origins*, xxii.

[23] Parker, *An Introduction to New Testament Manuscripts and Their Texts*, 169.

[24] Lin, *The Erotic Life of Manuscripts*, 77.

[25] See Lin, *The Erotic Life of Manuscripts*.

variation with impurity and contamination, textual criticism kept in step with the developments of natural science, adopting its language and borrowing its structures."[26] Lin points out that textual criticism followed the geneticists' advances in the twentieth century, seeing that textual difference should "in a Darwinian perspective, be understood as interesting in and of itself as a stage in a living text's adaptation to its environment."[27] She states that "By the latter half of the twentieth century, New Testament textual criticism would, for the most part, leave the methods of classification and rudimentary genealogy behind."[28] This marks the transition from the use of a handful of geographical text-types to relate manuscripts to one another, to methods that use computers to analyse large sets of complex data – such as the CBGM and phylogenetics.

Parker compares problems faced by those seeking to apply Lachmannian Stemmatics with those faced by geneticists. He points out that contamination is like genetic cross-mutation; coincidental emergence of readings is like independent mutations; and corrections to an earlier text are like a later mutation undoing another mutation's change. These similarities have meant that phylogenetic methods and software developed by geneticists can be adapted and used by textual critics.[29]

The two fields are often described as "analogous", for example in Parker's comparison of Streeter and Darwin.[30] Lin describes that genetics provides "a completely new metaphor through which to view texts and textual variants". She also quotes Dawkins as saying that "So similar are the techniques and difficulties in DNA evolution and literary text evolution, that each can be used to illustrate the other."[31] Lin states that "Because there are so

---

[26] Ibid., 62.

[27] Ibid., 64.

[28] Ibid., 65.

[29] See Parker, *An Introduction to New Testament Manuscripts and Their Texts*, 167ff.

[30] Ibid., 169; Streeter, *The Four Gospels: A Study of Origins*; Darwin, *Journal of Researches*.

[31] Lin, *The Erotic Life of Manuscripts*, 110. Lin is quoting Richard Dawkins, *The Ancestor's Tale: A Pilgrimage to the Dawn of Evolution*, New York: Houghton Mifflin, 2004, 128.

many apparent parallels between an organism's genes and a manuscript's text, the tools of genomics and cladistics are readily adaptable to text-critical work."[32]

The following quotation from Semple and Steel's introduction to Phylogenetics highlights just how appropriate phylogenetic tools are for studying manuscript traditions – and also shows the surprised reaction that this often elicits in biological fields:

> *"[The mathematical foundations of phylogenetics] date at least to the pioneering work by Peter Buneman, David Sankoff, and others in the early 1970s. Curiously, Buneman's early paper (1971) dealt not with biology but rather reconstructing the copying history of manuscripts."*[33]

In 2004, Howe et al. published a paper 'Parallels Between Stemmatology and Phylogenetics' with the specific aim of exploring "the similarities between the evolution of DNA sequences and the changes occurring in manuscript traditions."[34] This paper is an excellent introduction to why phylogenetic tools should be appropriate for manuscript traditions. They state that "It is remarkable how many parallels there are between the evolution of genetic material and the changes occurring in manuscripts… These include recombination, convergent evolution and transposition."[35] They show that recombination parallels a scribe changing his exemplar mid-way through the copying process. A more complicated form of recombination parallels contamination, or multiple exemplars. They argue that transposition (for example a virus inserting DNA from one species into another) has a parallel in manuscript studies where text from one passage is inserted into another. Convergent evolution is the name given to independent similar mutations, and is "comparable to convergence or parallelism in manuscript traditions", for example independent scribes

---

[32] Ibid., 118.

[33] Semple and Steel, *Phylogenetics*, ix.

[34] Howe et al., 'Parallels Between Stemmatology and Phylogenetics', 3.

[35] Ibid., 7.

following local contemporary customs regarding spelling or dialect.[36] They conclude that "The process of incorporation of changes into DNA mirrors the incorporation of changes into manuscripts… [and therefore] programs for phylogenetic analysis of sequence data can be exploited for stemmatic analysis of manuscript traditions…"[37] It appears that the two fields are more than "analogous" - they map onto each other very closely indeed.

The application of Phylogenetics to Manuscript Studies

In the 1970s Buneman, Platnick, Cameron and others began to explore the appropriateness of using phylogenetic techniques to address text-critical problems.[38] The 1980s saw the first applications of phylogenetics to manuscript traditions. Robinson reports that "the first scholar to apply phylogenetic methodology in a stemmatic context" was Lee whose paper 'Numerical Taxonomy Revisited: John Griffith, Cladistic Analysis and St. Augustine's Quaestiones in Heptateuchum' was published in 1989.[39]

Two clusters of researchers came to prominence in the 1990s in this area. A group of scholars in the Netherlands carried out what was described by Robinson as "the first concerted and methodical scholarly investigation into quantitative methods (including phylogenetics) for the exploration of textual scholarship… in the early 1990s".[40] Key among these scholars are van Reenen and van Mulken from Vrije Universiteit in Amsterdam, who edited two seminal works: *Studies in Stemmatology* in 1996 and, with den Hollander, *Studies in Stemmatology II* in 2004.[41] *Studies in Stemmatology* contains a selection of papers from the

---

[36] Ibid., 9.

[37] Ibid., 10.

[38] See Buneman, 'The Recovery of Trees from Measures of Dissimilarity'; Platnick and Cameron, 'Cladistic Methods in Textual, Linguistic, and Phylogenetic Analysis'.

[39] Robinson, 'Four Rules for the Application of Phylogenetics in the Analysis of Textual Traditions', 637; Lee, 'Numerical Taxonomy Revisited: John Griffith, Cladistic Analysis and St. Augustine's Quaestiones in Heptateuchum'.

[40] Robinson, 'Four Rules for the Application of Phylogenetics in the Analysis of Textual Traditions', 637.

[41] van Reenen and van Mulken, *Studies in Stemmatology*; van Reenen, den Hollander, and van Mulken, *Studies in Stemmatology II*.

annual Free University Stemmatological Colloquia 1990-93, convened to allow scholars from various disciplines to report their research on the genealogy of texts and "the application of the newest technology to manuscript traditions."[42]

The second group clusters around Robinson himself, who was a contributor to *Studies in Stemmatology*.[43] In 1991 Robinson posted a challenge online to "re-create by statistical or numerical means alone the table of relationships for some 44 manuscripts of the Old Norse narrative 'Svipdagsmal' established by Robinson on the basis of external evidence and traditional stemmatic methods."[44] A particularly successful response to the challenge came from O'Hara who used PAUP to carry out a cladistic analysis of the data.[45] Robinson and O'Hara went on to collaborate on a number of further projects, discussing and applying phylogenetic techniques to Chaucer's Canterbury Tales and Svipdagsmal.[46] As well as publishing his own individual papers, Robinson also collaborated with notable success with Barbrook, Blake and Howe on 'The Phylogeny of The Canterbury Tales', published in Nature in 1998. Here they used PAUP to create a phylogenetic tree from 58 manuscripts and identifed several distinct groups.[47] They reach a particularly interesting conclusion, that the root of the tradition is Chaucer's own copy which itself contained notes, deletions and additions – i.e. it was not a clean, finished text.

The 2000s saw *Studies in Stemmatology II*, as mentioned above. As discussed in Chapter 1, Mink's seminal work 'Problems of a highly contaminated tradition', his major publication on the CBGM, is found in *Studies in Stemmatology II.* Howe et al.'s 'Parallels

---

[42] van Reenen and van Mulken, *Studies in Stemmatology*, viii.

[43] See Robinson, 'Computer-Assisted Stemmatic Analysis and "Best-Text" Historical Editing'.

[44] M.W. Robinson and O'Hara, 'Report on the Textual Criticism Challenge 1991', 331.

[45] See Swofford, *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10*. This particular analysis was carried out using Macintosh Version 3.0r.

[46] O'Hara and Robinson, 'Computer-Assisted Methods of Stemmatic Analysis'; Robinson and O'Hara, 'Cladistic Analysis of an Old Norse Manuscript Tradition'.

[47] Barbrook et al., 'The Phylogeny of The Canterbury Tales'.

Between Stemmatology and Phylogenetics' (mentioned above) also comes from *Studies in Stemmatology II*.[48] Howe's co-authors of 'Parallels' were Barbrook, Mooney and Robinson. There is a strong collaborative trend in the articles on phylogenetics of manuscripts in the 2000s (continuing from the Nature article in 1998), and nearly all of them feature Spencer and Howe, or Robinson, or all three – along with other authors.[49] A notable two-fold collaboration between Spencer, Howe and Wachtel analysed manuscripts of the Letter of James from the New Testament. In their second paper, 'Representing Multiple Pathways of Textual Flow in the Greek Manuscripts of the Letter of James Using Reduced Median Networks', they argue that their results are consistent with the accepted theory that "most of the variation in the Greek New Testament arose early in the tradition."[50]

An attempt was made in 2011 to encourage the adoption of the term "phylomemetics" for the use of "phylogenetic analysis of non-genetic data" (from the word "meme") but it has not succeeded.[51] A clear introduction to the method of applying phylogenetics to manuscript traditions may be found in "Responding to Criticisms of Phylogenetic Methods in Stemmatology" by Howe, Connolly, and Windram.

In 2012 Perrin completed his PhD at Birmingham successfully using *SplitsTree* to analyse the relationships between the small number of Family 13 manuscripts in John.[52] In the same year De Vos, Macé, and Geuten published their paper 'Comparing Stemmatological and Phylogenetic Methods to Understand the Transmission History of the "Florilegium Coislinianum"' again with good results but with a small number of manuscripts. They used

---

[48] Howe et al., 'Parallels Between Stemmatology and Phylogenetics'.

[49] See, for example, Macé et al., 'Testing Methods on an Artificially Created Textual Tradition'; Howe et al., 'Parallels Between Stemmatology and Phylogenetics'; Spencer et al., 'Analyzing the Order of Items in Manuscripts of "The Canterbury Tales"'; Spencer et al., 'How Reliable Is a Stemma?'; Howe et al., 'Manuscript Evolution'; Spencer et al., 'Phylogenetics of Artificial Manuscripts'.

[50] Spencer, Wachtel, and Howe, 'Representing Multiple Pathways of Textual Flow in the Greek Manuscripts of the Letter of James Using Reduced Median Networks', 10; See also Spencer, Wachtel, and Howe, 'The Greek Vorlage of the Syra Harclensis'.

[51] Howe and Windram, 'Phylomemetics—Evolutionary Analysis beyond the Gene', para. 3.

[52] See, for example, Perrin Jr., 'Family 13 in Saint John's Gospel', 165 - a network of ten manuscripts.

PAUP, as did many of the papers above.

I will now discuss a special "Thematic Section on Studia Stemmatologica" from the *Digital Scholarship in the Humanities* (DSH) journal (2016, 31.3). This "Thematic Section" originated in a set of workshops organised by the section's editors in 2012 in Helsinki, Uppsala, Pisa, Cambridge and Rome.[53] Taken together these papers form an excellent overview of the task of applying phylogenetics to texts. Bordalejo sets out a clear introduction to stemmatology and the use of phylogenetics for analysing textual traditions.[54] For a full grounding the reader is directed to the articles themselves. Some key themes from these papers will be briefly discussed below, followed by a more in-depth look at Robinson's paper.

The question of the genealogical significance of a variant is tackled to varying extents in several papers in this special section. This is not only of interest for applying phylogenetics to textual traditions, but is also important for some arguments concerning the CBGM (see Chapter 2). Andrews concludes that philologists are not good at determining which variants are genealogically significant, and that "'insignificant' variation is really not that insignificant at all."[55] Her conclusions contrast with the related view from some scholars considering the CBGM that only "significant" variants should be included.[56] In my opinion, Andrews is right – for the significance of a variant in statistical analysis is not the same as its level of merit or interest from a text-critical perspective. This seems, to me, to be self-evident – for the definition of "genealogically meaningful" is not at all equivalent to saying that, for example, a variant reading is worthy of inclusion in a critical edition, or that it was made intentionally. It is easy to conceive of a hypothetical example where a simple spelling change in a place name

---

[53] See Heikkilä and Roos, 'Thematic Section on Studia Stemmatologica'; See also Macé, *The Evolution of Texts* for a similar work from a conference in 2004 featuring several of the same authors.

[54] See Bordalejo, 'The Genealogy of Texts'.

[55] Andrews, 'Analysis of Variation Significance in Artificial Traditions Using Stemmaweb', 538.

[56] See Chapter 2's discussion of, for example, Alexanderson, 'Problems in the New Testament: Old Manuscripts and Papyri, the New Genealogical Method (CBGM) and the Editio Critica Maior (ECM)'.

(a type of orthographic variation commonly regularised away) could be genealogically meaningful if it happened to be faithfully copied. Halonen discusses the question of the quality of a variant – is it "significant" or not? The computer-assisted methods generally work on the quantity of variants, and not their quality. So perhaps the question of significance is unnecessary anyway.[57] Robinson also addresses the question of whether variants should be classified as significant or not – or "indicative" to use the traditional term. He asks a key question: *is this truly a binary distinction?* Should variants be given various weights, rather than being simply included or excluded? He points out that "we are dealing with words, which are slippery things" and therefore the significance of a variant can be hard to determine.[58] His key point here is that "we can have confidence in [phylogenetic results] because our analysis does not rest on only these one or two variants ('indicative' as they might be), but on patterns within the whole mass of variation. We may be wrong sometimes about the classification of particular variants, but the great majority of times, we will be right."[59] Therefore applying phylogenetics to manuscript traditions is indeed valuable. A consequence of Robinson's point is that it is important to include the vast array of seemingly less indicative variants (regularised appropriately) to provide a weight of evidence for the phylogenetic software to use. Robinson and Andrews seem to disagree, to some extent, on the issue of the significance of variants – but their conclusions agree: It would be wrong to exclude variants for phylogenetic analysis merely because they seem less "significant" than others.

Halonen makes a bold claim, that in the future "there will be no need to compare stemma created by traditional and computer-assisted methods, one can simply rely on the

---

[57] See Halonen, 'Computer-Assisted Stemmatology in Studying Paulus Juusten's 16th-Century Chronicle *Catalogus et Ordinaria Successio Episcoporum Finlandensium*'.

[58] Robinson, 'Four Rules for the Application of Phylogenetics in the Analysis of Textual Traditions', 646.

[59] Ibid., 649.

latter."[60] Is this really likely? This Chapter, and Chapter 6 in particular will show that such computer-assisted methods can indeed identify relationships and trees that are very close to hand-made stemmata. He clearly believes that this time will come, although it will "take a long time and requires numerous stages of testing and retesting..."[61] He points out that transcription and normalisation of witnesses must still be carried out manually.

The "Thematic Section on Studia Stemmatologica" also includes papers describing the successful application of phylogenetics to manuscript traditions (and in one case the evolution of a folk tale from its oral beginnings). We will consider the final paper by Buzzoni, Burgio, Modena and Simion: "Open versus closed recensions (Pasquali): Pros and cons of some methods for computer-assisted stemmatology." A simple definition is that a closed recension has no contamination and therefore all manuscripts are copies of exactly one exemplar. An open recension, clearly therefore, has contamination – and this is clearly likely to be the reality for any significant corpus. Buzzoni *et al.* describe two case studies (one open and one closed) and the results of analysing them using phylogenetic software. For the open recension (*The Anglo-Saxon Chronicle*) they compare the results of using RHM, Stemstem, NeighborJoining and NeighborNet.[62] For a closed recension they chose the Latin *Meloine* tradition and applied RHM, Stemstem, NeighborJoining (as before) and also Maximum Parsimony.[63] PAUP was used for the Maximum Parsimony algorithm.[64] They conclude, for open recensions, with the thought that while these phylogenetic techniques have produced

---

[60] Halonen, 'Computer-Assisted Stemmatology in Studying Paulus Juusten's 16th-Century Chronicle *Catalogus et Ordinaria Successio Episcoporum Finlandensium*', 588.

[61] Ibid.

[62] See Roos and Heikkilä, 'Evaluating Methods for Computer-Assisted Stemmatology Using Artificial Benchmark Data Sets' (RHM); Roos and Zou, 'Analysis of Textual Variation by Latent Tree Structures' (Stemstem); Saitou and Nei, 'The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees'; Bryant and Moulton, 'NeighborNet'.

[63] There is a terminology problem here, which is perhaps best solved by considering these methods as proper names. For NeighborJoining is itself a method based on the principle of maximum parsimony (but not the method with the same name). See Saitou and Nei, 'The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees'.

[64] L. Swofford, *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10*.

valuable insights there were weaknesses with each one (i.e. places where the phylogenetic results differed from their philological understanding of the manuscript relationships). For closed recensions a Maximum Parsimony approach produced the best results. This would be of interest to New Testament scholars seeking to apply phylogenetics to a set of manuscripts believed to be free of contamination – and there may be such a set (although I suspect not a large one). For open recensions they found RHM and NeighborJoining produced the best results. They found that phylogenetic techniques were good at identifying major groups of witnesses, but less useful for deciding the priority of variant readings. The good news for the application of phylogenetics to the New Testament is that Buzzoni et al. found that phylogenetics worked better on open recensions than closed.

Robinson's paper "Four rules for the application of phylogenetics in the analysis of textual traditions" is to be recommended as an introduction to the topic, as indicated by the first sentence of the abstract: "This article surveys how phylogenetics may be applied effectively and productively to the analysis of textual traditions—and, by implication, how it might not be."[65] Robinson points out two key facts, that phylogenetic tools are powerful and useful for this field, but that their results are hard to interpret correctly. His four rules are:

- "It is an error to understand the results of any quantitative analysis of textual traditions as if they represent exactly what happened in the actual making of these copies."[66]
- "Quantitative methods can give useful results on vernacular and other traditions, where the spelling of individual words is relatively standard across copies, and can do so on the basis of computer-generated alignment and analysis alone."[67]
- "Before any analysis is done, indeed, before any collation, any transcription, any data

---

[65] Robinson, 'Four Rules for the Application of Phylogenetics in the Analysis of Textual Traditions', 637.
[66] Ibid., 641.
[67] Ibid.

preparation is started toward investigation of a textual tradition, you must develop an explicit model of the variation you expect to find in the copies which constitute that tradition."[68]

- "One can only be as certain, in any reconstruction of any textual tradition, as the data, the model of variation, the methods applied, and other evidence allow."[69]

My analysis below and in Chapter 5 (carried out before Robinson's paper was published) agrees broadly with Robinson's rules. His neatly phrased concluding thought is that phylogenetics and traditional philology should be used together:

> *"Phylogenetics may surprise us into new knowledge; it may confirm what we already suspect. But in textual traditions, no phylogenetic account, on its own, can ever be complete: phylogenetics needs philology. The counter is that philologists may use phylogenetics to question and confirm. It is a fundamental of philology, as with any science, that all conclusions are subject to test. Usually, the test is against what we know of manuscripts, of scribes, of the author, of meaning. Now, we have a new and remarkable test. We should use it."*[70]

Phylogenetic Software

Stripey

As has been discussed more fully in Chapters 1 and 2, the International Greek New Testament Project (IGNTP) is working in partnership with the Institut für neutestamentliche Textforschung (INTF) in Münster on the *Novum Testamentum Graecum: Editio Critica*

---

[68] Ibid., 645.
[69] Ibid., 649.
[70] Ibid., 650.

Maior (ECM).[71] As part of the process of creating the edition of John's Gospel, the IGNTP have made full text transcriptions of selected manuscripts. These are not all the manuscripts of the works in question, as that would be a prohibitively large amount of work as the transcriptions can only be made manually (at present).[72] Nevertheless the manuscripts chosen are believed to contain all significant textual evidence. A TEI-compatible XML format is used to encode the transcriptions and these XML transcriptions are the starting point for computerized analysis of the data.[73]

In the first weeks of my doctoral studies I began to write a small prototype computer program (called *Stripey*) to analyse manuscript transcriptions. The idea was to explore ways of visualizing the readings found in different texts in such a way that the texts (including corrections) could easily be compared – inspired by DNA sequences. Table 79 contains some example early output from *Stripey* showing Codex Sinaiticus's firsthand and five correctors. Each letter represents the state of the text in one verse of John 1. *Stripey* is named after these rows, or stripes, of letters that are not unlike DNA sequences – a core principle used in NEXUS files used by phylogenetic software.[74]

```
01 (firsthand)   abbbabbbbbbbcbcbbbbbbbbbbbbbcbbbbbbbbbbbbbbbbbbbbbb
01 (S1)               c        d  d  c    c     c     c
01 (S2)                        e
01 (ca)            a  c   a  d g cc  d    cc  c   c  cd cc ca cd   c
01 (cb2)                       f        a   c               c
01 (corrector)                                                 c
```

Table 79: Correctors as states of the text in John 1 of Codex Sinaiticus

Even in its early stages *Stripey's* output raised research questions. For example it was

---

[71] The Institute for Textual Scholarship and Electronic Editing (ITSEE) in Birmingham is responsible for producing the Greek and Latin editions of John (and later Paul) for the IGNTP.

[72] If someone could write a reliable automatic Greek manuscript transcription tool, then the benefits to the field would be immense.

[73] TEI is the "Text Encoding Initiative". It produces standards for representing digital texts. See http://www.tei-c.org.

[74] See, for example, Huson and Bryant, 'Application of Phylogenetic Networks in Evolutionary Studies'; Maddison, Swofford, and Maddison, 'Nexus'.

immediately clear from Table 79 that one verse in particular has attracted the attention of all but one of the correctors who were active in this chapter (verse 15). Correctors S1, S2, ca and cb2 all changed the text in the middle of the verse to a different form.[75] The firsthand reads ιωαννης μαρτυρι περι αυτου και κεκραγεν ουτος ην ο οπισω μου ερχομενος ος εμπροσθεν μου γεγονεν οτι πρωτος μου ην. The scriptorium corrector S1 added three words, changing κεκραγεν ουτος ην ο οπισω μου into κεκραγεν λεγων ουτος ην ο ειπων ο οπισω μου. The second scriptorium corrector S2 then changed ο ειπων to ον ειπων. Then between the fifth and seventh centuries two later corrections were made: Corrector ca deleted ος from a few words later in the verse, and corrector cb2 further changed ειπων to ειπον.

Stripey has developed significantly since the early prototype: It is now a database-driven web application into which can be imported the XML transcription data. *Stripey* allows the user to explore new ways of analysing and visualizing the data from the XML transcriptions.[76] *Stripey* no longer uses verses as the smallest units of text, but instead uses automatic collation to create a much more fine grained dataset. It uses the CollateX software, which was written “with the aim to create a successor of Peter Robinson's Collate”.[77] CollateX provides a solution for comparing texts that is analogous to the sequence alignment methods used in bioinformatics.

A straightforward step from the stripe display of identifiers in Table 79 was to make *Stripey* create NEXUS files, for use by phylogenetic software.[78] An abridged NEXUS file representing John 1:1 is shown in Table 80. The data is now ready to be input into phylogenetic software such as *SplitsTree.*

---

[75] See ‘Codex Sinaiticus: Experience the Oldest Bible. The Transcription.’ for information about 01’s scribes and correctors.

[76] The current version of *Stripey* is v1.2, See https://doi.org/10.5281/zenodo.1296295 or https://github.com/edmondac/*stripey*

[77] See https://collatex.net/about/ [Accessed 30 December 2016]

[78] See Maddison, Swofford, and Maddison, ‘Nexus’.

```
#nexus
BEGIN Taxa;
DIMENSIONS ntax=153;
TAXLABELS
_firsthand
01_firsthand
011_firsthand
[148 rows hidden]
P66_firsthand
P75_firsthand
;
END;
BEGIN Characters;
DIMENSIONS nchar=7;

FORMAT
    datatype=STANDARD
    missing=-
    gap=?
    symbols="a b c d"
;
MATRIX
_firsthand abaaa?d
01_firsthand aaaaa?a
011_firsthand aaaaa?a
[148 rows hidden]
P66_firsthand aaaaa?a
P75_firsthand aaaaa?a
;
END;
```

Table 80: Abridged NEXUS file representing John 1:1 (Dekker)

SplitsTree

SplitsTree is popular in this field, and like many popular phylogenetic applications is freely downloadable.[79] Once *SplitsTree* has analysed the data it can present it in many different forms. Figure 27 shows a small section of a Neighbour-Joining (NJ) tree (phylogram) for John chapter 1 (*Needleman-Wunsch* algorithm).[80] It is immediately obvious that some expected relationships are present – for example 03 and 𝔓75 are close to one another.[81] Note that Figure 27 represents roughly 7% of the full tree. *SplitsTree* includes many algorithms for generating trees and networks. For example, its UPGMA (Unweighted Pair

---

[79] See Huson and Bryant, ‘Application of Phylogenetic Networks in Evolutionary Studies’. See the bibliography for versions used.

[80] See Saitou and Nei, ‘The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees’. For a good introduction to reading phylogenetic trees, see Rambaut, ‘How to Read a Phylogenetic Tree’. The full tree is available at http://epapers.bham.ac.uk/3148/2/stripey_john_1_all.needleman-wunsch.splitstree.nj-phylogram.pdf

[81] Note that left-right distance represents the similarity (or difference) between manuscripts. Up-down distance has no such meaning, but is simply used to display the tree clearly.

Group Method using Arithmetic averages) tree of the same data is quite different from the NJ tree.

Stripey can also create a NEXUS file for an entire book. Consider the NJ tree for the whole of John (Figure 28), compared to that of just chapter 1.[82] There are obvious similarities (e.g. 03 and 𝔓75 are once again close) but other features are different. This is not surprising, as this tree is based on all 21 chapters of John – equating to 15,767 characters rather than the 769 of John 1.[83]

SplitsTree's trees are binary splits – meaning every branch splits into at most two more. Its trees are therefore always going to be complicated and difficult to interpret. For each different tree algorithm some features of interest can be identified by a textual scholar, but there is so much noise that they are hard to find. *SplitsTree* can also produce phylogenetic networks, but they are impenetrably complex with this number of manuscripts and cannot be represented well at a printable scale.

---

[82] The full tree is available at http://epapers.bham.ac.uk/3149/1/stripey_john_all.needleman-wunsch.splitstree.nj-phylogram.pdf

[83] A *character* in phylogenetics equates to a variant unit in a collation.

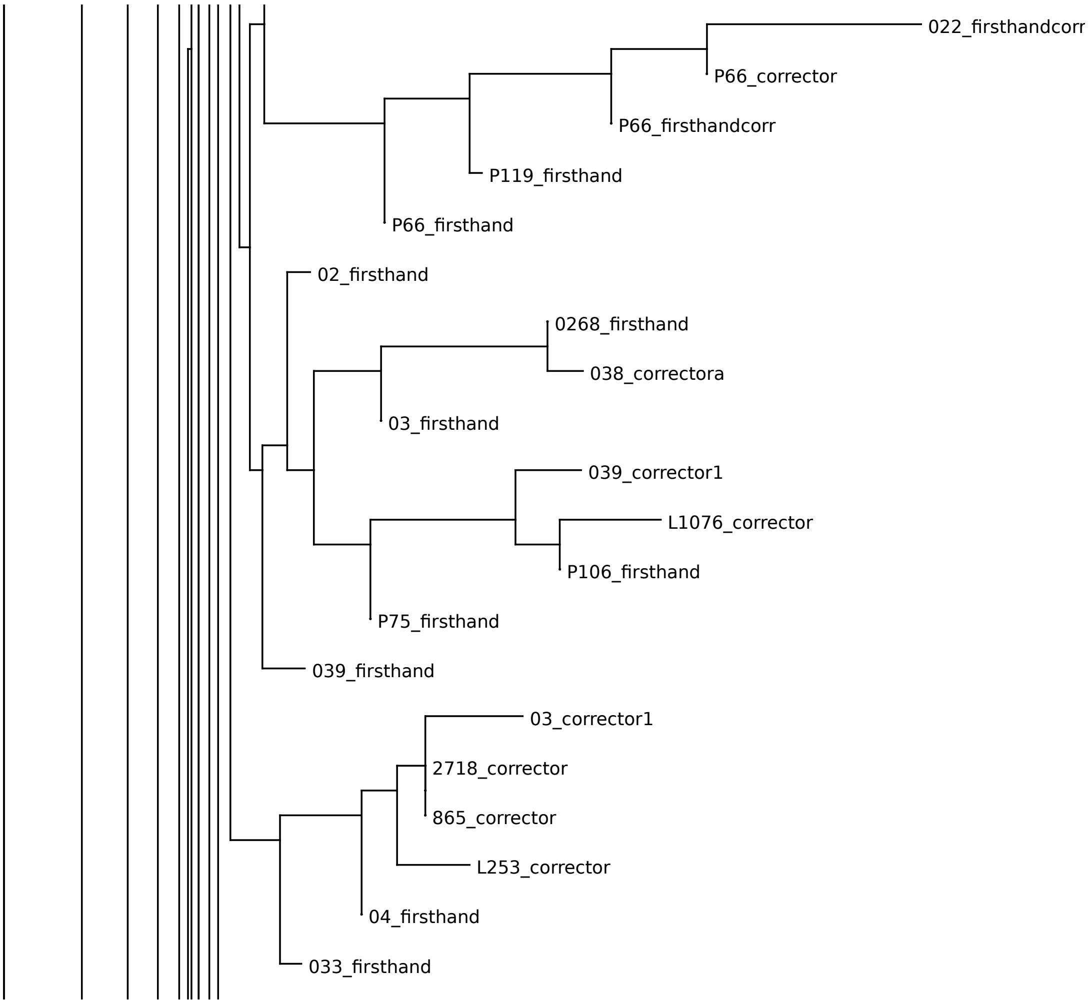


Figure 27: Part of SplitsTree NJ phylogram for John 1 (Needleman-Wunsch)

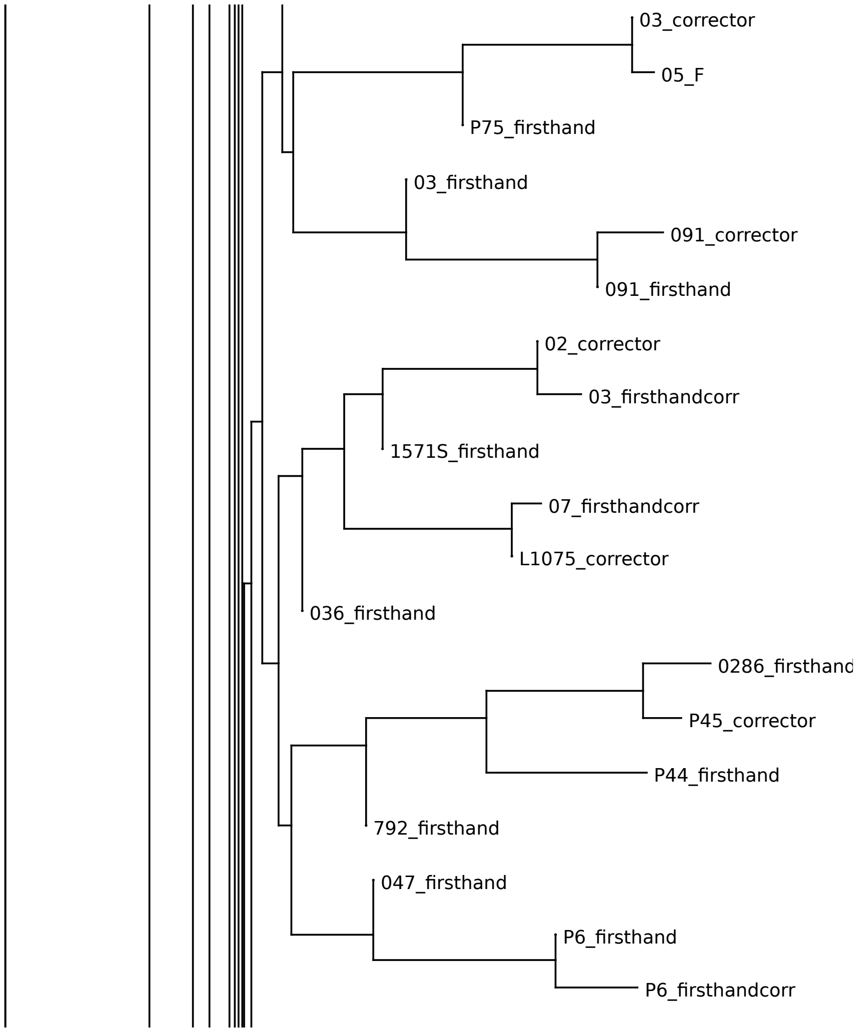


Figure 28: Part of SplitsTree NJ phylogram for John (Needleman-Wunsch)

Mesquite

A second example of free phylogenetic software is *Mesquite*. *Mesquite* is able to create circular trees similar to those used by Spencer et al. in ‘The Greek Vorlage of the Syra Harclensis.’[84] *Mesquite* can take the character matrix in the NEXUS file provided by *Stripey* and, using Maximum Parsimony (MP) algorithms, create a specified number of most parsimonious trees.[85] These may then be further analysed and a consensus tree created

---

[84] Spencer, Wachtel, and Howe, ‘The Greek Vorlage of the Syra Harclensis’. See Chapter 1. See also Wachtel, ‘The Coherence-Based Genealogical Method’, 125f.

[85] In Mesquite, select: “Analysis” > “Tree Influence” > “Tree Search” > “Heuristic (Add & rearrange)”. Select "Treelength" as the criterion for the tree search. Select “SPR” as the rearrangement algorithm as it has the

summarising the previous trees.[86] *Mesquite* uses much more computationally heavy algorithms than *SplitsTree* to create its trees (several hours as opposed to a few minutes on my laptop). Whether this extra effort implies a more trustworthy result is unclear.

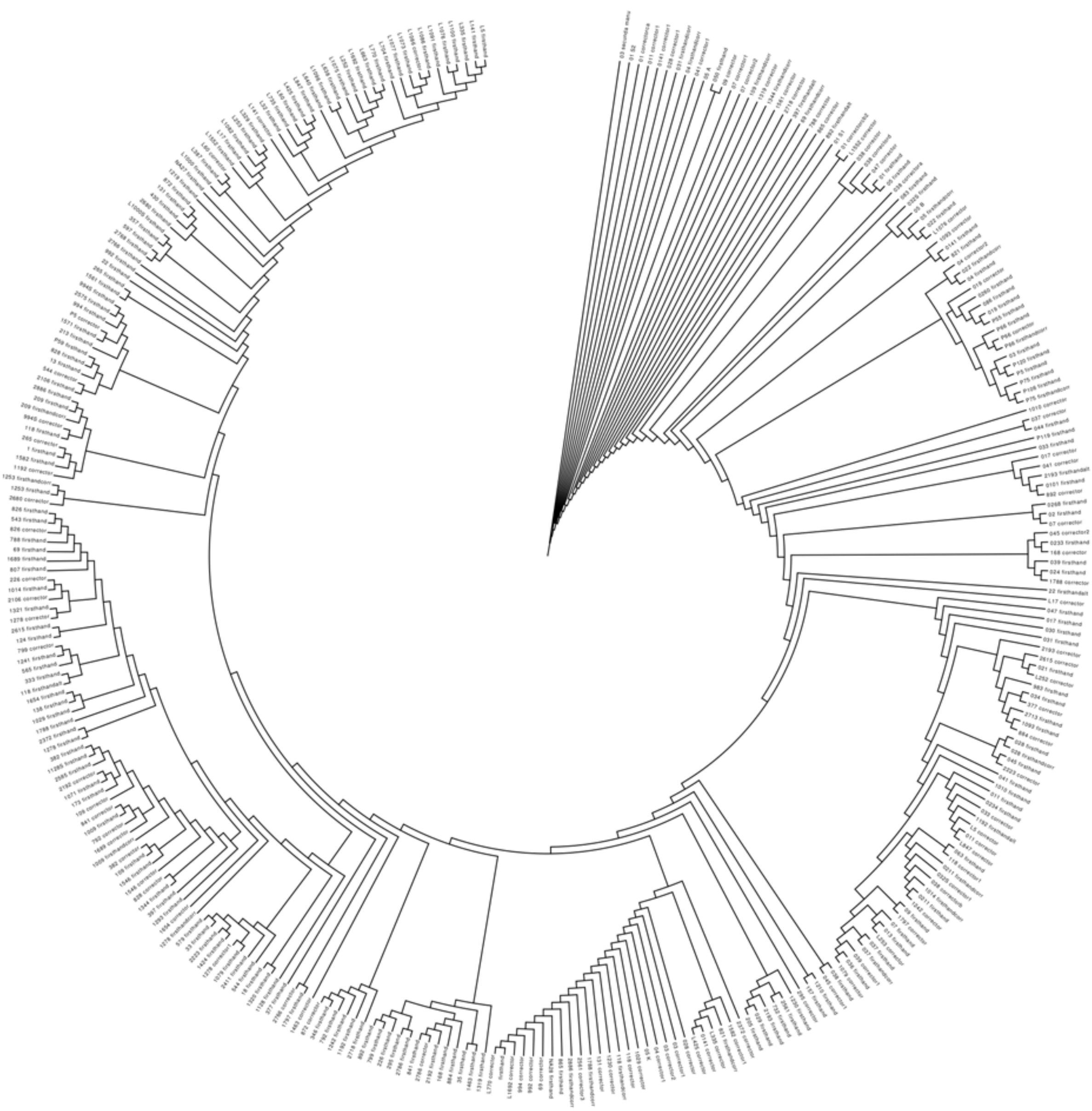


Figure 29: Mesquite consensus tree of MP trees of John 1 (Needleman-Wunsch)

Using the same NEXUS file used by *SplitsTree* to create the tree in Figure 27,

---

best chance of finding the most parsimonious tree(s) and choose the default MAXTREES=100. See 'Mesquite FAQ'.

[86] A "Majority Rules Consensus Tree" shows those clades present in the defined fraction (default > 0.50) of the trees.

Mesquite creates that in Figure 29.[87] Figure 30 shows (at a readable scale) the part of the tree corresponding to Figure 27 - and there are similarities (e.g. 03 and 𝔓75 again). However, there are also many differences, leaving open the question of whether either of them is correct. Unfortunately, *Mesquite's* documentation states this: "*Mesquite* is not designed for rigorous tree-estimation procedures, such as Maximum Parsimony, Maximum Likelihood, or Bayesian tree estimation. However, trees generated from programs such as PAUP, PHYLIP, and MrBayes can be read into and manipulated in Mesquite."[88] So let us now move on to try *MrBayes*.

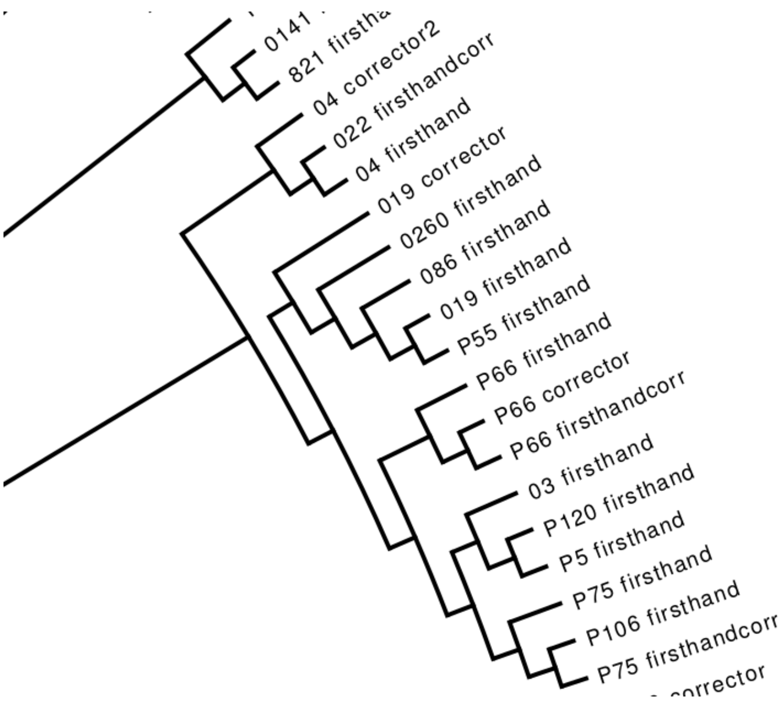


Figure 30: Part of the Mesquite consensus tree of MP trees of John 1 (Needleman-Wunsch)

MrBayes

MrBayes is extremely popular and powerful software for carrying out Bayesian Markov chain Monte Carlo (MCMC) analysis for phylogenetics. Unfortunately, *MrBayes*

---

[87] This tree is available at http://epapers.bham.ac.uk/3150/1/mesquite_john_1_tree.pdf

[88] See 'Mesquite FAQ' (Question: How do I create a publication-quality tree?).

does not allow custom symbols to be used, with only {0, 1, 2, 3, 4, 5, 6, 5, 7, 8, 9} being allowed for standard (morphology) data. Some other types (e.g. RNA) allow larger sets but none is big enough for the number of symbols required to represent (for example) the majority of NEXUS files created by *Stripey* for the Gospel of John.[89]

Nevertheless, a NEXUS file for a single verse is often simple enough for *MrBayes*. Figure 31 shows the consensus tree for John 2:1 (Needleman-Wunsch) using the version of *MrBayes* shipped with *Ubuntu Linux*, run for 1 million generations.[90] This circular representation of the tree has been drawn using *FigTree* - which can take *MrBayes'* output and draw high-quality trees of different types.[91] It is important to note that while the trees are rooted, the root point is arbitrary and does not represent the actual root of the tradition. Indeed *FigTree* provides a facility for changing the root to any place in the tree.

MrBayes creates a different type of tree to *SplitsTree* and *Mesquite* – inasmuch as it can have multiple leaves hanging off a single branch (as opposed to just two). This allows it to create much more readable trees for text-critical purposes. It does have the same restriction as *SplitsTree* and *Mesquite* in that it assumes that all the supplied taxa will be leaves of the tree – because it assumes that the data provided is from species alive today. This restriction means that the trees created by such phylogenetic software are very unlikely to be a true reflection of the relationships between manuscripts. This is easiest to see when considering a tree containing two manuscripts where we know one is the exemplar for the other. In that case such phylogenetic software will show them as closely related siblings, but cannot show them

---

[89] MrBayes uses the symbol set {A, C, G, U, R, Y, M, K, S, W, H, B, V, D, N} for RNA data.

[90] As of Feb 2017, this version is "mrbayes/xenial,now 3.2.6+dfsg-1build2 amd64" - or simply *MrBayes* v3.2.6.. The complete commands to *MrBayes* were:

```
mcmc temp=0.01 ngen=1000000 autotune=yes stoprule=yes stopval=0.01 nchains=4 nruns=2;
sump; sumt Calctreeprobs=yes nruns=2 burninfrac=0.5
```

Some of these settings were determined after the research in Chapters 4 and 5. This tree is available at http://epapers.bham.ac.uk/3151/1/stripey_john_2_1_all.needleman-wunsch.nex.con.tre.pdf

[91] *FigTree* v1.4.2

as parent and child. Nevertheless, these phylogenetic tools can give valuable information about how closely manuscripts are related to one another.

An important point is raised by Figure 31: there are not many branches or sub-trees (rather most manuscripts are attached directly to the trunk) and thus the eye is immediately drawn to the more interesting features. For example the large group of lectionaries at the bottom is clearly separated from the rest of the manuscripts. And indeed while the other witnesses begin with some form of και τη ημερα τη τριτη, the lectionaries read a form of τω καιρω εκεινω instead. The full collation of John 2:1 is shown in Table 81. One weakness of the automatic collation is immediately evident – the first και is split across two columns, or variant units, and the same is true of εν later. This introduces noise into the data and may limit the value of the output.

Patched MrBayes

In order to be able to use *MrBayes* more, since it has proved itself very useful, I patched the source code to allow up to 62 symbols (a-z, A-Z, 0-9) instead of just ten (0-9).[92] Now NEXUS files for much larger data sets can be loaded into my modified *MrBayes.*[93] For example John 2 (the whole chapter) requires a maximum of 23 character states (a-w) for its 386 characters and 254 taxa. The whole consensus tree is unreadable at this scale, even as a circular tree.[94] The bottom portion of the tree (showing all the sub-trees and therefore the most interesting part) is shown in Figure 32.

---

[92] The patch was to change the definition from line 15143 of command.c (int StandID (char nuc)) from the existing symbols (0-9) with values 2^n to (a-z,A-z) with values *n*. See the changeset at: https://github.com/edmondac/MrBayes/commit/a0ad10fd48a0555c6ee627829e414d2799b990de

[93] The version is now standard *MrBayes* v3.2.6 + my patch.

[94] Circular tree representations are frequently criticised anyway for being hard to read. They are a convenient way to show a lot of data on a single page, and they look attractive – but in general they are less useful than, the less spatially-efficient, vertically drawn trees. The full tree is available at http://epapers.bham.ac.uk/3152/1/stripey_john_2_all.needleman-wunsch.nex.con.tre.pdf

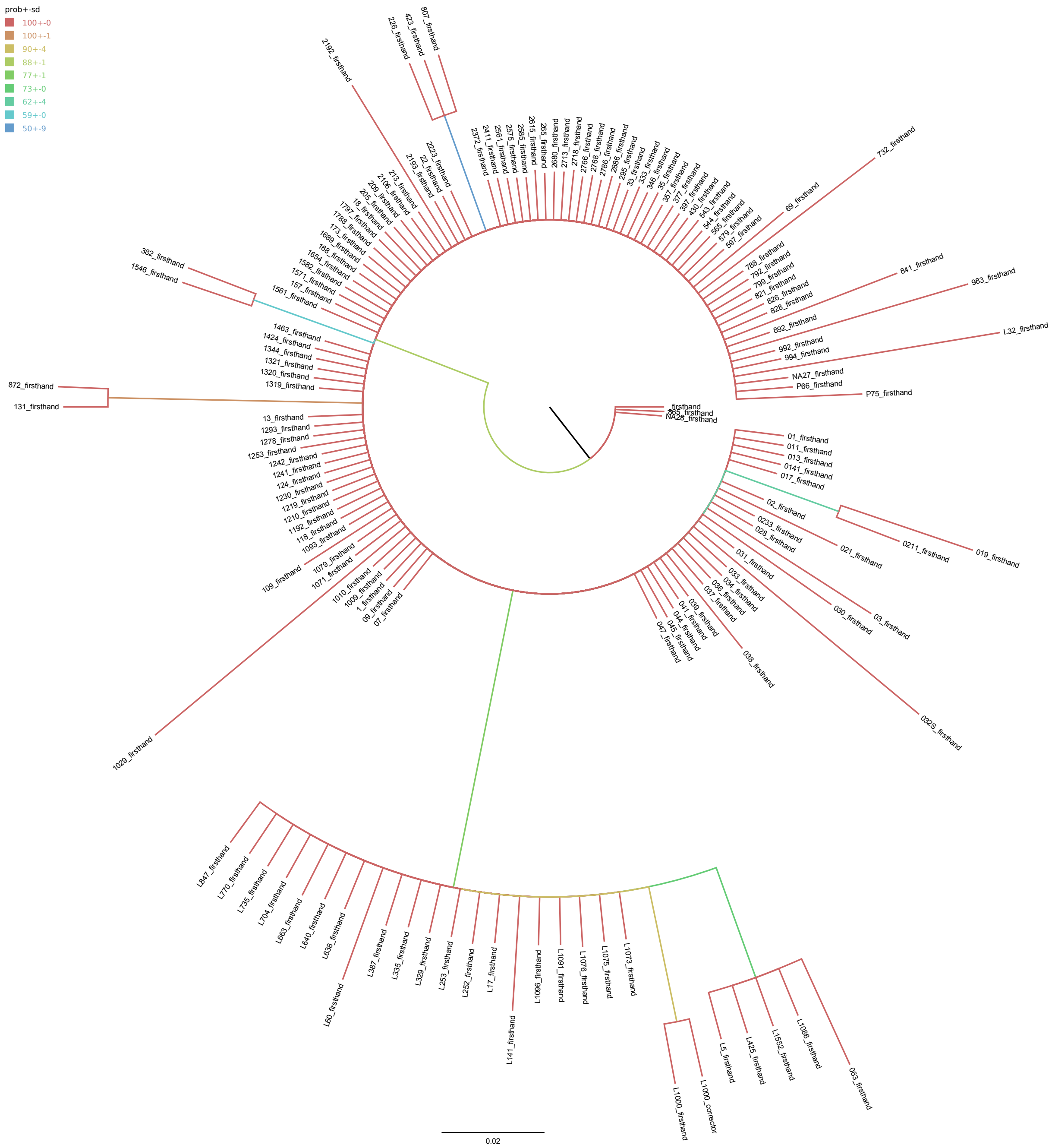


Figure 31: MrBayes consensus tree for John 2:1 (Needleman-Wunsch)

TR NA28 865	100%			και	τη	ημερα	τη	τριτη	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μητηρ του		ιησου εκει	
226 423 807	98.8%			και	τη	ημερα	τη	τριτη	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιησου εκει	
NA27 1071 1424	98.2%			και	τη	ημερα	τη	τριτη	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μητηρ του		ιυ εκει	
P66 01 034	98.2%		και		τη	ημερα	τη	τριτη	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μητηρ του		ιυ εκει	
1 18 22 33 35 118 168 209 213 265 333 357 377 397 430 544 565 579 597 792 799 992 994 1009 1079 1192 1219 1241 1242 1278 1293 1319 1320 1321 1344 1463 1571 1582 1654 1788 1797 2106 2223 2372 2411 2575 2615 2718 2766 2768 2886	97.0%			και	τη	ημερα	τη	τριτη	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
02 07 09 011 013 017 028 031 033 036 037 039 041 044 045 047 0141 0233 157 205 295 821 892 1010 1093 1210 1230 2193 2561 2585 2713 2786	97.0%		και		τη	ημερα	τη	τριτη	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
?75	96.6%		και		τη	ημερα	τη	τριτη	γαμος εγενετο	εν	τη	κανα		της	γαλιλαιας	και ην η μητηρ του		ιυ εκει	
841	95.9%			και	τη	ημερα	τη	τριτη	γαμος εγενετο		εν	γκανα		της	γαλλιλαιας	και ην η μηρ του		ιυ εκει	
382 1546	95.9%			και	τη	ημερα	τη	τριτη	γαμος εγενετω		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
0211	95.9%		και		τη	ημερα	τη	τριτη	γαμος εγενετο		εν	κανα		της	γαληλαιας	και ην η μηρ του		ιυ εκει	
2192	95.8%			και	τη	ημερα	τη	τριτη	γαμος εγενετο		εν	κανα		τη	γαλιλαια	και ην η μηρ του		ιυ εκει	
019	95.2%		και		τη	ημερα	τη	τριτη	γαμος εγενετο		εν	κανα		της	γαληλαιας	και ην η μηρ ου		ιυ εκει	
131 872	93.5%			και	τη	ημερα	τη	τριτη	γαμοσ εγενετο		εν	κανα		τησ	γαλιλαιασ	και ην η μηρ του		ιυ εκει	
1253 1561	93.3%			και	τη	ημερα	τη	γ	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
021	92.6%					και	τη	τριτη	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μητηρ του		ιυ εκει	

69	89.9%				και	τη	τριτη	ημερα	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιησου εκει	
032S	89.6%		και		τη	ημερα	τη	γ	γαμος εγινετο		εν	κανα		της	γαλιλεας	και ην η μηρ του		ιυ εκι	
109	89.3%				και	τη	τριτη	ημερα	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μητηρ του		ιυ εκει	
03	88.8%		και		τη		τριτη	ημερα	γαμος εγενετο		εν	κανα		της	γαλειλαιας	και ην η μητηρ του		ιυ εκει	
983	88.1%	και	τη		τριτη	ημερα		ο	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
13 124 173 346 543 788 826 828 1689 2680	88.0%				και	τη	τριτη	ημερα	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
038	88.0%		και		τη		τριτι	ημερα	γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
030	85.9%		και		τη		τριτη	ημερα	γαμος εγενετο	εν	κανα	της		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
L1000 corrector	82.7%				ω	καιρω	εκεινω		γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
L17 L252 L253 L329 L335 L387 L638 L640 L663 L704 L735 L770 L847 L1073 L1075 L1076 L1091 L1096	82.2%				τω	καιρω	εκεινω		γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
L5 L425 L1086 L1552	81.0%				τω	καιρω	εκεινω		γαμος εγενετο		εν	κανα		της	γαληλαιας	και ην η μηρ του		ιυ εκει	
L32	80.7%				τωι	καιρωι	εκεινωι		γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
L141	80.2%				τω	καιρω	εκεινω		γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκη	
063	79.8%				τω	καιρω	εκεινω		γαμος εγενετω		εν	κανα		της	γαληλαιας	και ην η μηρ του		ιυ εκει	
L1000	79.1%	ιπεν	ο	κς	ω	καιρω	εκεινω		γαμος εγενετο		εν	κανα		της	γαλιλαιας	και ην η μηρ του		ιυ εκει	
L60	79.0%				τω	καιρω	εκεινω		γαμος εγενετο		εν	κανα		της	γαληλεας	και ην η μηρ του		ιυ εκει	
1029	78.1%				και	τη	τριτη	ημερα	γαμος εγενετο	εν	κανα	της	γαλιλαιας	και		ην η μηρ του ιυ	εκει	και οι	αδελφοι αυτου
732	57.4%		και	τη	ημερα	τη	τριτη	γαμος εγενετο											

Table 81: Collation of John 2:1 (Needleman-Wunsch)

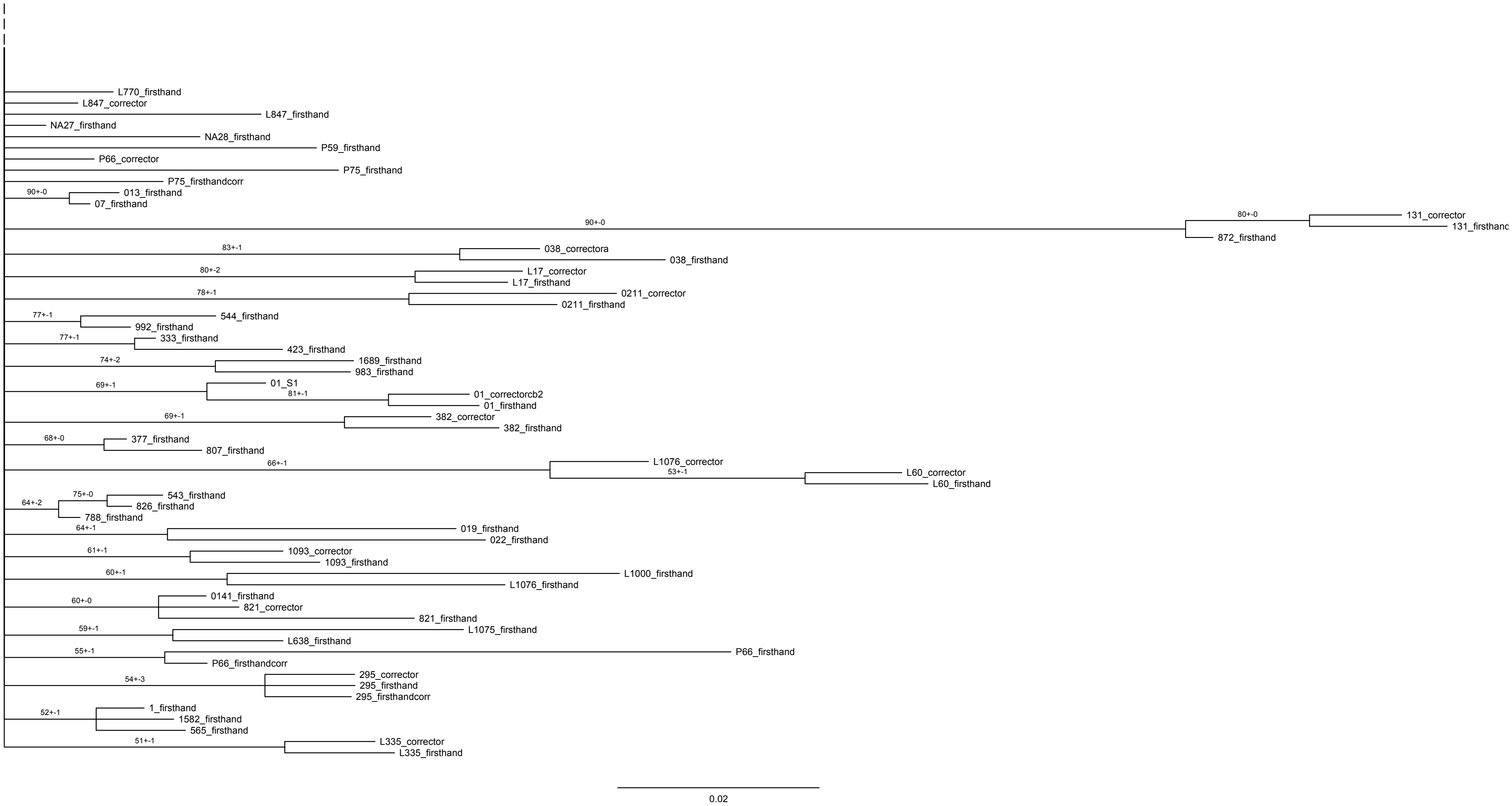


Figure 32: Lower portion of consensus tree for John 2 (Needleman-Wunsch)

Some sub-trees in Figure 32 are unremarkable – for example three hands of 295 are grouped together, as are the firsthand and corrector of each of ℓ17, 038, 0211, and several more. Some groups are more interesting, however:

- 131 and 872 are grouped far from the trunk on a branch with 90% probability.[95] This pair of manuscripts has sometimes been associated with Family 1, but Welsby argues that they should not be in the family.[96] This tree suggests that even if they are not in Family 1, they are closely related to one another.
- The group 1, 565 and 1582 represents a part of Family 1.
- The group 543, 788 and 826 represents a part of Family 13.
- As will be seen in Chapter 6, 821 and 0141 are a copy and its exemplar, 600 years apart.
- The pair 377 and 807 are very similar in John, yet 400 years apart. Could there be a copying relationship here?

Network (Fluxus Engineering)

The primary piece of phylogenetic software recommended to me by Dr. Albert Nobre de Menezes, from the University of Birmingham's Centre for Computational Biology, was *Network*, by Fluxus Engineering. This software can create Median Joining networks, a kind of phylogenetic network.[97] The key difference between phylogenetic trees and networks is that nodes in a network can have more than one parent – whereas in a tree they must have exactly one. This allows the software to represent the complexity caused by contamination in the data and not need to simplify it to a tree. An unfortunate consequence of this is that networks with

---

[95] See Chapter 5 for more about branch probabilities.

[96] See Welsby, 'A Textual Study of Family 1 in The Gospel of John.', 201.

[97] Other software, such as *SplitsTree* and *Mesquite* (above) can also create networks, but Fluxus Engineering's *Network* software is particularly good with the data used in this thesis.

many nodes can quickly become too complex to visualise (and therefore to use). I converted the NEXUS file for John 2 to an amino acid file in Roehl data format, setting the maximum fragmentation level to 35%: i.e. any manuscript that is less than 65% extant is excluded from this network.[98] This is necessary as the Network software excludes any character that is not found in all taxa – and so the analysis is limited to only those variant units found in the intersection of all witnesses. Taxa cannot, apparently, have arbitrary names when using RDF files in *Network*. So we use H_1 (for haplotype 1), H_2 (for haplotype 2) etc. and maintain a mapping from those names back to the manuscripts. These can then be reinstated in the final diagram. This file can then in turn be loaded into *Network*.

A key feature of this software is that, unlike many such programs, taxa/haplotypes from the supplied data can occur as *internal nodes* in the network – not just as leaf nodes. This means it has the *potential* to create networks that represent the actual copying history of the witnesses.

Network is not, in my opinion, particularly good at actually drawing the networks it creates – and they are almost always full of overlapping nodes and edges. However, it is relatively straightforward to convert the FDI output file from *Network* into a DOT file – which can then be used by *Graphviz* (and many other programmes) to visualise the network. The Median Joining network for John 2, drawn using *Graphviz*' SFDP algorithm is far too complex to represent at this scale, even though it is probably the best *Graphviz* algorithm to use for large graphs – but Figure 33 shows a very small corner of it for illustrative purposes nevertheless.[99] The five witnesses (shown in yellow) in the network are: to the left, 333 and then from top to bottom ℓ335, 423, 579 and 1797. The red dots represent hypothetical

---

[98] One character (variant unit) was too complex to represent in RDF – so was excluded. It had 21 character states.

[99] The full network is available at http://epapers.bham.ac.uk/3153/1/stripey_john_2_all.needleman-wunsch.f15.dot.sfdp.pdf

witnesses.

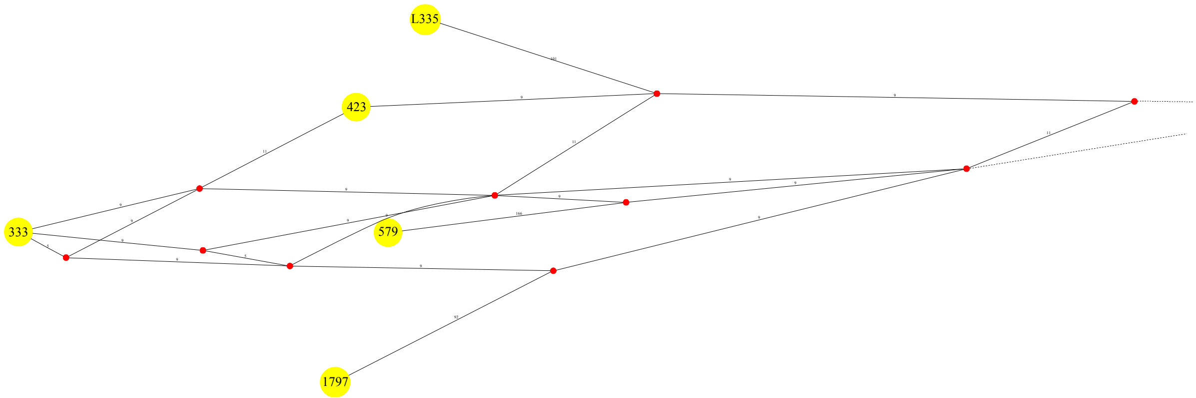


Figure 33: Part of the SFDP *drawing of the John 2 Median Joining network*

See Chapter 5 for some smaller, and much more valuable, Median Joining networks made using small numbers of witnesses from John 18.

Conclusion

These phylogenetic diagrams have brought up many interesting areas for textual study, which is typical when data is visualised in new ways. For example a pair of witnesses from different locations and different centuries can immediately be identified as closely related in a phylogenetic tree – and the human eye spots such patterns easily where it would not see the same pattern when presented as tables of numbers. Thus the tools and techniques discussed in this chapter allow the data to be explored in new ways, resulting in new research questions being raised.[100] The MCMC approach used by *MrBayes* seems to create more interesting and valuable results when including many witnesses (when compared to *SplitsTree* and *Mesquite*) – or perhaps simply more easily decipherable results, which is crucial. *Network* has significant potential, as will be shown in subsequent chapters, but with many witnesses

---

[100] Prof. Mark Viant, discussing a similar phenomenon in Omics studies (Genomics, Metabolomics, etc.) says “Omics studies are intended to be hypothesis generating (also called discovery driven), and this analysis can then lead to targeted, hypothesis driven research questions being addressed.” [Personal email 8/3/17, discussing his lecture at the “NERC-MDIBL Environmental Genomics and Metabolomics 2017” workshop, March 2017] This is analogous to the work of natural historians of previous centuries who set off to explore the unknown world, leading to countless important discoveries.

creates diagrams which are far too complex.

Phylogenetics has been applied successfully to manuscript traditions, as described in the earlier parts of this chapter – albeit commonly ones with relatively small numbers of manuscripts. It is possible that *MrBayes* may be able to create useful and meaningful trees with larger numbers of manuscripts, such as those involved in studying the gospel of John. It seems likely that the automatic collation introduces too much noise into the data – and to find out we will consider the results of running *MrBayes* and *Network* with manually collated texts in Chapter 5.

It is important to note, having considered both the CBGM and phylogenetics now, that while they have many similarities there are also some important differences. The first difference is that while phylogenetics (and indeed traditional stemmatics) introduce hypothetical nodes into their stemmata, the CBGM almost never does. The CBGM has the concept of the "intermediary node" (as can be seen in Chapter 3) but it is rarely necessary to incorporate one into the stemmata. As a result the CBGM's stemmata are generally much simpler than, for example, a phylogenetic network of the same data. A second difference is the manner in which the objects in the trees, networks and stemmata are related. Phylogenetic trees show the evolutionary descent of objects from a common ancestor – and this is true when considering the whole tree or just a small part. The CBGM takes the development of individual variants and combines them to create genealogical coherence tables, and then chooses the best ancestor in a stemma based on the percentage agreement between witnesses. These two tree-creation methods are methodologically distinct. The CBGM deals with texts and not manuscripts and is primarily (but not solely) concerned with identifying the initial text, and does not claim to represent historical manuscript relationships – an oft criticised feature of the CBGM that was discussed in Chapter 2. Conversely, phylogenetic software is good at hypothesising how manuscripts are related and identifying groups, but they are not designed to exactly identify their origin. What remains to be seen is the extent to which the CBGM and Phylogenetics can be used to validate one another's results. This will be explored in Chapter 10.

Chapter 5: Phylogenetics of John 18

Identifying manuscript families – new and well known

Introduction

In this chapter I describe the process of carrying out Bayesian phylogenetic analysis, using *MrBayes*, of sixteen hundred witnesses to John 18. Some remarkable groups of witnesses can be seen in the resulting phylogenetic tree – some already known about such as Families 1 and 13, and some new. A further phylogenetic network technique is then applied to interesting groups and families to shed further light on the relationships between the manuscripts.

Methodology

All 1,663 extant witnesses to John 18 were manually collated by the International Greek New Testament Project (IGNTP), and stored in a database in Münster.[1] I then generated NEXUS files from this database to be analysed by *MrBayes*, using the witnesses as taxa and variant units as characters.[2] The most varied of these variant units has fifty-two different readings (character states), post collation and regularisation. *MrBayes* is extremely popular and powerful software for carrying out Bayesian Markov chain Monte Carlo (MCMC) analysis for phylogenetics.[3] However, *MrBayes* restricts the number of character states for standard (morphology) data to only ten. I therefore had to modify the source code of

---

[1] See Morrill, ‘A Complete Collation and Analysis of All Greek Manuscripts of John 18’.This includes four special witnesses: A , NA27, MTJ18 and TR.

[2] See Maddison, Swofford, and Maddison, ‘Nexus’; Huelsenbeck and Ronquist, ‘MRBAYES’; Ronquist and Huelsenbeck, ‘MrBayes 3’; Altekar et al., ‘Parallel Metropolis Coupled Markov Chain Monte Carlo for Bayesian Phylogenetic Inference’. See Chapter 4 for an introduction to, and an analysis of, using phylogenetic software on manuscript transcription data, and Chapter 6 for an investigation into optimising *MrBayes*.

[3] Technically, *MrBayes* runs Metropolis-coupled MCMC – or MCMCMC, or MC^3.

MrBayes to make it compatible with this dataset. My modified version of *MrBayes* accepts up to sixty-two character states, represented by the lower and upper case Latin letters and digits.

The computations described in this chapter were performed using the University of Birmingham's BlueBEAR HPC service, which provides a High Performance Computing service to the University's research community.[4] I used *MrBayes* to carry out 78M generations of MCMC analysis on the data, in several stages.[5] After each stage I checked for MCMC convergence by considering the average standard deviation of split frequencies (ASDSF), the effective sample size (ESS), and the potential scale reduction factor (PSRF).[6] For this data set, it appears that the MCMC may never achieve an ASDSF of 0.01 (a typical metric for convergence) as Figure 34 shows it levelling off just below 0.11. Figure 35 shows ESS values growing slowly at first and then increasing steadily after approximately 36M generations. By 36M generations the ESS had passed a common rule of thumb value of 200, suggesting convergence. By 78M generations, the ESS had risen to 1,311.

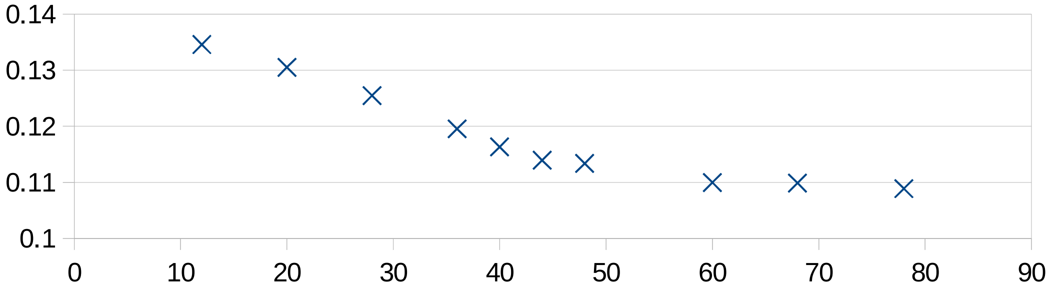


Figure 34: ASDSF by million generations

---

[4] See http://www.birmingham.ac.uk/bear for more details.

[5] This analysis was carried out in March 2015.

[6] *MrBayes*' "sump" command calculates these convergence metrics. For definitions of these terms, and a discussion on the best ways to check for MCMC convergence, see Ronquist and Deans, 'Bayesian Phylogenetics and Its Influence on Insect Systematics'; Gelman and Rubin, 'Inference from Iterative Simulation Using Multiple Sequences'.

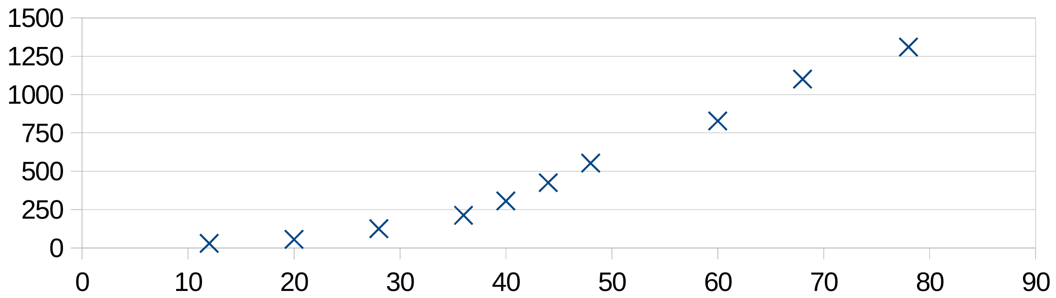


Figure 35: ESS by million generations

Finally, the PSRF (preferred by Ronquist and Deans) had fallen from 1.028 at 12M to 1.000 by 48M generations and measured between 1.000 and 1.003 thereafter.[7] These metrics all suggest that by 36M generations the MCMC analysis has probably converged, and by 48M it can be said to have converged. It is therefore reasonable to expect the following results to be meaningful. They are based on the consensus tree created by *MrBayes* after 48M generations - primarily because the analysis had converged at this point and secondly (and pragmatically) because simply creating the trees after 78M generations requires more than forty-eight hours on BlueBEAR and therefore exceeds the maximum allowed runtime at the time of this analysis. For a further discussion of *MrBayes* convergence and optimising the time required to reach it see Chapter 6.

Results

The size of the resulting consensus tree generated by *MrBayes* representing the complete manual collation of John 18 presents a challenge. A spherical representation of the tree is the only version that could be printed on a single page, and is included as Figure 36.

---

[7] See Ronquist and Deans, 'Bayesian Phylogenetics and Its Influence on Insect Systematics', 191.

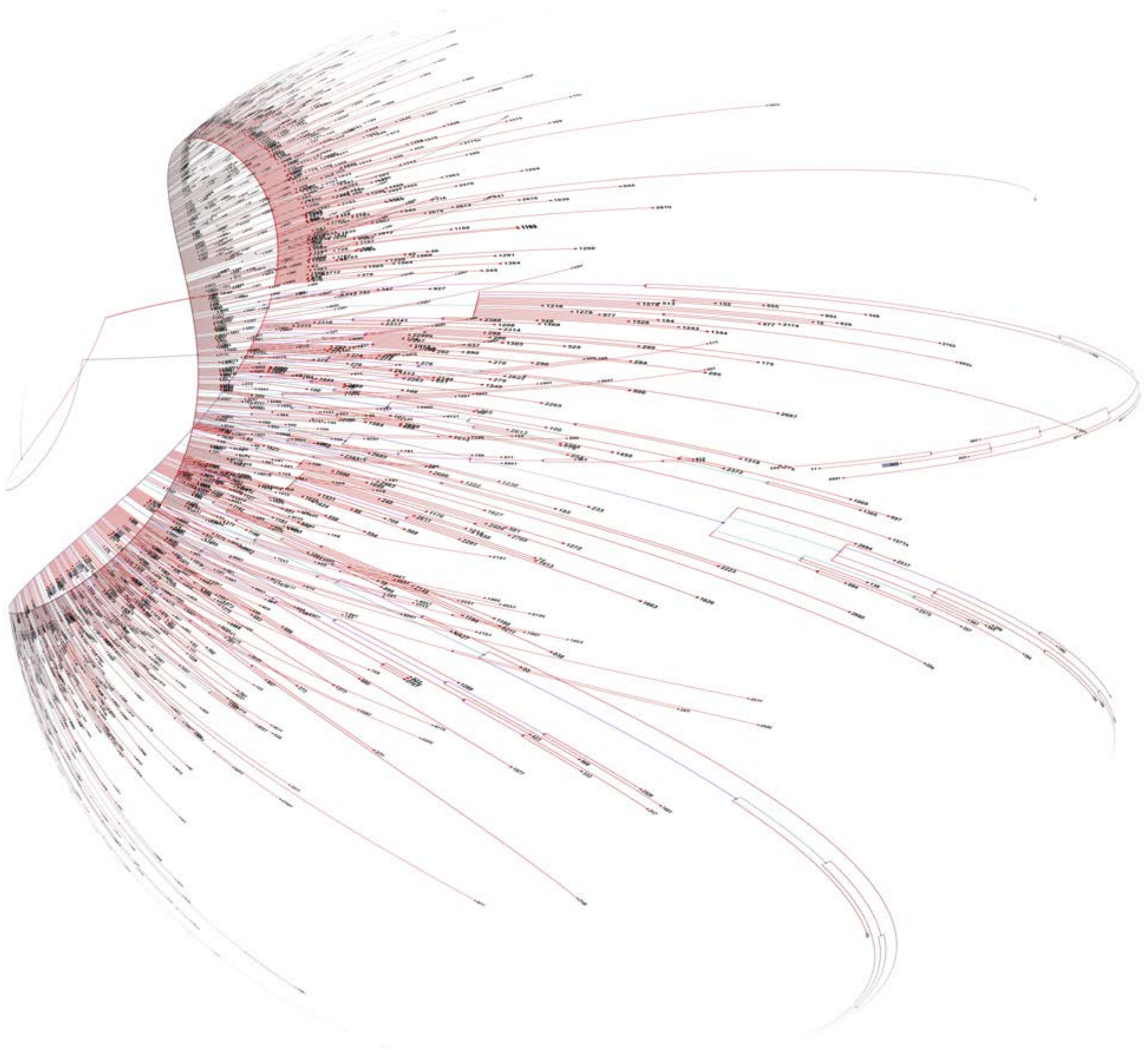


Figure 36: Spherical representation of the consensus tree of John 18

While Figure 36 is visually attractive, it is impossible to examine at this scale. An extract from the traditionally formatted consensus tree is shown in Figure 56, which represents only about 2% of the full tree. As a result the whole tree will not be included here in this format, but instead some interesting features will be presented in individual figures.[8] The bold labels to the right of branches are Gregory-Aland numbers, identifying the

---

[8] The full tree is available here: http://epapers.bham.ac.uk/3105/1/John18MrBayesConsensusTree.pdf

manuscripts. The small labels above branches represent the probability and standard deviation of the branch and thus represent a level of confidence in that branch. Any branch with no label carries 100% probability (i.e. is found in every tree from which the consensus tree was made).

Before proceeding, however, two things should be noted about this tree. First the tree represents closeness of relationship between witnesses and not true parent/child relationships.[9] As such, a pair of witnesses where one was copied from the other will show up as close siblings, but not parent and child.[10] Secondly, the tree is rooted arbitrarily at the *Textus Receptus* as doing so makes it comparatively easy to read. It is possible to represent the same tree with any given witness as the root.

Family 13 (*f*^{13})

At the very top of the full phylogenetic tree, is a familiar group: *f*^{13} or the "Ferrar Group". The relevant sub-tree is shown in Figure 37.

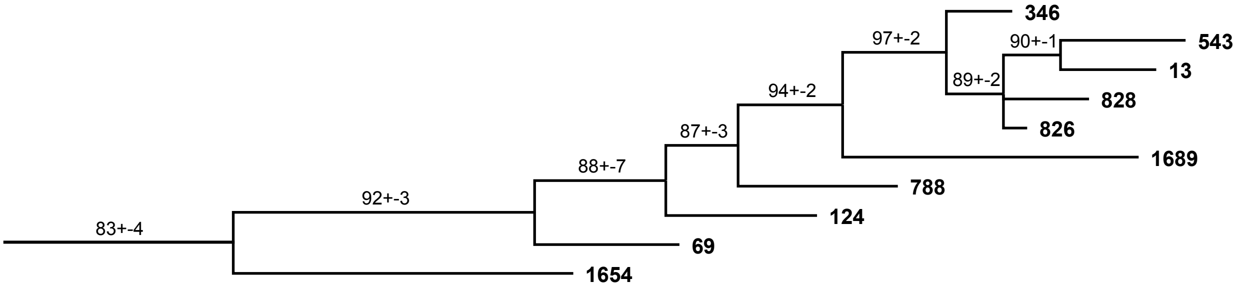


Figure 37: Family 13

NA28 lists *f*^{13} membership (for the whole New Testament) as 13, 69, 124, 174, 230,

---

[9] It is common for phylogenetic software to be based on the assumption that all supplied taxa will represent living organisms. As such it would be expected that these taxa will be leaves of any resultant phylogenetic tree.

[10] See Chapter 4 for a discussion of the merits of visualising data this way to identify such close relationships.

346, 543, 788, 826, 828, 983, 1689, 1709 – ending with "etc.".[11] Jac Perrin established 13, 69, 124, 346, 543, 788, 826, 828, 983, and 1689 as valid members (in John's Gospel) in his PhD thesis in 2012 on *f*^{13} in John – so excluding 174, 230 and 1709 from NA28's list.[12] Didier Lafleur explored five criteria, from the Synoptics, for family membership (also in 2012) and his list agrees with NA28 excluding 1709 and adding ℓ547.[13] Perrin and Lafleur each list a number of other manuscripts that have previously been associated with *f*^{13} but which they argue are not members. Figure 37 shows *f*^{13} as identified by *MrBayes*; a group very similar to Perrin's list. The expected members 13, 69, 124, 346, 543, 788, 826, 828 and 1689 are present. 983 is missing John 11:34-19:9 and thus not present.

Of NA28's list 174, 230, and 1709 are also missing from Figure 37. 174 is lost after John 8 and therefore not present in this analysis. 230 and 1709 are both present but not associated with *f*^{13} and so appear in different parts of the tree.

Figure 37 shows one further manuscript in the sub-tree: 1654. 1654 is not normally associated with *f*^{13}. Visually it is closer to the main trunk of the tree than to *f*^{13} (and the point of such a tree is that it represents closeness or difference by the length of the horizontal lines). In fact 1654 differs from 69 (the *f*^{13} member closest to the trunk) in thirty-three variant units. For comparison, 69 and 124 differ in only seventeen variant units, and the greatest difference between any two members of this sub-tree is twenty-seven variant units. It appears that 1654 is a remote cousin of *f*^{13} at best. It is not easy to answer the question "so why did *MyBayes* put 1654 there?" This highlights something about the way this kind of phylogenetic tree is best used: allow the tree to suggest questions, but do not expect it to provide answers . . .

So, in conclusion, out of the thirteen *f*^{13} members listed in NA28, eleven are extant in

---

[11] See "Witnesses, Signs, Abbreviations" supplement to Aland et al., *NA28*.

[12] See Perrin Jr., 'Family 13 in Saint John's Gospel', 276.

[13] See Lafleur, 'Which Criteria for Family 13 (f ^{13}) Manuscripts?', 140. Lectionaries are not included in this data set, so ℓ547 is not present.

John 18. Nine of these have been successfully identified as a group by this phylogenetic analysis, exactly agreeing with Perrin's analysis. This is encouraging, in that it suggests that the method is valid. It is also good news for any keen supporters of f^{13} as the group has been so clearly identified out of 1,663 witnesses to John 18 – *a chapter devoid of any accepted markers of family membership*.

Median Joining Network

It is possible to carry out further phylogenetic analysis of this group of witnesses representing f^{13}. If we take the subset of the NEXUS file corresponding to f^{13}, and then remove any variant units (i.e. characters) where all nine witnesses agree, and then create a new NEXUS file of just this data we find that it has only sixty characters and no character has more than four states.[14] This is crucial, as this can be represented exactly as DNA data – see Figure 38. A standard NEXUS file representing DNA data will use only the symbols ACGT, along with ? for missing data. In this case, the character states are simply transformed into these letters when creating the NEXUS file. There are many more tools available for analysing NEXUS files of DNA data than there are for morphological data.

```
#NEXUS
begin data;
dimensions ntax=10 nchar=77;
format missing=? gap=- matchchar=. datatype=dna;
matrix
13 AACCCACACCTCCACCCAACAAATCCAATACAACAAAAACCAATAACAACAAACATACACCAAACACAACAAAAAAA
788 AACCCAAAACTCAAACAATCAAAACCAATCCAACAAAACACAATACAAACCACCCAACCCCAACCACAAAAAAAACA
543 AACCCAAAACTCCAACAAACAAAACCACTACAACACAAAACAAAAACAACAACCAGACACCAAACACAAAAAAAAAA
346 AACCCCAAACACCAACAAACAAAATCAATACAAAAAAAAACAATAACAACAACCATACACCCACCCCAAAAAAAAAA
124 C?CACAAAACTCAAACAATCAAAAACCA?CCAACCAAACACAATAAAAACCAACATACACCCCCCGCAAAAACCAAA
828 AACCCAAAACTCCAACAACCAAATCCAATACAACCAAAAACAATAACAACAACAATACACCAACCACAAAAAAAAAA
1689 AACCCATCACTCAAAAAAACAAACCCAACAAAAACAACAACACTCACAACCACCATAAACCCAAAAAAAAC?TAAAA
69 AACCCAAAACACACACACTC??ATCCAAACCA?CAA?ACAAAATAAAAACCACCATACAACCAACACCAAACCCAAA
826 AACCCAAAACTCCAACAAACAAAACCAATACAACAAAAAACAATAACAACA?CCATACACCAACCACAAAAAAAAAA
1654 AAAAAAACAACAAAACACTAAACAAAAACACCCAAAAAAATCACAAACCACACCACCCACACAACTCACAAACCCAC
;
END;
```

Figure 38: NEXUS file for f^{13} in John 18

---

[14] If we also include 1654 then there are 77 characters. However, this witness is a far outlier in the resulting network, and so can be safely discarded at this stage. The network is otherwise unchanged.

```
     12345678911111111112222222222333333333344444444445555555555666666666677777777
              01234567890123456789012345678901234567890123456789012345678901234567

H_1  AACCCACACCTCCACCCAACAAATCCAATACAACAAAAACCAATAACAACAAACATACACCAAACACAACAAAAAAA 1
H_2  AACCCAAAACTCAAACAATCAAAACCAATCCAACAAAACACAATACAAACCACCCAACCCCAACCACAAAAAAAACA 1
H_3  AACCCAAAACTCCAACAAACAAAACCACTACAACACAAAACAAAAACAACAACCAGACACCAAACACAAAAAAAAAA 1
H_4  AACCCCAAACACCAACAAACAAAATCAATACAAAAAAAAACAATAACAACAACCATACACCCACCCCAAAAAAAAAA 1
H_5  CNCACAAAACTCAAACAATCAAAAACCANCCAACCAAACACAATAAAAACCAACATACACCCCCCGCAAAAACCAAA 1
H_6  AACCCAAAACTCCAACAACCAAATCCAATACAACCAAAAACAATAACAACAACAATACACCAACCACAAAAAAAAAA 1
H_7  AACCCATCACTCAAAAAAACAAACCCAACAAAAACAACAACACTCACAACCACCATAAACCCAAAAAAAACNTAAAA 1
H_8  AACCCAAAACACACACACTCNNATCCAAACCANCAANACAAAATAAAAACCACCATACAACCAACACCAAACCCAAA 1
H_9  AACCCAAAACTCCAACAAACAAAACCAATACAACAAAAAACAATAACAACANCCATACACCAACCACAAAAAAAAAA 1
H_10 AAAAAAACAACAAAACACTAAACAAAAACACCCAAAAAAATCACAAACCACACCACCCACACAACTCACAAACCCAC 1

11111111111111111111111111111111111111111111111111111111111111111111111111111
```

Figure 39: Roehl data file for f^{13} in John 18

This entirely standard NEXUS file can be read by many phylogenetic software packages. It can then be converted into Roehl data format (see Figure 39), which can then in turn be loaded into the *Network* software from *Fluxus Engineering* for creating phylogenetic networks.[15] A key feature of this software is that, unlike many such programs, taxa/haplotypes from the supplied data can occur as *internal* nodes in the network – not just as leaf nodes. See Figure 40 for the Median Joining network for f^{13}.

---

[15] See https://doi.org/10.5281/zenodo.1296329 or https://github.com/edmondac/nexus_scripts, for my script `nexus_to_rdf.py` to convert from NEXUS format to Roehl data format. The taxa cannot, apparently, have arbitrary names for this file format when using *Network*. So we use H_1 (for haplotype 1), H_2 (for haplotype 2) etc. and maintain a mapping from those names back to the manuscripts. These can then be reinstated in the final diagram using `relabel_fdi.py` from the same repository.

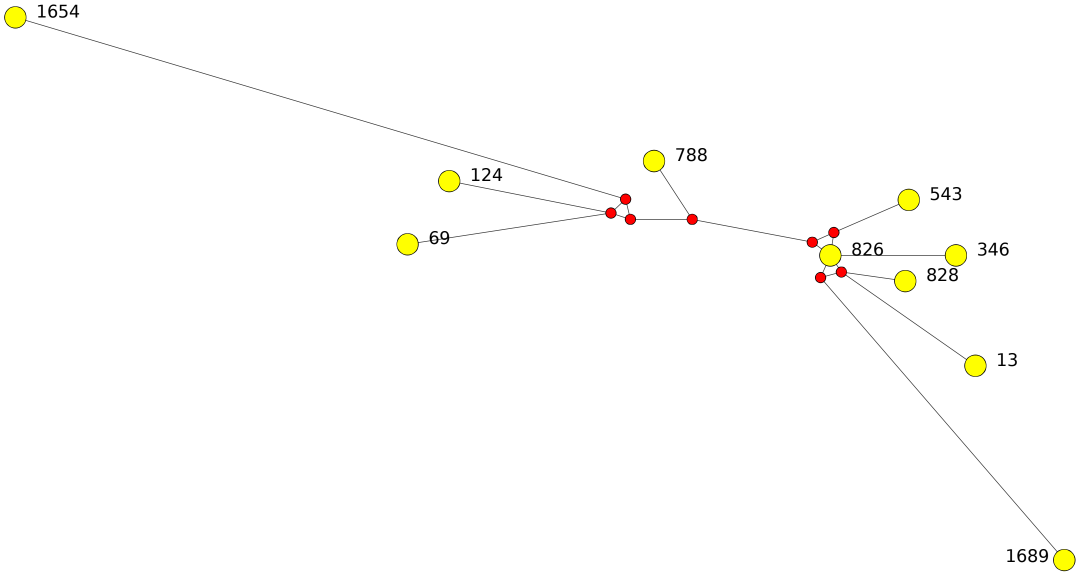


Figure 40: Median Joining Network of f^{13} in John 18

What is striking about this network, is that there is one witness very much at the centre: 826. This, for a bioinformatician, is a clear indication that 826 is the ancestral haplotype, or most recent common ancestor, of this group. Interestingly, this is in broad agreement with the conclusion of Wisse, when using the Claremont Profile Method to analyse Luke. Wisse comments that 826 is the only perfect member of the group, but that either it or 543 could represent f^{13} in an apparatus.[16] He lists 13, 346, 543 and 788 as core members (in addition to 826). Perrin's core group of f^{13} in John is 13, 346, 543, 826, and 828.[17] On the basis of our network in Figure 40 it would seem reasonable to state that the core group is actually the union of Perrin's and Wisse's lists, namely: 13, 346, 543, 788, 826 and 828 – with 826 at the very centre.

---

[16] See Wisse, *The Profile Method for the Classification and Evaluation of Manuscript Evidence, as Applied to the Continuous Greek Text of the Gospel of Luke*, 55ff, 106ff.

[17] See Perrin Jr., 'Family 13 in Saint John's Gospel', 273.

Family 1 (*f*¹)

There is perhaps only one manuscript family as famous as *f*¹³: namely *f*¹ or the "Lake Group". Figure 41 shows that *MrBayes* also found something akin to *f*¹.

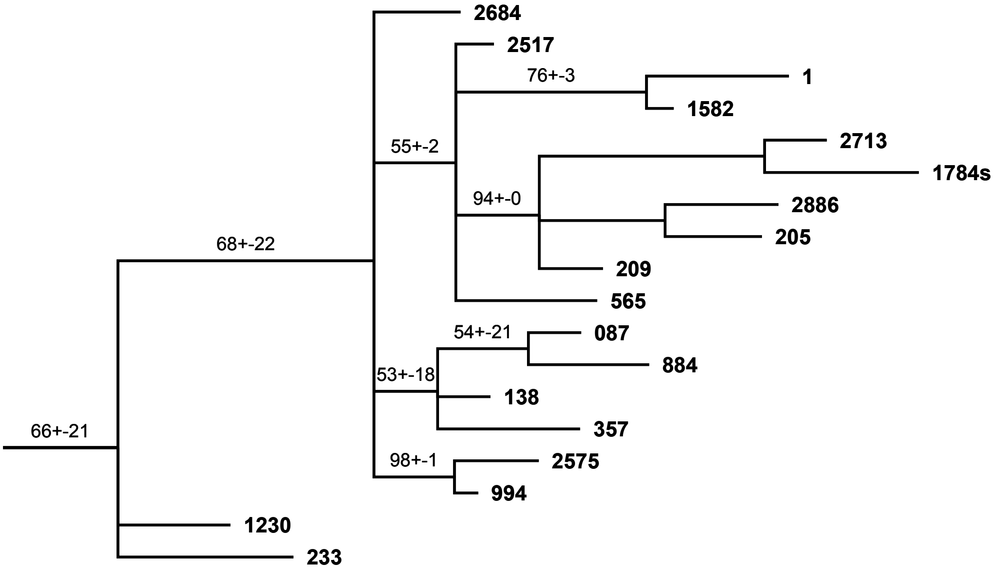


Figure 41: Family 1

NA28 lists *f*¹ membership as "1, 118, 131, 209, 1582, etc."[18] Welsby, in 2011, studied seventeen candidate members of *f*¹ in John. She concluded that fifteen should be considered family members, namely: 1, 22, 118, 205, 205abs (now known as 2886), 209, 565, 884, 1192, 1210, 1278, 1582, 2193, 2372, and 2713. For the remaining two manuscripts, she argued that 131 has only a weak family affinity in John, and 872 has none.[19]

Several of the witnesses listed by Welsby (1, 2886, 205, 209, 565, 884, 1582 and 2713) are grouped together by this phylogenetic analysis along with a number of other manuscripts. Before turning to the others, first let us consider the stemma proposed by Welsby and compare it to our phylogenetic tree. Welsby's stemma of *f*¹ is shown in Figure 42.[20]

---

[18] See "Witnesses, Signs, Abbreviations" supplement to Aland et al., *NA28*.
[19] See Welsby, 'A Textual Study of Family 1 in The Gospel of John.', 201.
[20] Ibid., 202.

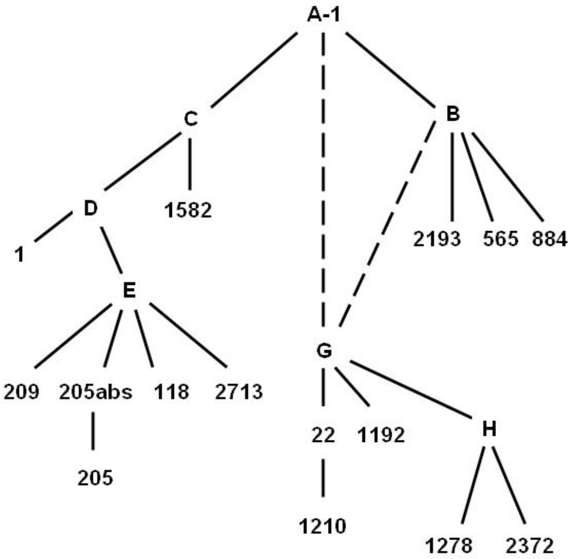


Figure 42: Welsby's stemma of f^1

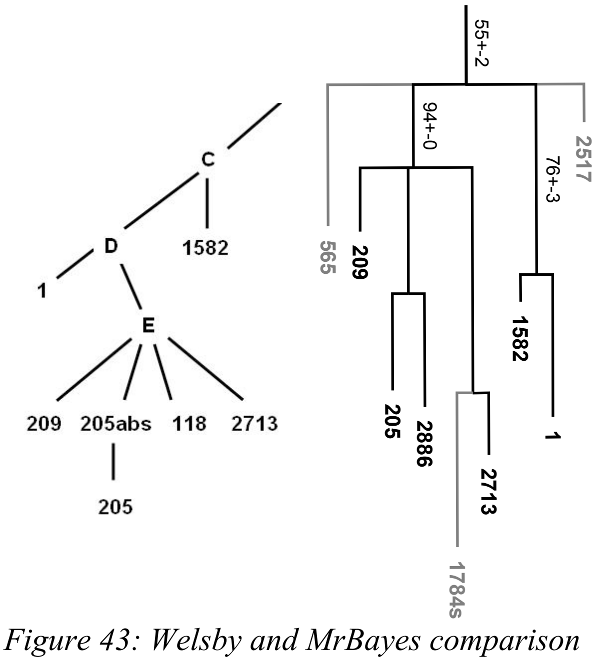


Figure 43: Welsby and MrBayes comparison

Consider the witnesses on the left of Welsby's tree. First, Welsby has 1 as a descendant of 1582 via the hypothetical witness D, where *MrBayes* has them as siblings in Figure 41 – but with 1582 having a much shorter branch length. This phylogenetic analysis only creates hyparchetypes (speciation events) as needed and would not postulate the C-D-E set from Welsby's results. As such this may be considered as an equivalent result. Secondly, Welsby has 2886 (205abs) as 205's parent whereas *MrBayes* again shows them as siblings (on a branch with 100% probability). Since the phylogenetic analysis will always show real witnesses as leaves of the tree this is as close to Welsby's relationship as can be achieved by this method. Thirdly, 209 and 2713 also appear in an equivalent relationship in both trees. Remarkably, these two sub-groups even appear to be related to one another in a similar way in both trees. The similarities between the two trees are clearly shown if the relevant part of Figure 41 is extracted and rotated, extra witnesses greyed out and then the image placed beside the relevant part of Welsby's tree – as in Figure 43.

The only other two witnesses held in common between the two trees are 565 and 884. Welsby has these two as close siblings but Figure 41 shows them somewhat further apart. Of the remaining witnesses, 118 is supplemented in John 18 with 118S and this supplement differs in thirty-two places of variation in John 18 from 1582 (for example). This is the same number of differences as between 1582 and 01 (Codex Sinaiticus). As a comparison, 1 and 1582 differ in only seven places in John 18. Thus 118S is clearly not a member of *f^1*, but this study can say nothing about the family membership of 118 itself.

Manuscript 22 is extant, but grouped with 134 and 1210 (another of Welsby's *f^1* members) but not with *f^1*. 1192, 1278 and 2193 are extant yet appear as individual branches off the trunk and not with *f^1*.

2372 is part of an interesting looking group (see Figure 44) and not with *f^1* although it is lacunose in the last 114 variant units. In John 18, 2372 and 1005 (its closest visual neighbour in Figure 44) differ in only four variant units. By contrast, the minimum number of variant units of difference between 2372 and any witness in Welsby's study and also the tree in Figure 41 is nineteen (with both 884 and 1582). Hence it seems, based only on data from John 18, that 2372 might not belong with *f^1*.

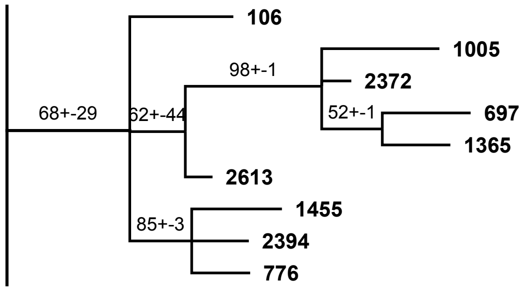


Figure 44: Group containing 2372

The phylogenetic tree in Figure 41 shows ten further manuscripts not listed by Welsby: 087, 138, 223, 357, 994, 1230, 1784S, 2517, 2684 and 3575. Perhaps the most surprising of these is the sixth century majuscule 087. But, 087 is extant in only seven verses of John 18 (29-35), along with fragments of John 19, 20, 21; and some parts of Matthew and Mark. This paucity of data means that the association with *f^1* manuscripts must be treated with caution. 1784S, however, is shown right at the heart of the core *f^1* group (varying from 2713 in only eight places) – was it perhaps copied from a member of this core group to supplement 1784? Further study would be required to answer that question, and also to investigate the claims on family membership for the remaining manuscripts.

A median-joining network can be created for this group of manuscripts like we did for f^{13} above. There are a few key differences, however. First, two manuscripts (087 and 2517) must be excluded as they are sufficiently lacunose that *Network* reports an error with the data. Secondly, the resultant network (Figure 45) does not show an extant manuscript in a central location as it did for f^{13}. Nevertheless the locations of the remaining witnesses in the network show the same agreements with Welsby as did *MrBayes*. Also, interestingly, 1 is shown as descending from 1582; and 2575 from 994.

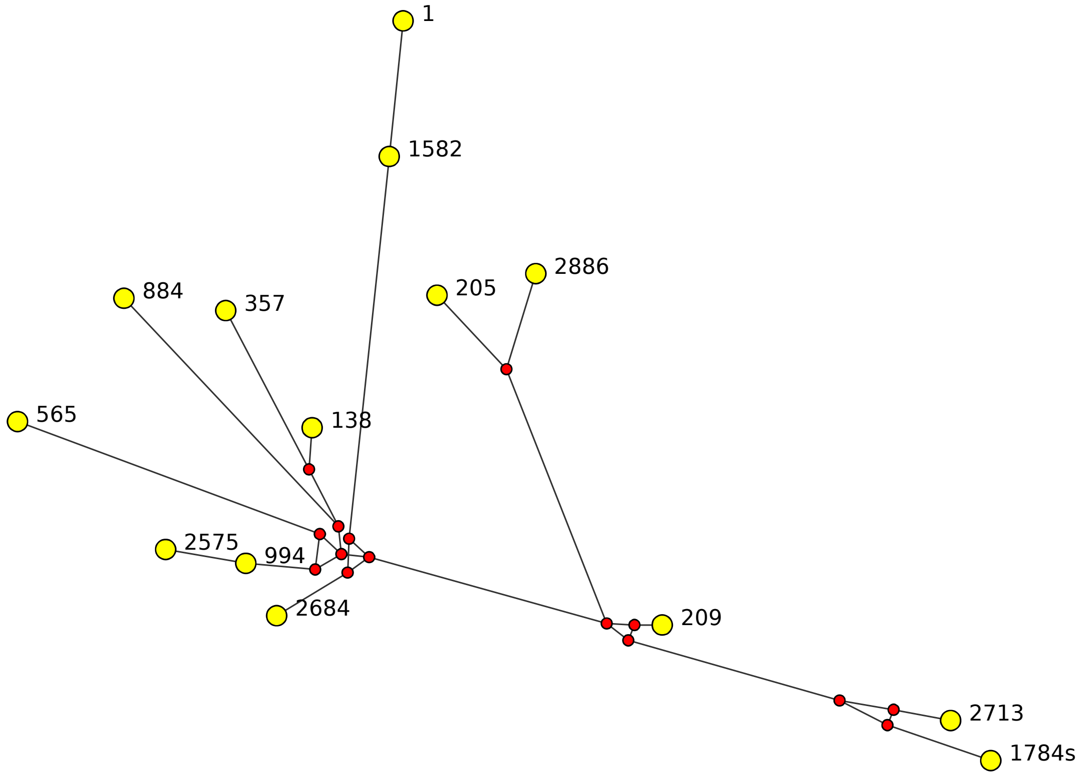


Figure 45: Median-joining network for f^1

A Group of Catena Manuscripts

The full phylogenetic tree created by *MrBayes* contains many interesting sub-trees. These may represent hitherto unknown manuscript groups. They may well, however, just be an artificial product of the algorithms and not represent anything significant about the manuscript relationships. As such, to evaluate each group proposed by *MrBayes* it is necessary to return to the manuscripts themselves. First, we will consider the group of witnesses highlighted with black branches in Figure 46.

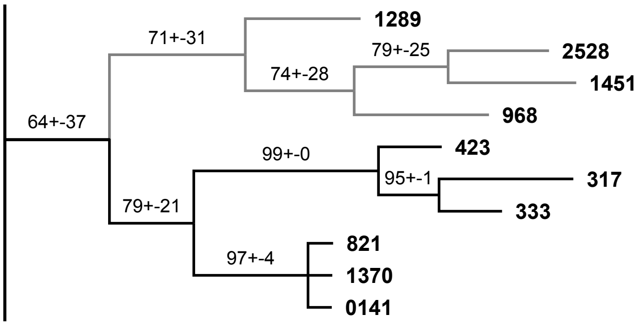


Figure 46: Group 0141

The most striking feature of this group is the trio of Byzantine witnesses at the bottom: 821, 1370 and 0141:

- 0141 is a tenth century catena manuscript in the Bibliothèque Nationale in Paris and written in one column. It has majuscule lemmata (with occasional exceptions) and a minuscule commentary.[21]
- 821 is a sixteenth century minuscule catena manuscript in Biblioteca Nacional de España in Madrid, again written in one column.
- 1370 is a sixteenth century (specifically 1542) minuscule catena manuscript in the Staatsbibliothek zu Berlin, written in two columns.

---

[21] There are a few cases where the scribe has apparently forgotten to change script and instead marked the lemma with diplai or coronis. Taylor Farnes lists a few examples as John 3:3, 3:4, and 3:5. See Taylor Farnes, 'Scribal Habits in Selected New Testament Manuscripts', 225–26.

In John 18, 0141 and 821 have almost identical Biblical text, as also shown by Bruce Morrill.[22] It is worth noting that while Morrill's extensive tables contain the information that would demonstrate that these two manuscripts are closely related it is almost impossible to notice such a relationship. However, visualisation of the relationships (for example in the tree in Figure 46) allows such details to be spotted easily and intuitively. This is one of the important strengths of such a technique. Compare this to a pre-computer age where a scholar would have to travel to libraries in Madrid, Paris and Berlin searching out manuscripts. He would somehow have to realise that a copying relationship might exist between these three – which is virtually impossible unless he was already looking for it. Realistically, this could not have been achieved without computers.[23]

In John 18, 821 and 0141 differ in only two places: In John 18:34/2-4, 821 reads απεκριθη ο ιησους, but 0141 does not have the article; and in John 18:38/22-24, 821 reads παλιν εξηλθεν, but 0141 duplicated the second word (presumably by mistake as it occurs at the end of one page and the start of another). Even in the unregularised, raw data these two manuscripts, *created six hundred years apart*, differ in only four minor places. More remarkable is that the commentary text agrees (at least in the handful of test passages I have inspected). For example consider these pages: 0141 folio 292r; and 821 folio 497v.[24] These pages both contain three biblical lemmata: John 18:31a, John 18:31b-32, and John 18:33-34. So could 821 be a copy of 0141, six centuries later? Alan Taylor Farnes has subsequently explored this question and has conclusively shown that 821 is a direct copy of 0141.[25]

---

[22] See Morrill, 'A Complete Collation and Analysis of All Greek Manuscripts of John 18', 128, 224.

[23] I was corrected here by a visiting noted textual scholar, who said that he would have remembered sufficient detail from one trip to make the connection on a second. So perhaps I should revise my statement to saying that I cannot imagine achieving that feat myself without computers!

[24] For 0141 folio 292r see http://ntvmr.uni-muenster.de/community/modules/papyri/?zoom=38&left=5&top=-906&site=INTF&image=20141/155635/5900/10/2441;
and for 821 folio 497v http://ntvmr.uni-muenster.de/community/modules/papyri/?zoom=63&left=-443&top=-1227&site=INTF&image=30821/155635/4750/10/2321

[25] See Taylor Farnes, 'Scribal Habits in Selected New Testament Manuscripts'.

Results

Sadly, there are no online images of John in 1370 available, to my knowledge.[26] The database asserts that 1370 agrees with 0141 in all but six variant units: In John 18:3/16-30, 1370 removes the second article from *εκ των αρχιερεων και των φαρισαιων υπηρετας*; in John 18:13/2-10, 1370 adds *αυτον* making *και απηγαγον αυτον προς ανναν πρωτον*; in John 18:16/22-28, 0141 reads *ο αλλος μαθητης* and 1370 *ο μαθητης ο αλλος*; in John 18:18/30-48, 1370 removes the *εστως* from *ην δε πετρος μετ αυτων εστως και θερμαινομενος*; in John 18:22/10-16, 0141 reads *εις των παρεστωτων* (the only reading in John 18 that is unique to 821 and 0141) and 1370 *εις των υπηρετων παρεστηκως*; and in John 18:38/22-24, 1370 agrees with 821 as above. Interestingly, in John 18:34/2-4, 1370 agrees with 0141 against 821. It seems plausible, therefore, that 1370 is another copy of 0141 albeit with a different page layout.

In the middle of Figure 46 is another trio of manuscripts: 317 (twelfth century), 333 (thirteenth century) and 423 (sixteenth century), all witnessing to Nicetas of Heraclea's catena. Michael Clark has shown that 333 was the exemplar used by the copyist of 423 and indeed they differ in only five variant units in John 18.[27] 333 and 423 also share a unique (although unremarkable) reading in John 18:29/6-18. All five witnesses in this group (0141, 821, 333, 317 and 423) share another seemingly unremarkable yet otherwise unique reading in John 18:18/30-48: *ην δε πετρος*....

The group of four witnesses greyed out at the top of Figure 46 are 968 (sixteenth century), 1289 (thirteenth century), 1451 (twelfth/thirteenth century) and 2528 (thirteenth/fourteenth century). They are not catena manuscripts and it is not obvious why this group exists. This is a sometimes problematic characteristic of this kind of statistical method; there is no identifiable "reason" behind features in the results. The four witnesses differ

---

[26] As of May 2018, the first few folios of 1370 are now available through the NTVMR.

[27] See Clark, 'The Catena of Nicetas of Heraclea and Its Johannine Text', 29ff.

pairwise in between twelve and seventeen variant units and have no otherwise unique reading to set them apart from the corpus.

While a median-joining network does not add significantly to the results from the Bayesian analysis here, it is nevertheless interesting and confirms those results, and so is included as Figure 47. Consider, for example, the closeness of the pair 821 and 0141, with 1370 just a little distance away.

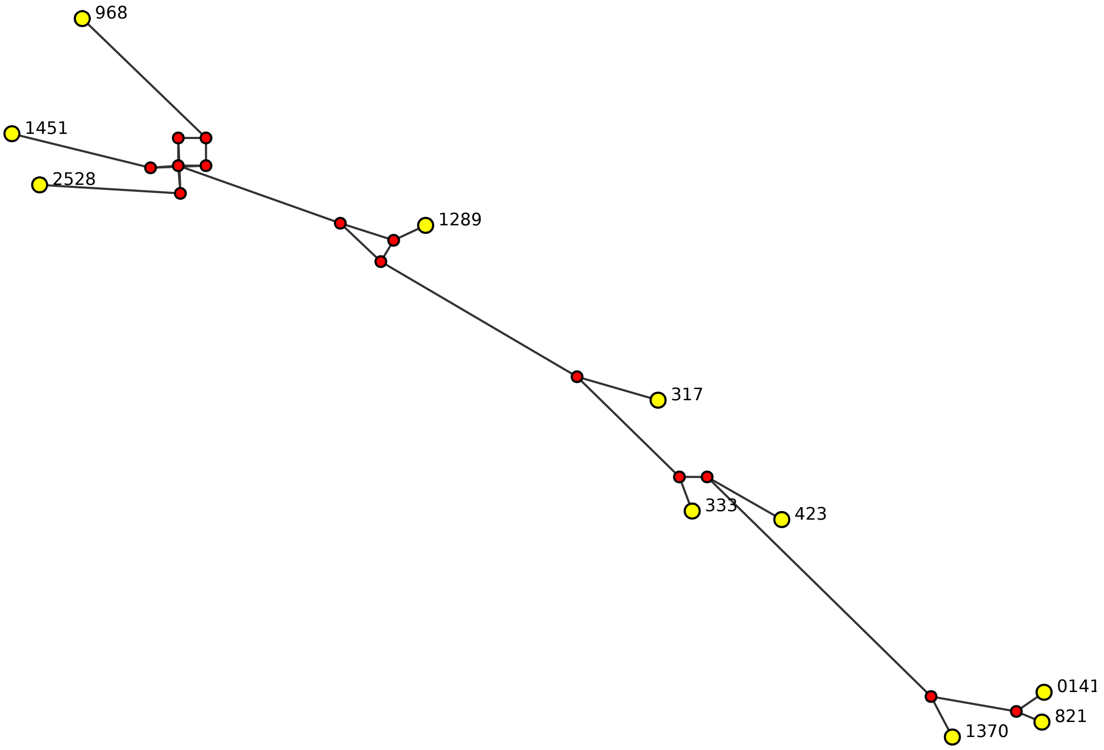


Figure 47: Median joining network of the catena manuscripts

Further Groups

The following figures depict other interesting groups. They are presented here to provide a further taste of the potential of this methodology, and are only briefly commented on. For convenience, each group is named after the lowest numbered member.

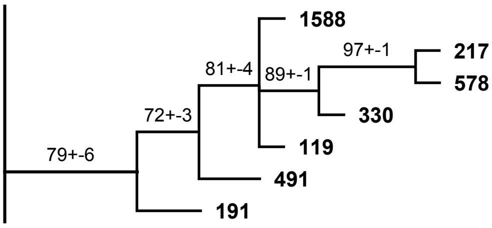


Figure 48: Group 119

Figure 48 shows "Group 119". The pair of witnesses 217 (twelfth century) and 578 (fourteenth century) differ in only one word in John 18 – suggesting a very close relationship. The average pairwise difference in variant units in this group as a whole is only 6.9, with maximum fourteen. More light can be shed on the relationships in this group by creating a median-joining network, which shows that 578 may be descended from 217 which in turn may be descended from from 330. Also 191 is shown as having descended from 1588 (see Figure 49).

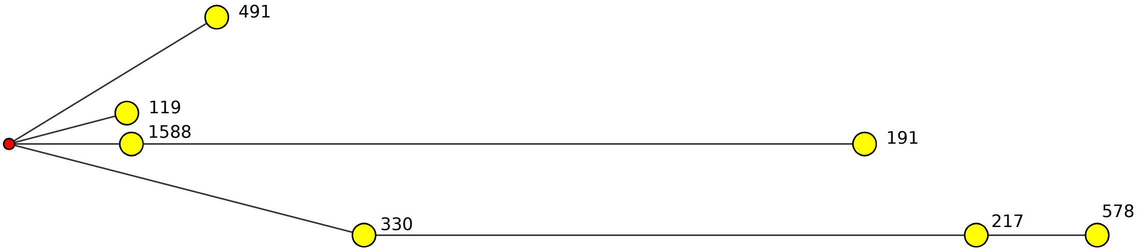


Figure 49: Median-joining network for group 119

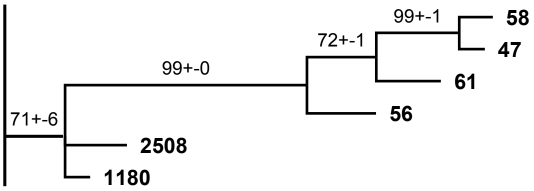


Figure 50: Group 47

Figure 50 shows four witnesses identified as close by Gregory: 47, 56, 58 and 61.[28] Codex Montfortianus (61) was famously created in the sixteenth century in Oxford specifically to force Erasmus to include the Johannine Comma in the *Textus Receptus*.[29] Out of the other three witnesses, in John 18, 61 is closest to 47, differing in only six variant units. Was it perhaps copied from 47, with a handful of changes (deliberate or accidental) drawing it closer to the imminent *Textus Receptus* text? Figure 51 lends weight to this suggestion, showing 61 descended from 47 in the Median Joining network.

---

[28] Gregory, *Textkritik des Neuen Testaments*, p142.

[29] An additional clause in 1 John 5:7-8 [KJV]: 7 For there are three that bear record in heaven, the Father, the Word, and the Holy Ghost: and these three are one. 8 And there are three that bear witness in earth, the Spirit, and the water, and the blood: and these three agree in one.

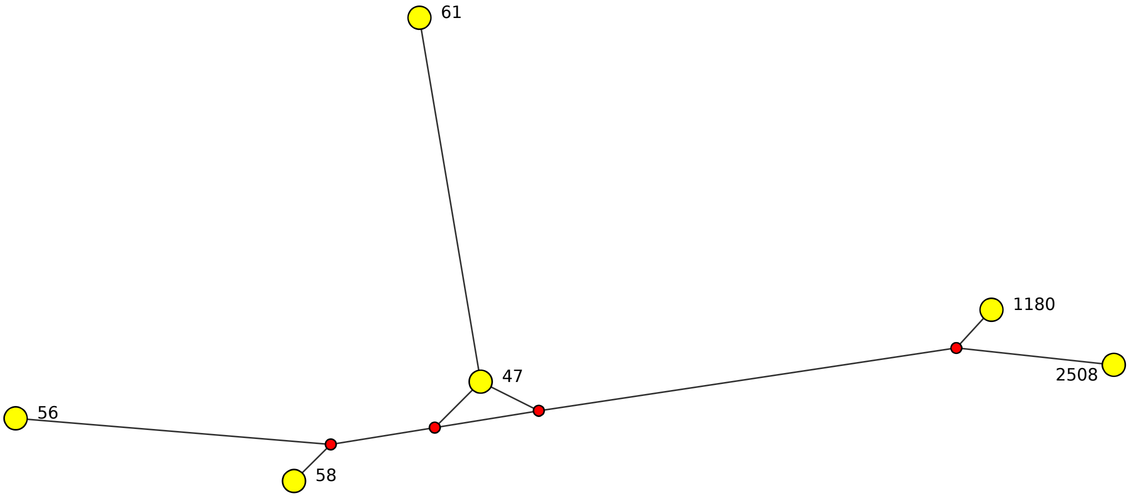


Figure 51: Median-joining network for group 47

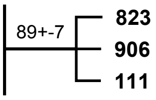


Figure 52: Group 111

Figure 52 shows a trio of twelfth and thirteenth century witnesses that differ, pairwise, in only one (111 and 823), two (823 and 906) and three (111 and 906) variant units. These differences are very minor, and a very close relationship is implied.

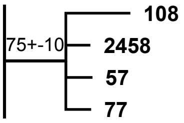


Figure 53: Group 57

Figure 53 shows four eleventh and twelfth century manuscripts. 57, 77 and 2458 have an average pairwise difference of 3.3 variant units while 108 increases this average to 5.5 – still a very close group.

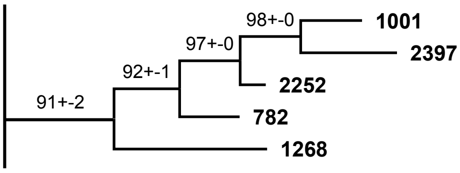


Figure 54: Group 782

Finally, Figure 54 shows a sub-tree with remarkably high confidence values. 782 and 2252 come from the twelfth century, 1001 and 1268 from the thirteenth, and 2397 is dated, according to the *Liste*, to the year 1303.[30] The median-joining network for this group is shown in Figure 55. This network implies that 1268 and 782 are descended from 2252.

---

[30] Aland et al., *Kurzgefaßte Liste der griechischen Handschriften des Neuen Testaments*, 185.

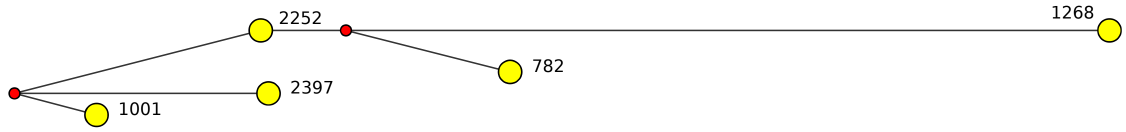


Figure 55: Median-joining network for group 782

There are many similar sub-trees in the results that could warrant further study, highlighting hitherto unknown manuscript relationships.

Chapter 6: Optimising MrBayes

Introduction

Running *MrBayes* on the IGNTP's John 18 data (see Chapter 5) on my laptop required several months for the results to converge and therefore be meaningful. Simply carrying out an identical calculation on the University of Birmingham's BlueBEAR HPC service (which provides a High Performance Computing service to the University's research community) brought the time required down to two weeks, using eight processors.[1] But BlueBEAR has thousands of processors, and so a question presents itself: can this be done any quicker?

This chapter explores several ways of potentially speeding up *MrBayes*' convergence – and looks for some correlation between convergence statistics and the subjective quality of the phylogenetic trees produced at the end of each job. As a dataset we will take Parker's collation of John 2 for the ECM, excluding any fragmentary witnesses that are only extant in less than 25% of variant units.[2] This dataset is considerably smaller than the IGNTP's John 18 data (and therefore faster to process), yet large enough for *MrBayes* to produce quality results.

The Solution Space

It is important to be able to visualise what *MrBayes* is actually doing in order to understand what convergence means. A helpful way to do this is to consider the set of all possible results (i.e. every possible phylogenetic tree having these witnesses at the ends of the branches) in terms of a mountain range. The higher the mountain peak, the better the result will match reality (the higher the mountain the better the view). *MrBayes*' job is to explore the

---

[1] See http://www.birmingham.ac.uk/bear for more details.

[2] Including witnesses that are highly fragmentary can create misleading results – or can hide otherwise useful knowledge to be gained from the consensus trees etc.

mountain range looking for the highest peak. To do this analytically (e.g. by mapping out the entire mountain range) would take for ever and so *MrBayes*' approach is (broadly speaking) to explore the landscape and try to climb upwards.[3] The key to *MrBayes*' success is that it uses multiple chains (see below) which allow it to jump from one mountain to another if the second one is found to be taller.

Convergence

MrBayes was compiled with OpenMPI support allowing it to run in parallel on many of BlueBEAR's processors at once.[4] In 2004, Altekar, Dwarkadas, Huelsenbeck and Ronquist wrote a paper entitled *Parallel Metropolis-coupled Markov chain Monte Carlo for Bayesian phylogenetic inference*. They described their motivation for writing the paper as follows:

> *Bayesian estimation of phylogeny is based on the posterior probability distribution of trees. Currently, the only numerical method that can effectively approximate posterior probabilities of trees is Markov chain Monte Carlo (MCMC). Standard implementations of MCMC can be prone to entrapment in local optima. Metropolis coupled MCMC, a variant of MCMC, allows multiple peaks in the landscape of trees to be more readily explored, but at the cost of increased execution time.*[5]

In other words, Metropolis coupling is required in order to increase the quality of the results for MCMC Bayesian analysis of phylogenetic data – but it is slow. The reason it is

---

[3] The total number of possible trees with 176 leaves is huge. Calculating this number is described in Felsenstein, 'The Number of Evolutionary Trees'. We can simply infer that there are considerably more than 10^{370} possible multifurcating trees with 176 leaves. Consider that the number of atoms in the universe is approximately 10^{80} and the scale of the problem becomes clear.

[4] This is my modified version of *MrBayes*. *MrBayes* only supports ten character states for standard morphology data. I modified *MrBayes* to support sixty-two character states. See Chapter 4 for details. See also Gabriel et al., 'Open MPI: Goals, Concept, and Design of a Next Generation MPI Implementation'.

[5] Altekar et al., 'Parallel Metropolis Coupled Markov Chain Monte Carlo for Bayesian Phylogenetic Inference', 407.

slow is that instead of doing the MCMC analysis once, it is done several times at once in different “chains”. Each of these chains has a different level of “heat” applied – basically making it traverse the solution space in bigger jumps. The first chain is “cold” - just like a simple MCMC analysis – but it can essentially learn from what is happening in the other chains by swapping places with them: “Successfully swapping states allows a chain that is otherwise stuck on one peak in the landscape of trees to explore other peaks. For example, if the cold chain is stuck on a peak in the posterior distribution of trees, swapping states with another (heated) chain allows the cold chain to jump to another peak in a single cycle. As a result, the cold chain can more easily traverse the space of trees.”[6]

The *MrBayes* tutorial describes convergence like this:

> *By default, MrBayes will run two simultaneous, completely independent analyses starting from different random trees (Nruns = 2) ... The idea is to start each run from different randomly chosen trees. In the early phases of the run, the two runs will sample very different trees but when they have reached convergence (when they produce a good sample from the posterior probability distribution), the two tree samples should be very similar.*[7]

We will consider the following three ways of identifying whether or not the analysis has converged: the average standard deviation of split frequencies (ASDSF); the effective sample size (ESS); and the potential scale reduction factor (PSRF).[8] Basically, in an ideal world, when the analysis has converged the ASDSF should be low (preferably < 0.01), the ESS should be high (> 200) and the PSRF should approach one (rounding to 1.000 is desired).

---

[6] Ibid., 409.

[7] See ‘MrBayes Tutorial, v3.2’ (section ‘Setting up the Analysis’).

[8] For definitions of these terms, and a discussion on the best ways to check for MCMC convergence see Chapter 5, and Ronquist and Deans, ‘Bayesian Phylogenetics and Its Influence on Insect Systematics’; Gelman and Rubin, ‘Inference from Iterative Simulation Using Multiple Sequences’.

Convergence

The goal of this optimisation task is to have well-converged analysis in the shortest time frame.

Benchmark

As a benchmark, I configured *MrBayes* as shown in Table 83 and executed it six independent times on the John 2 data, i.e. on a NEXUS file created from Parker's ECM collation of John 2, including all witnesses from the ECM that are extant in at least 25% of variant units in this chapter.[9] See Appendix 3 for the full NEXUS file. This configuration means that two independent runs will be carried out with eight chains each.[10] They will calculate a maximum of 100 million generations but stop early if the ASDSF gets as low as 0.01. These terms will be explained as we go along. The benchmark results are shown in Table 83.

#Runs	#Chains	Target ASDSF	#Generations	Burninfrac
2	8	0.01	100,000,000	0.5

Table 82: MrBayes configuration for the benchmark

Analysis ID	ASDSF	#Generations	ESS	PSRF	MCMC time
920421	0.009995	86470000	14944.03	1.000	16.75
920450	0.009848	58150000	9765.85	1.000	10.5
920451	0.009993	77865000	10329.27	1.000	14.75
920452	0.009997	43040000	7582.82	1.000	7.75
920453	0.009988	37780000	7612.02	1.000	6.5
920454	0.009998	67575000	11970.11	1.000	13

Table 83: Benchmark results

---

[9] This analysis was carried out between May and November 2016, using data from 20th May 2016.

[10] Each node on BlueBEAR has sixteen processors (at time of writing), so each benchmark run will require just a single node – which reduces the queue time waiting for the job to be started.

The first key conclusion of the benchmarking process is that some of the statistics are significantly different between the six benchmarks. While they all converged to less than 0.01 ASDSF, they took between 6.5 and 16.75 hours to do so – and between 37,780,000 and 86,470,000 generations. The jobs that took longer had a larger ESS than those that took less time – which is to be expected. The ASDSF of the six jobs is shown in Figure 57 starting from generation 2,970,000. This graph highlights some interesting features of MC^3. First consider 920450. It is converging at around the same number of generations as 920452 around the middle of the lower axis, but then a significant event happens and it diverges again. It seems reasonable to postulate that a heated chain for one of the two runs had found a better peak, some distance away in the solution space and the cold chain swapped with it – temporarily diverging from the other run. The same seems to have happened for 920454.

Benchmark Trees

All six benchmarks produced very similar trees. For an example tree see Appendix 2. Job 920450 includes one extra sub-tree (when compared to all the other jobs; the pair 33 and 579). The full set of sub-trees from 920450 (as opposed to the rest of the tree which consists of single witnesses on the end of independent branches) is shown in Figure 58.

Branch Probabilities

Before looking at these trees in detail a brief mention of the numbers on the branches is in order. These represent the branch probability – or the confidence that can be assigned to any given branch. It is calculated by taking the second 50% of generations (the idea being that the data has largely converged by then) and considering the tree created by each generation. The trees shown here are consensus trees – which are made by including any sub-tree that occurs in more than half of the trees considered. Thus the minimum branch probability will be

50%, and such a branch would indicate that nearly as many trees did not have this branch as had it. The higher the branch probability, the more confidence can be ascribed to the branch being a genuine feature of the data.

Sub-tree: Family 1 (f^1)

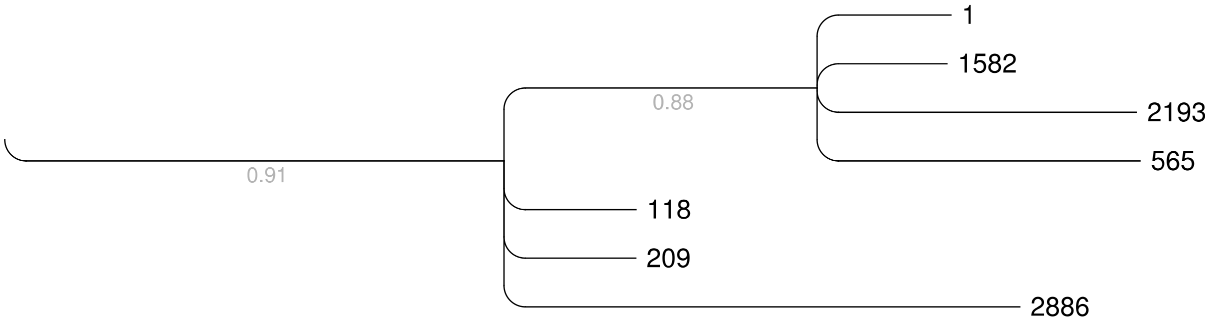


Figure 56: Family 1

Family 1 was discussed in Chapter 5, where *MrBayes* identified a number of f^1 witnesses in a sub-tree in John 18 data. It has identified such a sub-tree again here in John 2, although with subtly different membership than before. Here 1, 118, 2886, 209, 565, 1582 and 2193 are found in a sub-tree, while the other extant (and included) witnesses 884, 1192, 1210, 1278, 2372 and 2713 are found as singleton branches.

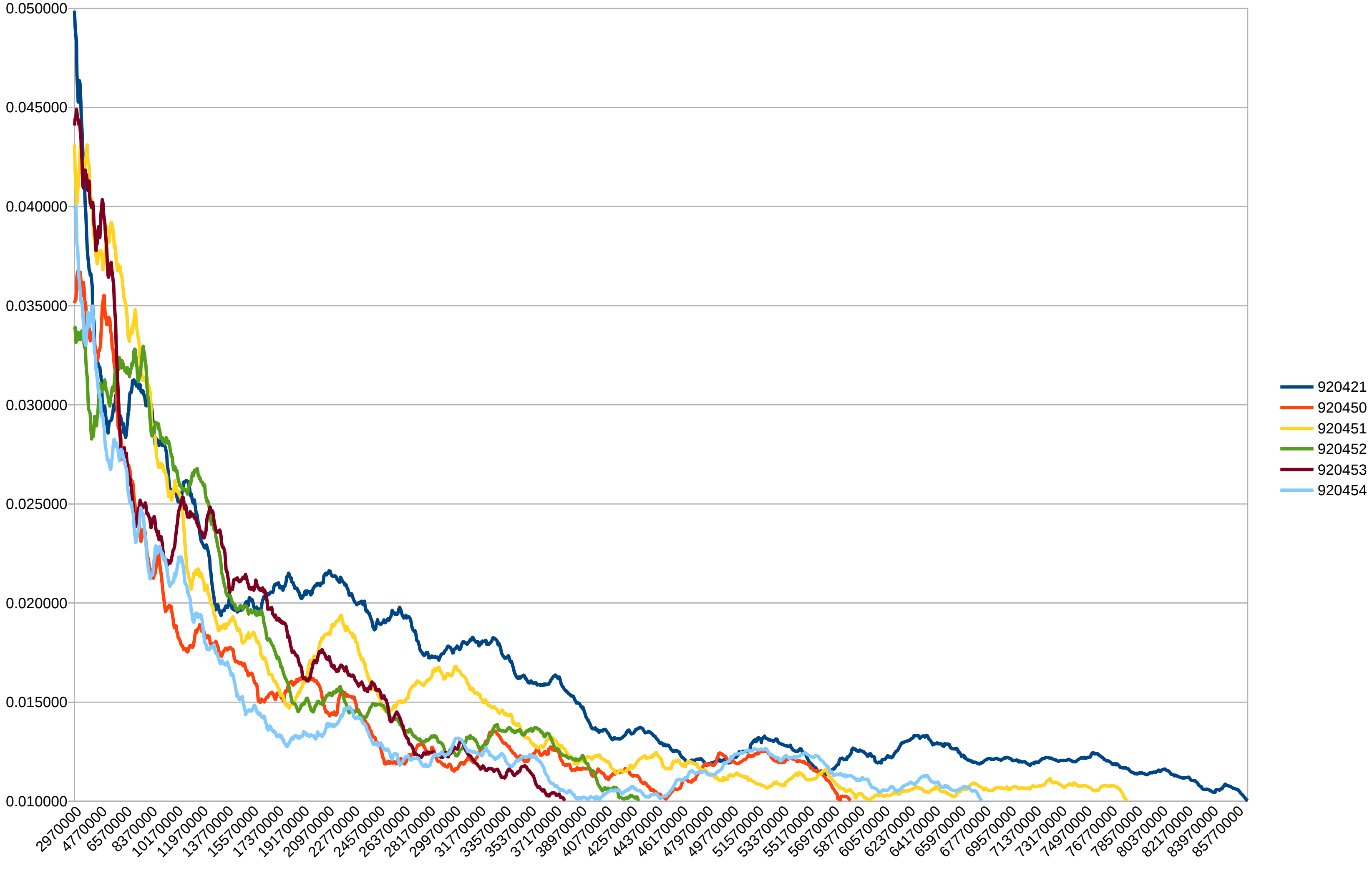
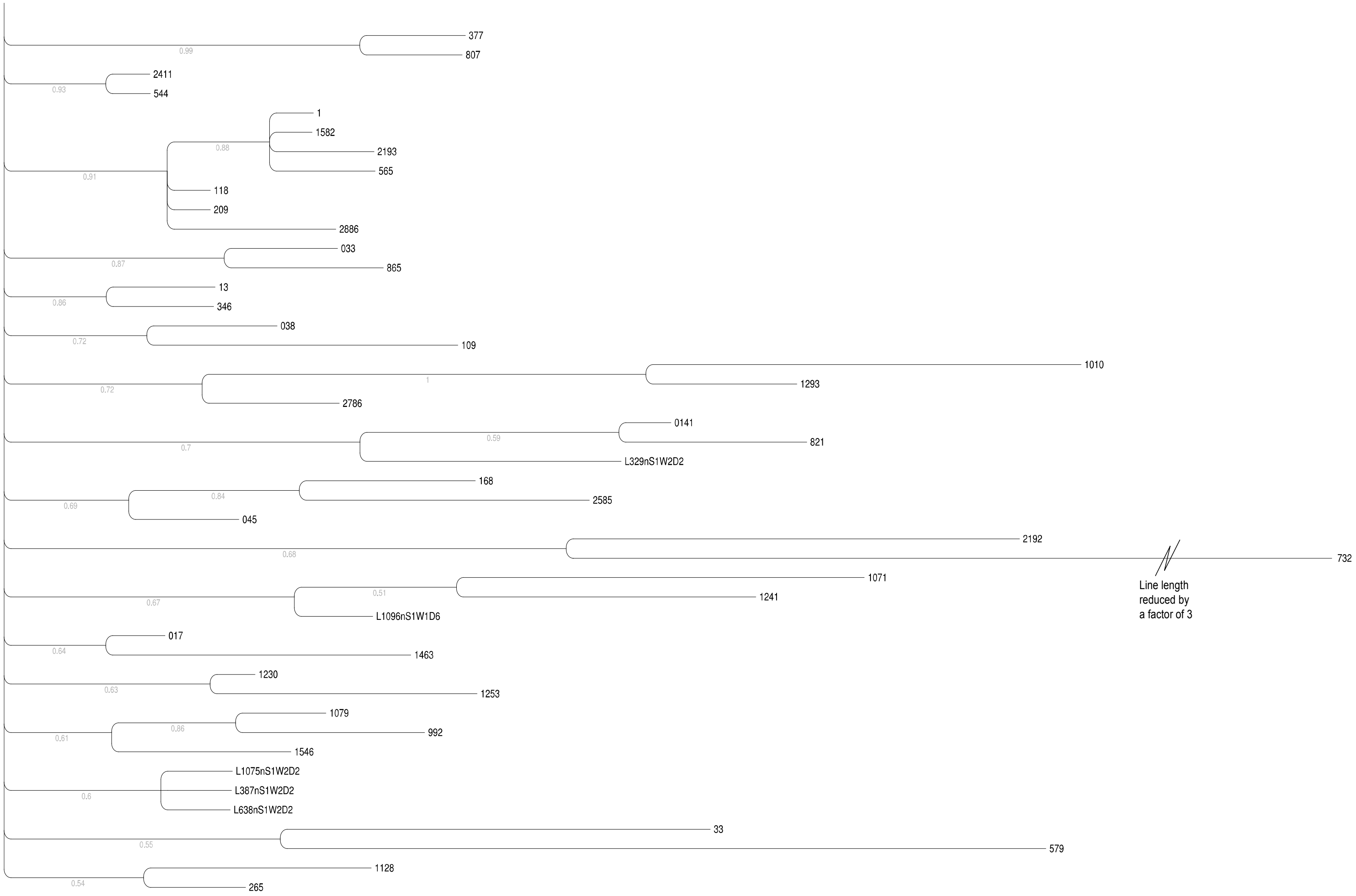


Figure 57: ASDSF of the benchmarks, by generation


377
0.99
807
2411
0.93
544
1
1582
0.88
2193
565
0.91
118
209
2886
033
0.87
865
13
0.86
346
038
0.72
109
1010
1
0.72
1293
2786
0141
0.59
0.7
821
L329nS1W2D2
168
0.84
0.69
2585
045
2192
0.68
732
Line length reduced by a factor of 3
1071
0.51
0.67
1241
L1096nS1W1D6
017
0.64
1463
1230
0.63
1253
1079
0.86
0.61
992
1546
L1075nS1W2D2
0.6
L387nS1W2D2
L638nS1W2D2
33
0.55
579
1128
0.54
265

Sub-tree: Family 13 (f^{13})

Again in Chapter 5, f^{13} was discussed – where Perrin's list of members (13, 69, 124, 346, 543, 788, 826, 828, 983, and 1689) was largely identified by *MrBayes*. Here there is no such identifiable tree. Only 13 and 346 are grouped together in Figure 58. 1689 is not included in the ECM, but all the other witnesses listed above exist as singleton branches. Interestingly, jobs 920452, 920453 and 920454 also include a small sub-tree containing the pair 69 and 828, although separately from 13 and 346.

Other Notable Features

Some features among the remaining simple sub-trees are worthy of note. For example the pair 0141 and 821 (see Chapter 5) occurs again – this time grouped with a lectionary witness: L329-S1W2D2 (labelled as L329nS1W2D2 in the tree as *MrBayes* does not allow hyphens in a taxon name) – an eleventh century manuscript currently residing in the British Library.

The only multi-lectionary sub-tree is a trio of witnesses: L1075-S1W2D2, L387-S1W2D2 and L638-S1W2D2 – although the sub-tree only has 0.6 (i.e. 60%) probability indicating that caution should be exercised when talking about the group. These three eleventh century lectionaries currently reside in the Great Lavra (Mt. Athos), the National Library of Greece (Athens) and Dionysiou (Mt. Athos) respectively.

An intriguing pair is also found, although there is more distance between them than there is from the trunk. 017 is a ninth century manuscript in the National Library in Paris, while 1463 dates from the thirteenth century and resides in the Great Lavra (Mt. Athos).

Conclusion

Even though the convergence data vary significantly, the resulting phylogenetic trees from all six benchmark jobs are remarkably similar. We will, therefore, use these main sub-trees as a way of measuring the relative quality of the results for the rest of this chapter. The goal is *to get equivalent (or better) quality results in significantly less time*.

Number of Processors

Introduction

There are two ways in which *MrBayes* can make use of multiple processors – multiple "runs" and multiple "chains". We will explore increasing and decreasing both settings.

Chains

In addition to the benchmark jobs (two runs, eight chains) I scheduled three jobs each with two runs and four, twelve, sixteen, twenty, twenty-four, thirty-two, forty-eight and sixty-four chains. The results were a mixture and did not show any strong pattern. Figure 59 shows the time required for the MCMC analysis to converge (ASDSF < 0.01) with the different numbers of chains. Figure 60 shows the number of generations required in each case.

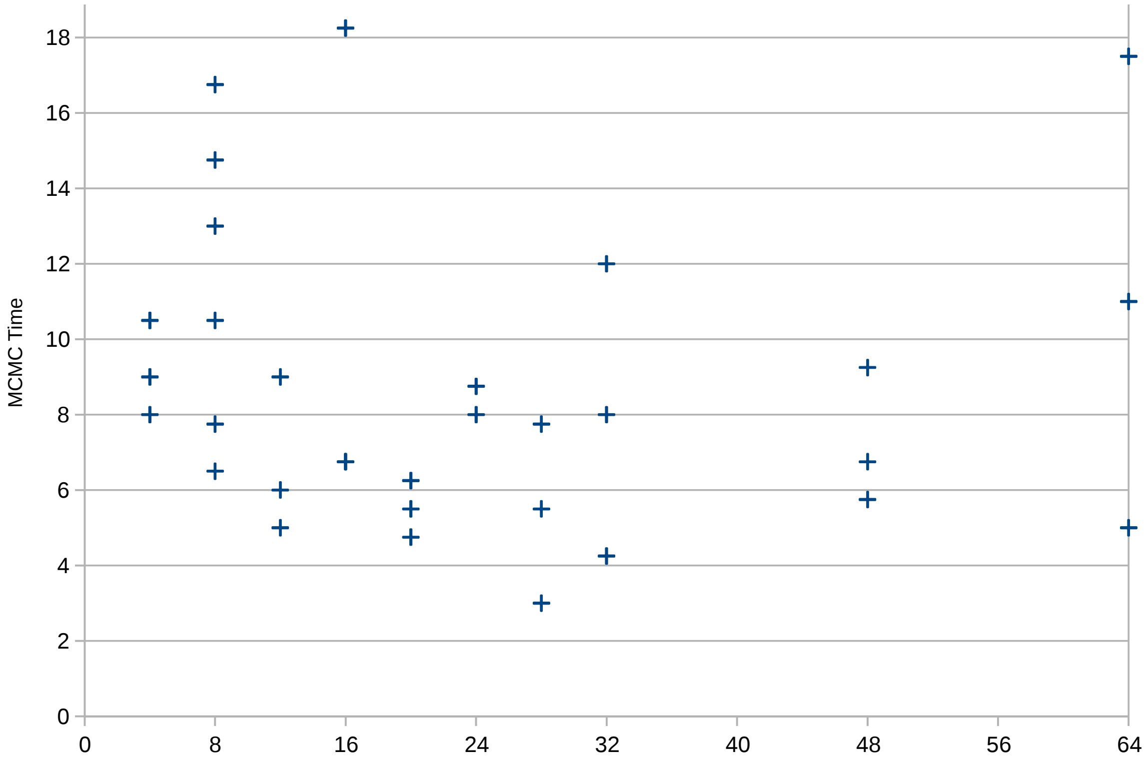


Figure 59: Time for MCMC analysis to converge with different numbers of chains

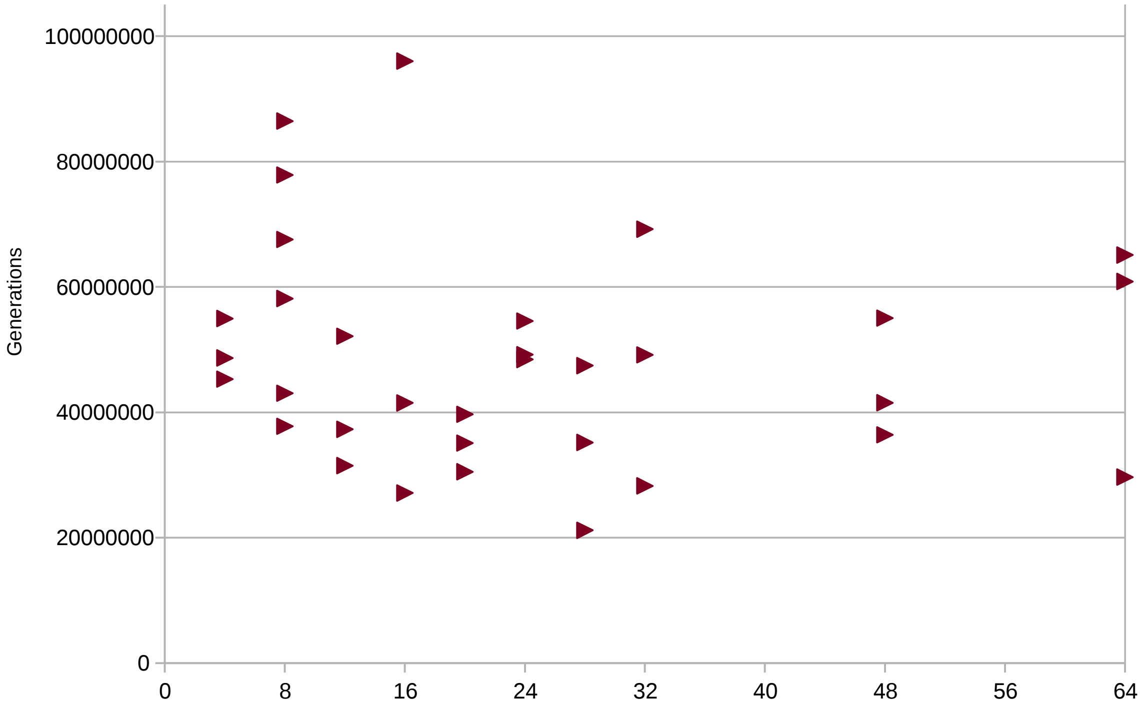


Figure 60: Generations for MCMC analysis to converge with different numbers of chains

Number of Processors

There are a number of peculiarities arising from these graphs. First there are some outliers – notably one job at sixteen chains took over eighteen hours to converge and required 96,035,000 generations. One job at sixty-four chains took nearly as long, but significantly fewer generations (65,100,000).[11] This highlights two important facts about this kind of analysis – first any given job could choose a good route through the solution space or a bad one – and a good one might converge quickly to a good answer while a bad one might take a very long time.[12] This is a feature of this kind of non-analytical, sampling-based method. The second feature is that there does not seem to be a simple way of translating the number of generations required into the time taken. For some reason it sometimes takes a long time to do a few generations. This is made plain by the other two jobs with sixteen chains. One took 6.75 hours to do 27,130,000 generations while the other took 6.25 hours to do 39,685,000. Whether this is a feature of *MrBayes*, or BlueBEAR or some other factor is not clear at this time.

The second peculiarity is that there seems to be a general downward trend in both graphs up to twenty chains, or possibly twenty-eight chains, but then it goes up again. Twenty and twenty-eight chains have about the same average time and generation results, but the standard deviation is significantly greater with twenty-eight chains (thus it also has the quickest individual job). Given the discussion in the previous paragraph it would be unwise to claim any kind of pattern from this data. It may be that this dataset can converge easily with relatively few chains, and so the extra chains do not add much to the Metropolis coupling. Or it may be that the apparent randomness of the Bayesian method outweighs the effects gained by adding more and more chains.

---

[11] These two facts come from the underlying data, as individual jobs cannot be identified from the graphs.

[12] Of course it does not "choose" a "route" at all – but attempts to converge on peaks while the Metropolis coupling helps it to jump to other peaks. But the end result can be thought of as a route through the solution space.

While most of these jobs produced the expected trees, two produced trees with an additional interesting feature: 920944 and 920607, both with twenty-eight chains. Figure 61 shows their sub-tree representing f^{13} – which is significantly more exciting to a text critic than that found in the benchmark results.

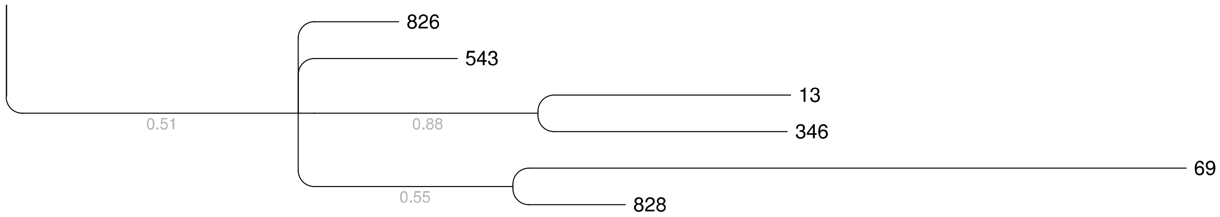


Figure 61: Family 13 in 920944

Here we find 13, 69, 346, 543, 826 and 828 grouped together. That leaves out only 124, 788 and 983 of the available witnesses. Sadly, this sub-tree sits on a branch with only 51% probability – and so no great claims can be made. This is either a hard-to-find result reflecting real-world relationships (and therefore exciting) or a statistical anomaly seeming to confirm a flawed result we were hoping to find (and therefore very shaky ground). Interestingly, 920607 was both the fastest result (three hours) and also contained this sub-tree. If this was the only job we had run then we might have claimed a great result, which highlights the importance of repeating such analysis multiple times.

Runs

The benchmark jobs used two runs and eight chains each. To test the effect of changing the number of runs I ran three jobs each with one, three, four, six and eight runs. In a similar vein to the Chains tests above, the results were inconclusive. Figure 62 shows the time required for the MCMC analysis to converge (ASDSF < 0.01) with the different numbers of runs. Figure 63 shows the number of generations required in each case. Note that, by definition, a single run cannot converge – so the first point on each graph represents the full

Number of Processors

100M generations.

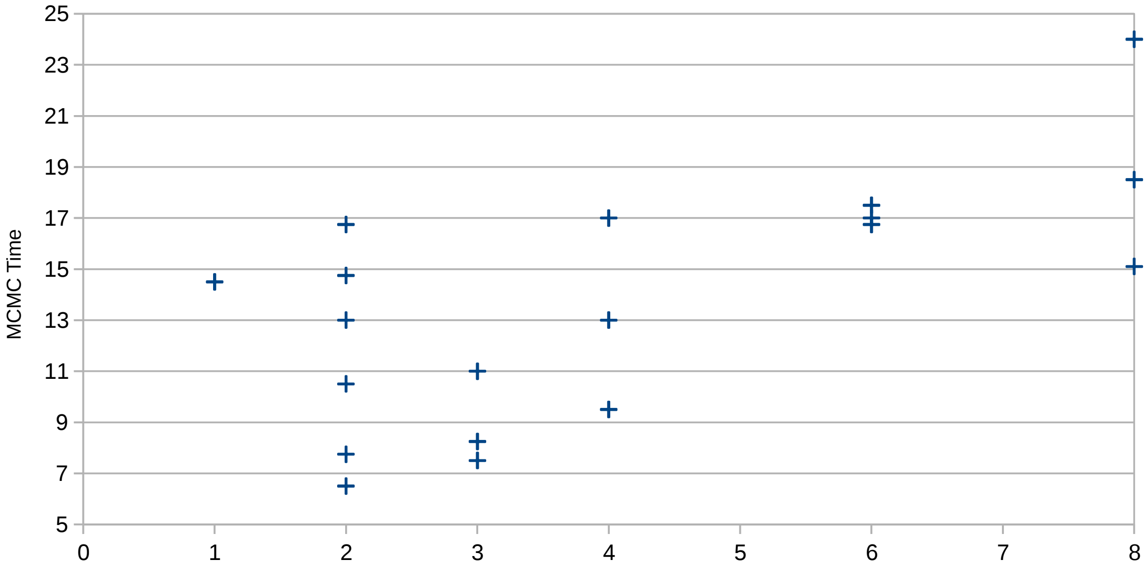


Figure 62: Time for MCMC analysis to converge with different numbers of runs

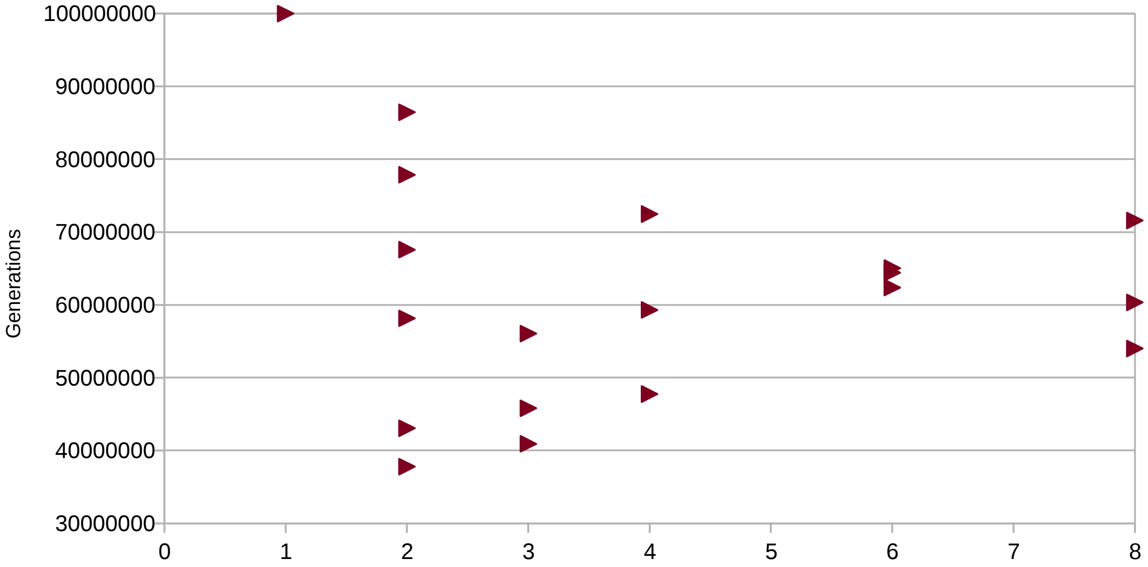


Figure 63: Generations for MCMC analysis to converge with different numbers of runs

The jobs with eight runs completed the MCMC analysis fine, but failed to create a consensus tree. It seems that BlueBEAR killed the process in each case - due to memory exhaustion on the first compute node. As such we can consider how quickly they converged, but they are currently unusable for the real end-purpose. The jobs with six runs were not killed in this way, but they did not complete their analysis of the trees in time and BlueBEAR killed them anyway at forty-eight hours.

While there may be some doubt about a trend in Figure 63 (generations required to converge), Figure 62 seems conclusive: using more runs requires more time. This is not unexpected. It is likely that more independent runs would have a greater chance of finding the peaks in the solution space but, crucially, they would all have to converge on the same peak, which is much less likely to happen quickly. So more runs makes it harder to converge and *MrBayes* thus takes longer to do so. The jobs with one, two, three and four runs were able to complete and create consensus trees. The resulting trees were similar to the benchmarks. Therefore, it seems that increasing the number of runs above two is unnecessary.

Configuration Options

Burn-in

One key aspect of making a good consensus tree is making sure that early generations are not included. In these early generations the standard deviation will be changing significantly, as the two runs explore the solution space. If a consensus tree is created using these early generations then it is unlikely to contain many interesting sub-trees. Matzke's *MrBayes Lab* explains how to interpret the ".p" files created by *MrBayes* to identify a suitable burn-in setting. In this case the standard deviation becomes stationary early in the run and the "50%" burn-in setting that was used in all our test jobs was in no danger of including unstable results. As a result there is little to be gained by experimenting with different burn-in values at this time.

Temperature

It is possible to change the temperature settings applied to the heated chains: “The higher the temperature, the more likely the heated chains are to move between isolated peaks in the posterior distribution. However, excessive heating may lead to very low acceptance rates for swaps between different chains.”[13] The *MrBayes Tutorial* explains the value of chains swapping states and that insufficient swaps are an indication that the temperature is too high.[14] The more chains that are used, the smaller the change in temperature there is between the chains – which explains why running the analysis with more chains can speed up convergence, although it may not do so every time (see the Chains section above). Consider Figure 64 and Figure 65 - the heat applied to a run with eight chains compared with twenty chains.[15]

```
ID -- Heat
-----------
 1 -- 1.00  (cold chain)
 2 -- 0.91
 3 -- 0.83
 4 -- 0.77
 5 -- 0.71
 6 -- 0.67
 7 -- 0.62
 8 -- 0.59
```

Figure 64: Temperatures for 8 chains with T=0.1

```
ID -- Heat
-----------
 1 -- 1.00  (cold chain)
 2 -- 0.91
 3 -- 0.83
 4 -- 0.77
 5 -- 0.71
 6 -- 0.67
 7 -- 0.62
 8 -- 0.59
 9 -- 0.56
10 -- 0.53
11 -- 0.50
12 -- 0.48
13 -- 0.45
14 -- 0.43
15 -- 0.42
16 -- 0.40
17 -- 0.38
18 -- 0.37
19 -- 0.36
20 -- 0.34
```

Figure 65: Temperatures for 20 chains with T=0.1

---

[13] *MrBayes* internal help (run “help mcmc” from within *MrBayes* itself). See also Altekar et al., ‘Parallel Metropolis Coupled Markov Chain Monte Carlo for Bayesian Phylogenetic Inference’.

[14] ‘MrBayes Tutorial, v3.2’, para. 4 in section “Running the Analysis”.

[15] The heat is calculated as follows: Heat = 1 / (1 + T * (ID – 1)). Where T = 0.10 is the temperature and ID is the chain number.

It seems that modifying the temperature (from the default T=0.1) might be useful. First we will run two runs with eight chains at T=0.1 (the benchmark jobs), 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, and 0.5. Intuitively, it seems likely that raising the temperature will produce the best result – as it will produce hotter chains that traverse the solution space faster. The results are shown in Figure 66 - and it appears that (at least for this dataset) raising the temperature has made things worse.

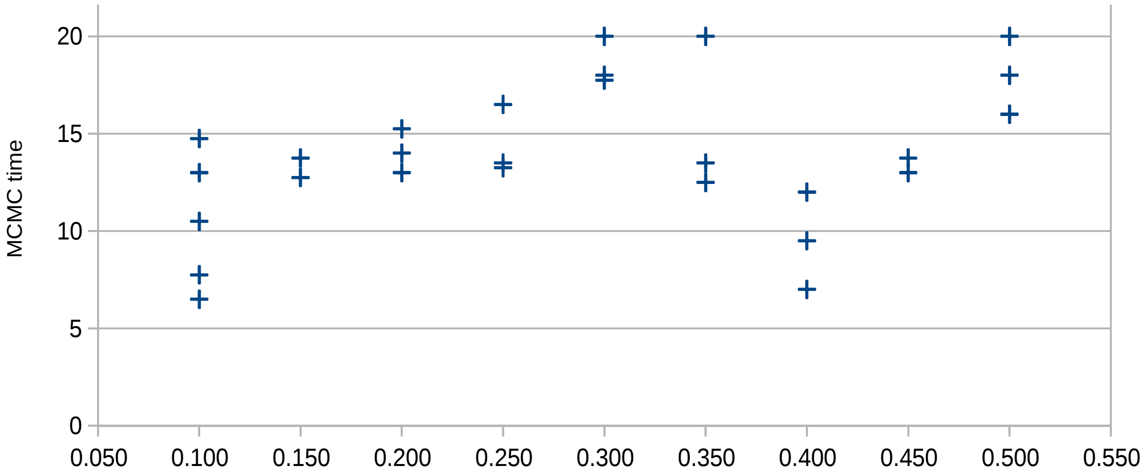


Figure 66: MCMC time against Temperature (hotter)

Now we will try *lowering* the temperature – and it turns out that this is the key for this dataset. Figure 67 shows the MCMC time taken with T ≤ 0.1 – and it would seem, the colder the better:

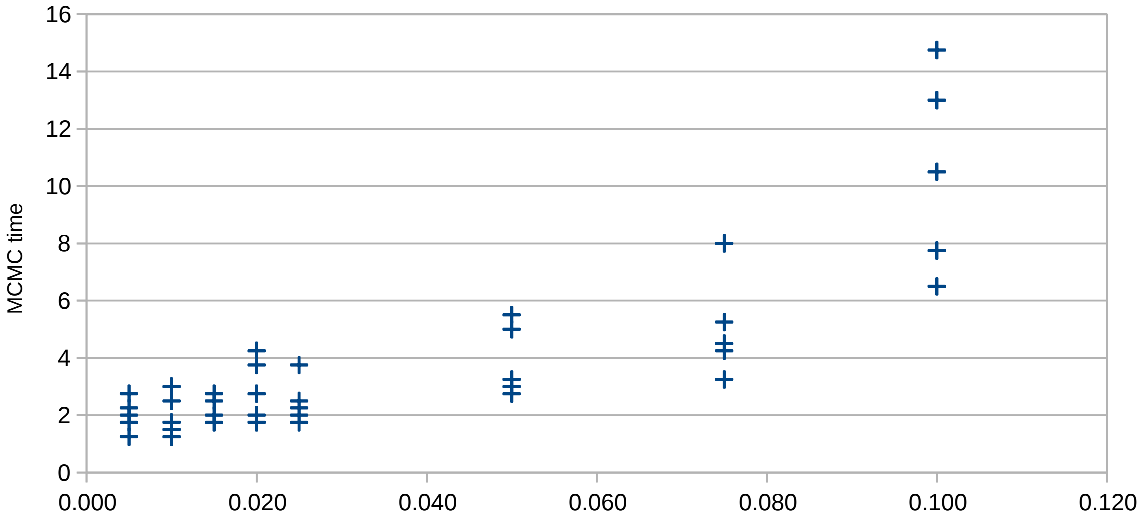


Figure 67: MCMC time against Temperature (colder)

It is clear from this data that lowering the temperature results in considerably faster convergence. All jobs with T<0.02 converged in less than four hours, and the average time for T=0.005 was just two hours (the quickest of which took seventy-six minutes).[16] So lowering the temperature is highly effective in reducing the time required to produce an equally good result. The reason must be that an increase in the frequency of chain swaps results in faster convergence for this dataset. Figure 69 shows the chain temperatures with T=0.025, and for comparison Figure 68 shows them with T=0.1 (duplicating Figure 64 above).

```
ID -- Heat
-----------
 1 -- 1.00  (cold chain)
 2 -- 0.98
 3 -- 0.95
 4 -- 0.93
 5 -- 0.91
 6 -- 0.89
 7 -- 0.87
 8 -- 0.85
```

Figure 69: Temperatures for 8 chains with T=0.025

```
ID -- Heat
-----------
 1 -- 1.00  (cold chain)
 2 -- 0.91
 3 -- 0.83
 4 -- 0.77
 5 -- 0.71
 6 -- 0.67
 7 -- 0.62
 8 -- 0.59
```

Figure 68: Temperatures for 8 chains with T=0.1

If we reduce the temperature significantly further, to 0.0001, we find that the results get somewhat worse again. I ran three jobs with T=0.0001 with an average convergence time

---

[16] The fastest job of all was actually ID 946853, with T=0.01, which finished in 1 hour 11 mins.

of just under three hours.[17] So it seems that a little heat is good, for traversing valleys, but too much and the chains refuse to swap. Even the default of T=0.1 seems considerably hotter than optimal, and T=0.005 has given the quickest convergence in this test. Note that T=0.01's results have an average convergence time of only about ten minutes higher than that of T=0.005. We could continue interpolating these values, but we have found a significant improvement and now must deal with the problem of too little heat.

Many Chains and Lower Temperatures

There is a methodological worry with lowering the temperature so much: what happens if the chains are no longer able to traverse valleys? They may climb to their local peak and stay there. It would then be possible to find that the two runs converged on a peak that was far from the highest in the space, but they just happened both to focus on it. This problem does not appear to have manifested in this analysis, considering the six test jobs performed. This risk can, however, be mitigated as follows.

Recall that the six benchmark jobs (two runs, eight chains, T=0.1 (default)) converged in between 6.5 hours (actually six hours twenty-four minutes and thirty-three seconds) and 16.75 hours (actually sixteen hours, forty-seven minutes and forty-nine seconds). Changing the number of runs could not improve on this time. Changing the number of chains did improve matters, with the best single result being with twenty-eight chains where it converged in three hours, six minutes and fifty-four seconds. The faster convergence was found by changing the temperature, where one job with T=0.01 took only one hour, ten minutes and fifty-seven seconds. A good solution would be to use a low temperature (say 0.01 – a tenth of the default) but more chains – so that there is still a suitably hot chain (e.g. chain thirty-two

---

[17] Note that *MrBayes* reports the chain temperatures as all being 1 when T=0.0001, evidently rounding T to 0 at this value.

where the heat is 0.76) that is able to traverse the valleys. I ran six jobs with these settings, the results of which are shown in Table 84.[18] These include the first sub-one hour results we have seen, and the average is an order of magnitude improvement on the average time from the benchmarks (which was over eleven hours). The average time is fifty-seven minutes, with standard deviation (SD) twenty-one. The consensus trees created are of equal quality to those from the benchmarks.

Job ID	**MCMC time**	**Avg. ESS**	**PSRF**
947022	1 hour 29 mins 28 seconds	10867.94	1.000
947023	45 mins 31 seconds	5757.06	1.000
947024	46 mins 12 seconds	6604.11	1.000
947698	38 mins 20 seconds	3976.48	1.000
947699	42 mins 29 seconds	6048.18	1.000
947700	1 hour 18 mins 14 seconds	10193.42	1.000

Table 84: Results of 32 chains with T=0.01

This method can be pursued further, although with diminishing returns. I executed three runs with T=0.005 and sixty-four chains (i.e. 128 processors, or eight nodes), the results of which are shown in Table 85. The average run time was thirty-eight minutes (SD seventeen). This is a small improvement in time, speaking in absolutes, but requires twice as many BlueBEAR nodes.

Job ID	**MCMC time**	**Avg. ESS**	**PSRF**
948915	55 mins 35 seconds	6013.83	1.000
948916	35 mins 58 seconds	3100.96	1.000
948917	22 mins 58 seconds	3476.74	1.000

Table 85: Results of 64 chains with T=0.005

---

[18] I.e. with settings: temperature=0.01, 32 chains, 2 runs, stop when ASDSF <= 0.01.

Conclusion

Table 86 shows the compute resources required for the benchmark, and the two fastest experiments. The “sweet spot” identified for this dataset is T=0.01 with thirty-two chains per run, completing the task in the fewest compute minutes. The best improvement (to T=0.005) yields an eighteen-fold increase in speed, considering the averages.

Temperature	Chains	Avg (SD) minutes	Compute minutes
0.1	8 (per run)	690 (243)	8 x 2 x 690 = 11,040
0.01	32 (per run)	57 (21)	32 x 2 x 57 = 3,616
0.005	64 (per run)	38 (17)	64 x 2 x 38 = 4,907

Table 86: Compute resource required

We have shown that, for a given dataset, the time required for *MrBayes*’ analysis can be reduced by at least an order of magnitude. Exactly which settings to use will differ for different data sets – but there will be a “sweet spot” which is likely to be found using low temperatures and many chains. The final decision on which settings to use will include a trade-off between speed of completion and compute resource required.

PART THREE

ANALYSING THE CBGM

Chapter 7: The CBGM: Connectivity

Introduction

The CBGM's textual flow diagrams all use a specified maximum connectivity value.[1] For example a diagram with connectivity ten would only allow witnesses to be linked to ancestors of rank ten or less, where the rank comes from the genealogical coherence data for the witness. Mink observes that the default options for connectivity are: five (Low); ten (Average); and unlimited ("Absolute", although in reality 499).[2] Users can choose their own value, but when presented with default options, most people will select from those. The premise of this chapter is that using the rank for connectivity is insufficient, and it would be better to include a minimum coherence percentage. Also, using a single connectivity (rank) for all variant units is insufficient. As such, all textual flow diagrams in this section will show both the rank and genealogical coherence percentage on the link between witnesses.

Why does this matter? Consider, for example, Wasserman's paper from 2015: 'The Coherence Based Genealogical Method as a Tool for Explaining Textual Changes in the Greek New Testament.' In this paper the CBGM's ability (in its textual flow diagrams) to show multiple emergence of a variant reading is crucial to Wasserman's arguments concerning the development of the text. If textual flow diagrams are showing multiple emergence where there is none, or hiding it where it should be shown, then arguments such as Wasserman's could fail.

A typical implication of a textual flow diagram showing multiple emergence for a connective variant, is that the scholar must re-visit the relevant local stemma. If multiple

---

[1] See Mink, 'Guide to "Genealogical Queries"' section 3.b for a brief discussion of connectivity and how to select it.

[2] See Mink, 'Contamination, Coherence, and Coincidence in Textual Transmission', 175.

emergence is deemed unlikely/impossible (given the philological analysis of the variant readings) then the local stemma must be wrong, commonly with the result that the only remaining choice is to invert a relationship, potentially changing the critical text at this place.

Using the rank of a potential ancestor for deciding whether or not two witnesses may be linked in a textual flow diagram is highly problematic: it leaves one a hostage to fortune, and at the mercy of the vagaries of history.[3] The rank would be radically different for a witness found in an archaeological cache of similar witnesses compared to exactly the same witness being found alone (with the others all destroyed). This is the same sandy foundation that underlies studies based on "singular readings", which can only know about readings in extant manuscripts (or references to manuscripts from secondary sources). A reading is not singular any more if another witness to it is found. Also consider the related question of the inclusion or exclusion of witnesses: the ECM of John has excluded (after careful study) the majority of witnesses (consider that John 18 has over 1,500 extant witnesses yet the ECM of John editorial team selected 229) - so is a reading singular if it only appears once in the selected witnesses?[4] Similarly, if all the witnesses were included then most witnesses would gain a number of very similar relatives and many ranks would change (but not the related coherence percentages). Indeed the witnesses were excluded in order to avoid such near-duplicates as they add little to the edition.

Take a hypothetical example where a witness has many highly similar potential ancestors. Its fiftieth ranked potential ancestor could have a coherence of 95%. Yet a witness

---

[3] Mink highlights this problem in Mink, 'The Coherence-Based Genealogical Method (CBGM) – Introductory Presentation, 1.0', 270, 293 etc. He cautions that high ranking potential ancestors with lower coherence "are in the same range of agreement percentages where coincidental agreement may be considered." Yet the CBGM only uses rank for textual flow diagrams, and so the reader is presented with a diagram that hides this crucial information. See also Mink, 'Problems of a Highly Contaminated Tradition', 32.

[4] See Chapter 5 for the John 18 data. The ECM's 229 manuscripts are 27 papyri, 55 majuscules, 109 minuscules and 38 lectionaries, including supplements. See Parker et al., 'The Selection of Greek Manuscripts to Be Included in the International Greek New Testament Project's Edition of John in the Editio Critica Maior', 328 for a full list of included manuscripts.

with few similar ancestors could have a third ranked potential ancestor with coherence 75%. Now suppose that forty-nine of the fifty close relatives of the first witness had been destroyed in a fire 500 years ago... The rank is too easily affected by the destruction or loss of manuscripts over the centuries, or the relative popularity (and thus the frequency of copying) of a witness. The coherence percentage is not changed by these factors.

This chapter describes a simple experiment aimed at proving the hypothesis that the coherence percentage should be used in preference to (or in combination with) the rank.

Data

The data used in this section comes from Parker's work on the ECM of John's Gospel.[5] This is an intermediate stage in the process, where the majority of local stemmata have been defined, but are yet to be confirmed by reviewing the textual flow in each variant unit. This makes it an excellent set of data for analysing the impact of the methodology used for creating textual flow diagrams, for a change to the methodology could yield a change to the confirmation process, and thus a change in the ECM text of John.

Examples

John 2:10/12-20

The local stemma for John 2:10/12-20 is shown in Table 87. The interesting question, as far as this chapter is concerned, is whether reading *d* is a case of multiple emergence. Again, as is very often the case in the New Testament, the variant reading here is a simple one-word change; the omission of *τον*. Consider the textual flow diagrams in Figure 70. If the connectivity is set to less than forty-three then the left-hand diagram is produced – implying

---

[5] This analysis was performed in November 2017 using my CBGM tag 1.10, with data from October 31 2017.

multiple emergence. If it is set to forty-three or higher then we have the right-hand diagram. In other words, multiple emergence is probable unless 544 is the ancestor (considering textual flow) of 1797. A connectivity setting of forty (used here) is high, and thus multiple emergence is strongly implied by the textual flow. This seems entirely appropriate, as the addition or omission of *τον* after *πρωτον* could easily happen multiple times independently.

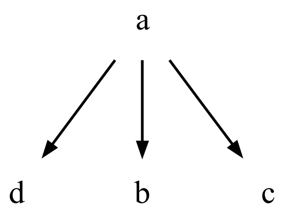
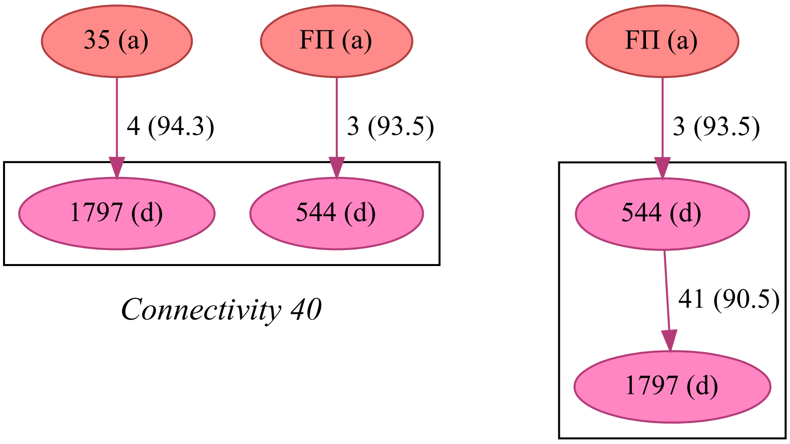
<table>
<tr><td colspan="3">
</td></tr>
<tr><td>a</td><td>πρωτον τον καλον οινον τιθησιν</td><td>FΠ MT F1 F13 𝔓66 01 02 03 07 011 019 022 028 032S 033 037 038 044 045 063 083 0141 0211 0233 18 33 35 109 157 168 173 213 226 295 333 357 377 397 430 579 597 732 792 799 807 821 841 865 992 994 1009 1010 1029 1071 1093 1230 1241 1242 1253 1293 1319 1320 1321 1344 1424 1463 1546 1561 1571 1654 1788 2106 2192 2223 2411 2561 2575 2615 2680 2718 2766 2768 2786</td></tr>
<tr><td>b</td><td>πρωτον τιθησιν τον καλον οινον</td><td>382</td></tr>
<tr><td>c</td><td>τον καλον οινον πρωτον τιθησιν</td><td>𝔓75 892 2585</td></tr>
<tr><td>d</td><td>πρωτον καλον οινον τιθησιν</td><td>544 1797</td></tr>
<tr><td>lac</td><td></td><td>. . .</td></tr>
</table>

Table 87: Local stemma for John 2:10/12-20



Connectivity 40

Absolute Connectivity

Figure 70: Textual flow diagrams of John 2:10/12-20 (reading d)

The difficult question is "what is an appropriate connectivity setting to use?" Normally, in the CBGM, a connectivity of ten or twenty is used – and so forty would be highly unusual. But notice the coherence values in brackets (which are not shown in conventional textual flow diagrams).[6] If 90% coherence is acceptable then a textual flow diagram based on connectivity="90% coherence" (instead of the rank) would show perfect coherence. Should we insist on 95%? Wasserman points out that "it is up to the critic to adjust in each variation-unit the level of [connectivity]".[7] But how to choose? In the simple case of looking at these partial textual flow diagrams, the problem is masked, as they are small and simple. To see why, consider the full textual flow diagram for this variant unit (see Figure 71). When viewing a full textual flow diagram (or indeed a traditional one showing a single reading, but with a hundred witnesses) it is not always possible to determine whether a relationship is admissible. We have considered 544 → 1797 (both thirteenth century, rank forty-one, 90.5%). But on the same diagram we find, for example:

1. A (initial text) → 01 (fourth century) both reading *a*, rank four, 88.3%;

---

[6] The tables are accessible in "Genealogical Queries", and in the new CBGM interface they are easier to find. The key point here, however, is that the data is not shown in the diagram itself.

[7] Wasserman, 'The Coherence Based Genealogical Method as a Tool for Explaining Textual Changes in the Greek New Testament', 209.

Examples

2. 35 (eleventh century) → 173 (twelfth century) both reading *a*, rank five, 89.4%; and
3. 18 (fourteenth century) → 2718 (twelfth century) both reading *a*, rank five, 88.4%. This is an interesting case of a later manuscript (18) presumably preserving an older text, as it has a twelfth century descendant (2718).

Surely these examples are not inherently more trustworthy than 544 → 1797? They are examples of considerably less close relationships (considering coherence) even though they have ranks that would mean they will be shown as parent/child on any standard CBGM textual flow diagram. Rank forty-one might simply be dismissed out of hand, yet rank four or five could well be accepted without further investigation or question. Now we might trust that in each situation where a low rank (i.e. a large number) is shown, the ECM's editors would consult the genealogical coherence tables and examine the percentage coherence values. But how would they know to challenge the apparently strong relationship shown by rank four? Why would a user explore further into the data tables underlying the diagram? The real strength of a diagram is that it conveys a large amount of information ready to be quickly (and implicitly) interpreted. Why would anyone further investigate a relationship with a low and apparently acceptable rank?

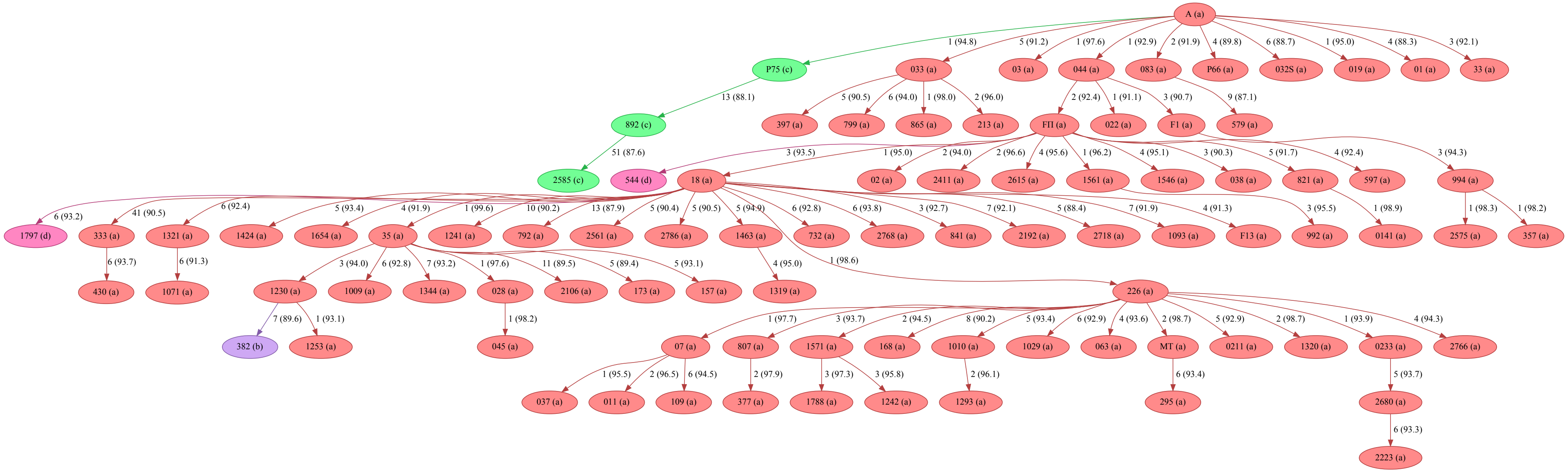
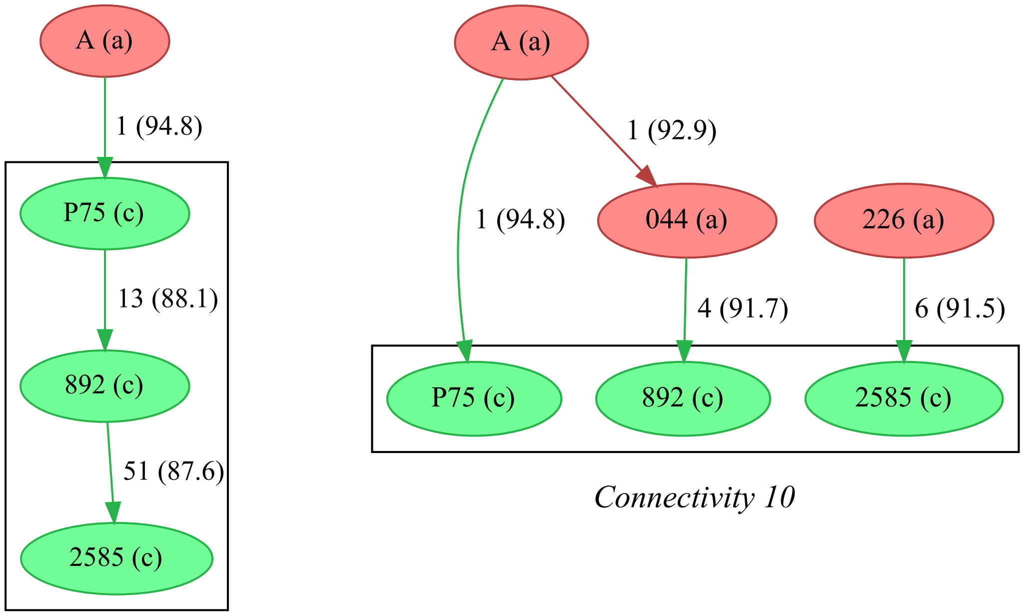


Figure 71: Full textual flow for John 2:10/12-20 with absolute connectivity

Note: This diagram has been edited to fit on the page. Normally witnesses of the same generation (e.g. the immediate descendants of 18) are shown on the same row. But here 226 has been moved lower so that the diagram is less wide.

Examples

Consider, too, reading *c* here (see Figure 72). What connectivity setting should be chosen in this case? The reading is a change in word order, and so not highly connective. There is no problem with using rank here, as a diagram with rank ten shows three-fold multiple emergence. But consider the relationships shown with absolute connectivity. First 𝔓75 → 892 (rank thirteen, 88.1%) has a modestly high rank, yet 892 → 2585 (rank fifty-one, 87.6%) has low rank – but a similar coherence to the first relationship.



Absolute connectivity

Connectivity 10

Figure 72: Textual flow diagrams for John 2:10/12-20 (reading c)

Thus these two relationships have roughly equivalent strength considering coherence percentages, yet very different strength considering rank. In other words, rank alone is insufficient to determine the strength of a relationship. This is not a new statement, and will not surprise any CBGM-proficient scholar. The focus here is the information conveyed implicitly in the textual flow diagrams, where even such a scholar may miss important facts.

John 7:1/24

Here we find a genuine connective variant. This variant is part of the phrase *Καὶ μετὰ ταῦτα περιεπάτει ὁ Ἰησοῦς ἐν τῇ Γαλιλαίᾳ· οὐ γὰρ ἤθελεν ἐν τῇ Ἰουδαίᾳ περιπατεῖν, ὅτι ἐζήτουν αὐτὸν οἱ Ἰουδαῖοι ἀποκτεῖναι*. It is hard to imagine that reading *b*, ειχεν εξουσιαν (*he did not have the authority to go to Judea*) could have emerged multiple times independently from ηθελεν (*he did not wish to go to Judea*). John 7:1 has no synoptic parallel from which such a variant could be harmonised. Perhaps John 19:11 could be the source of the wording, although the context is hardly similar: *ἀπεκρίθη [αὐτῷ] Ἰησοῦς· οὐκ εἶχες ἐξουσίαν κατ' ἐμοῦ οὐδεμίαν εἰ μὴ ἦν δεδομένον σοι ἄνωθεν· διὰ τοῦτο ὁ παραδούς μέ σοι μείζονα ἁμαρτίαν ἔχει.* (NA28). The Greek manuscript support for reading *b* is interesting. Michael Clark has shown that 249, 333, 430 and 869 are all Nicetas catenae and were copied in the eleventh to thirteenth centuries - and thus are closely related and share a recent common ancestor.[8] 032, however, is a fourth/fifth century majuscule, continuous text manuscript. There is also some Old Latin support for this reading, with the group VL 4, VL 14 and VL 22; the pair VL 8 and VL 11; and also VL 3 reading *non enim habebat potestatem*.[9] These attestations imply that this reading was known in Italy at least by the late fourth century. So perhaps 032 (and other lost witnesses) are the source for both that and Nicetas' text – either directly or as a result of contamination.

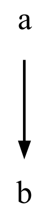
Considering pre-genealogical coherence data for 032 (see Table 89) we find that the catena manuscripts have the following similarity to 032: 249 (rank thirty-four, 85.1%); 333 (rank fifty, 84.6%); 430 (rank ninety-six, 83.0%); and 869 (rank twenty-eight, 85.4%). Those are not strong relationships to 032, but since we have decided (from a text-critical

---

[8] See Clark, 'The Catena of Nicetas of Heraclea and Its Johannine Text', 50 for Clark's stemma of these manuscripts.

[9] See Houghton, *The Latin New Testament*, 166.

perspective) that this must be a connective variant we are left assuming that some "missing link" manuscripts must be lost in time.



a	ηθελεν	FΠ MT F1 F13 𝔓66 𝔓75 01 03 04 05 07 011 019 022 028 033 037 038 044 045 0141 0211 18 33 35 109 138 157 168 173 213 226 295 357 377 382 397 544 579 597 732 792 799 807 821 841 865 892 992 994 1009 1010 1014 1029 1071 1093 1128 1230 1241 1242 1253 1293 1319 1320 1321 1344 1424 1463 1546 1561 1571S 1654 1788 1797 2106 2192 2223 2411 2561 2575 2585 2615 2680 2718 2766 2768 2786 2790
b	ειχεν εξουσιαν	032 249 333 430 869
	lac	...

Table 88: Local stemma for John 7:1/24

Chrysostom cites this passage as ειχεν εξουσιαν in three of the four places where he quotes it – and this must be Nicetas' source. 032 and Chrysostom presumably share some ancestry, inasmuch as an older Greek witness (or witnesses) now lost attested the reading. But 032 and the four Nicetas manuscripts are not related. So this raises the question: "what is multiple emergence?" This reading is found in unrelated manuscripts – but a sound theory exists for how it got there without needing it to have emerged multiple times. Perhaps this is a limitation caused by considering only the Greek witnesses, and not the entire contents of the ECM at this stage.[10] Further research is required on these questions. Even though *b* is the

---

[10] Of course, to include all possible witness types in the CBGM would radically increase the complexity, and so the time and cost of making the edition.

harder reading, the overwhelming support (both numerically and based on age) for the *a* reading is enough to convince the textual scholar of its primacy.

W2	NR	PERC1	EQ	PASS
07C	1	97.727	86	88
1797C	2	97.590	81	83
017C	3	97.561	40	41
892SC	4	95.652	66	69
2192C	5	94.118	144	153
01Ccb2	6	92.727	51	55
037C	7	92.424	61	66
168C	8	91.447	139	152
1546C	9	91.304	84	92
A	10	91.237	4144	4542
2766C	11	91.111	82	90
019C	12	90.385	47	52
029	13	89.281	683	765
𝔓75	14	88.381	2708	3064
𝔓66C*	15	88.248	413	468
03	16	88.196	4386	4973
019	17	87.808	4278	4872
1788C	18	87.805	72	82
04	19	87.737	1760	2006
33	20	87.247	4262	4885
044	21	86.924	4321	4971
892	22	86.558	2228	2574
F1	23	86.169	4274	4960
1344C	24	86.111	62	72
070	25	85.747	758	884
033	26	85.721	3398	3964
FΠ	27	85.700	4231	4937
869	28	85.430	1800	2107
597	29	85.394	4233	4957
213	30	85.343	3575	4189
02	31	85.320	3551	4162
109C	32	85.294	58	68
2561C3	33	85.209	265	311
249	34	85.095	2649	3113

Table 89: Pre-genealogical coherence for W1=032 (showing up to rank 34)

So, let us look at the textual flow diagrams for reading *b*. The difficult question again, is "what is a suitable connectivity setting to use?" Is ten a good choice, or twenty? Indeed, is 90% acceptable or should we insist on 95% (or allow 85%)? The default connectivity of ten results in the diagram in Figure 73.[11] Here 869 descends from its first ranked potential ancestor 397 (92.6%). Note that this rank refers to genealogical coherence, whereas the previous paragraph and table refer to pre-genealogical coherence. Hence the ranks are not comparable.

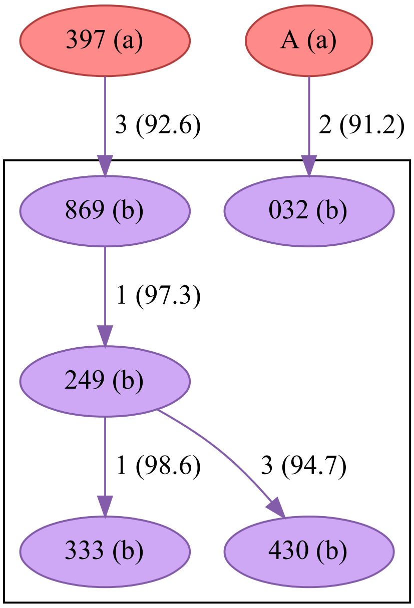


Figure 73: Textual flow diagram for John 7:1/24 (reading b) with connectivity 10

Since we have already decided that this is a connective variant, and therefore the multiple emergence shown here must be considered unlikely, we are left with three ways forward (generally speaking): first we could try a different connectivity setting; secondly we could invert the local stemma and see whether the coherence improves (this is often the case as manuscripts commonly have witness A as a relatively high ranking ancestor); thirdly we could consider the strength of the textual flow.

Now connective variants allow a looser connectivity setting by definition, as text

---

[11] The default for Acts was changed to five, because "For this writing a more cautious approach is appropriate."Wachtel, 'Text-Critical Commentary' in Strutwolf et al., *ECM III (Acts)*, pt 3, 2.

critical knowledge is not overridden by the CBGM. Yet we are forced to increase the connectivity to thirty-three before reading *b* gains perfect coherence (see Figure 74). Recall that 032 is a fourth/fifth century codex, while the other witnesses to *b* come from the eleventh to thirteenth centuries. This is an interesting example of an old reading being preserved in much later manuscripts, with only one early manuscript surviving to attest it. Now rank thirty-three seems low (i.e. a large number), and a coherence of 85.4% is not much better. Given that this is a connective variant then perhaps any other witnesses to reading *b* (from the intervening six or seven centuries) have been "corrected" before they were copied. This is not pure speculation, as we know that 430's corrector changed reading *b* to *a*. Perhaps the missing link witnesses have just not been unearthed yet, or maybe they exist but have been excluded from the ECM edition of John. If we invert the textual flow, so *b* becomes the initial text, we find a situation that is not much improved (see Figure 75). We still require a connectivity setting of sixteen or 88.6% to get perfect coherence. Now sixteen seems like a significant improvement on thirty-one, but 88.6% is only marginally better than 85.4%. These values are not conclusive.

The final thing to do is to consider the strength of the textual flow for each relationship. A → 032 has strength 322/0, so is strong and indisputable.[12] 032 → 869 is 128/65 and still strong. 869 → 249 is 23/20 and very weak. All that would be required would be the reversal of three local stemmata in the whole Gospel for this textual flow to be reversed. However, that does little to help as even though 869 would have a strong ancestor, 249 would descend from 032 with an even worse rank forty-six and 85.1%.[13] 249 → 333 is 16/10 and so quite weak. 249 → 430 is 51/21, or moderate strength. The full genealogical

---

[12] I will use this notation for the strength of a textual flow. X/Y where X is the number of variants where W2 (ancestor) is prior, and Y is the number where W1 (descendant) is prior. The bigger the difference, the stronger the flow.

[13] These values might be slightly changed with the change of other variant units, but presumably not much.

coherence data for these witnesses can be found in Appendix 4.

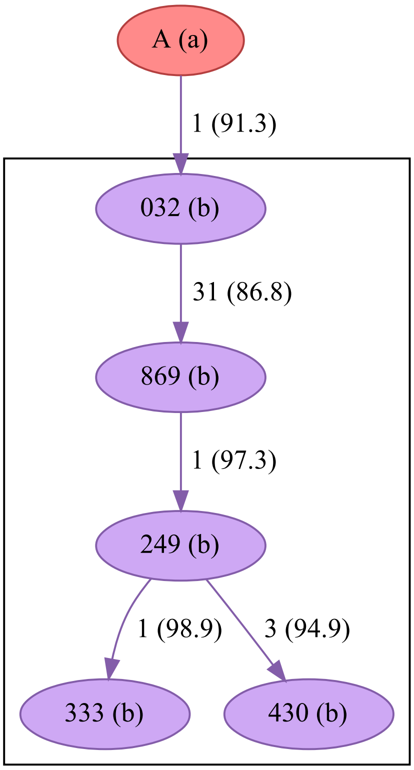


Figure 74: Textual flow diagram for John 7:1/24 for reading b with absolute connectivity

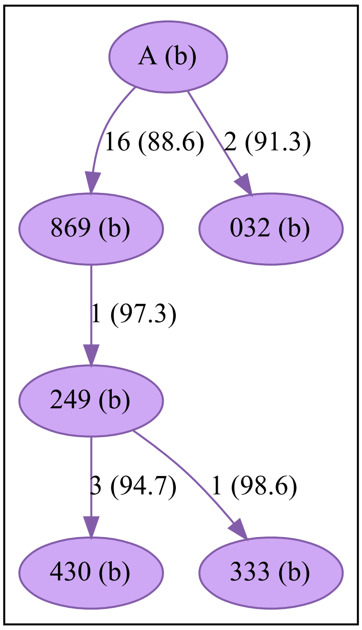
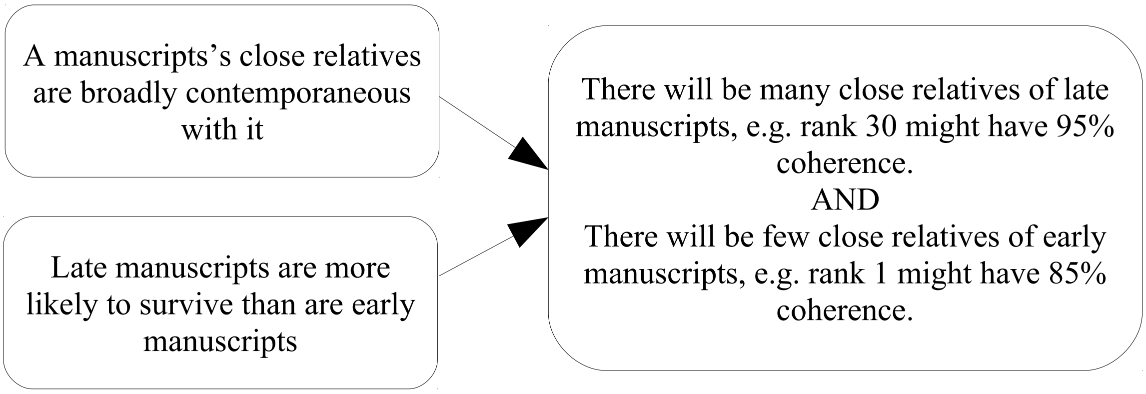


Figure 75: Textual flow diagram for John 7:1/24 for reading b (as initial text) with absolute connectivity

All in all, the textual flow suggests that the extant witnesses do not support this as a connective variant – and imply multiple emergence. Of course, if more witnesses were extant from the intervening centuries then the story might be quite different.

Percentage versus Rank

In each of the two examples above I have shown the date of the manuscripts. Now this might have caused some concern to the reader, as the CBGM deals in texts and not manuscripts (see Chapter 1). But actually this is a case where the manuscript itself, and particularly its date, matters. This statement is, I believe, undeniable: the later a manuscript the more likely it is to have been preserved. Or conversely, early manuscripts are more likely to have been lost than later ones. And manuscripts are copied to create new manuscripts. A second, uncontroversial statement is: *most of the time*, manuscripts were copied from other manuscripts only slightly older than themselves. In other words, the close relatives of a *late* manuscript are likely to be *late* manuscripts, and the close relatives of an *early* manuscript are likely to be *early* manuscripts. These two ideas can be combined as follows:



So setting a particular connectivity value (maximum rank) will lead to textual flow diagrams that are more likely to show multiple emergence for variants in late manuscripts. And early variants in early witnesses are likely to appear to have good or perfect coherence because relatively distantly related witnesses may have rank one or two. This is a crucial point for the interpretation of textual flow diagrams.

The converse of the previous paragraph is true if the percentage is used instead of the rank for connectivity. Straightforwardly, earlier manuscripts do not have many extant close relatives but late manuscripts (generally) have many. So it is much easier to find a potential ancestor with an acceptable coherence percentage in later witnesses.

In summary, the CBGM is likely to suggest multiple emergence in later readings if the rank is used but in earlier readings if the percentage is used. If the CBGM allowed the user to specify a maximum rank *and* a minimum coherence percentage then multiple emergence would be indicated in both situations (i.e. not hidden).

Experiment: John 18

Introduction and Coherence Tables

Several years ago the IGNTP created local stemmata for the complete collation of 1,662 witnesses to John 18 (see Chapter 5) yielding 267 variant units.[14] In contrast, Parker's ECM collation has 316 variant units for only 133 witnesses. For the current purpose, I have excluded correctors from this data, since the IGNTP data did not include them, and thus the comparison will be more appropriate. Of these 133, ninety-seven (non-corrector) witnesses are extant in at least 5% of variant units in John 18 of the ECM, and hence are included here.[15] I have imported both sets of local stemmata into my CBGM implementation.[16] The abridged pre-genealogical coherence for W1=032 is shown for each dataset in Figure 73 (IGNTP) and Figure 74 (ECM).[17]

The tables are difficult to compare, since so many witnesses exist only in the IGNTP data. The ranks change considerably between the tables yet the coherence percentages do not – for example:

- 03 and 1071 – ECM equal rank five (88.9%), IGNTP equal rank twenty-six (87.3%)
- 01 – ECM rank sixteen (87%), IGNTP rank forty-eight (85.8%) (See Appendix 5)

---

[14] Note that in Chapter 5 I refer to 1,663 witnesses, the extra one being 'A', the critical text.

[15] 𝔓108, 𝔓52, 𝔓59, 𝔓60, 𝔓66, 𝔓90, 01, 0109, 011, 0141, 019, 02, 0211, 022, 028, 0290, 03, 032, 033, 037, 038, 04, 044, 045, 05, 054, 05S, 07, 087, 1009, 1010, 1014, 1029, 1071, 109, 1093, 1128, 118S, 1230, 1241, 1242, 1293, 1319, 1320, 1321, 1344, 138, 1424, 1463, 1546, 1561, 157, 1571S, 1654, 168, 1788, 1797, 18, 2106, 213, 2192, 2223, 226, 2411, 2561, 2575, 2615, 2680, 2718, 2766, 2768, 2786, 295, 317, 33, 333, 35, 357, 377, 397, 544, 579, 597, 732, 792, 799, 807, 821, 841, 865, 892S, 992, 994, F1, F13, FΠ, MT. Note, F1 is Family 1, F13 is Family 13, FΠ is Family Pi and MT is the majority text. See Chapter 5 for details of Families 1 and 13.

[16] CBGM tag 1.11 was used for this experiment for both datasets (IGNTP and Parker's ECM). Note that this is John 18 only, and so the coherence tables and textual flow diagrams will differ from the whole-gospel ones earlier in this chapter.

[17] For the full tables, see Appendix 5.

W2	NR	PERC1	EQ	PASS
𝔓108	1	100.000	10	10
𝔓52		100.000	9	9
2418	3	94.828	55	58
1803		94.828	55	58
1143	5	93.220	110	118
𝔓90	6	92.857	13	14
1590	7	91.346	95	104
A	8	90.763	226	249
1804	9	90.698	78	86
2649	10	89.423	93	104
2634	11	89.091	49	55
2307	12	89.062	114	128
1349	13	88.889	120	135
475s		88.889	24	27
546	15	88.679	141	159
748	16	88.281	113	128
96	17	88.000	154	175
2717	18	87.975	139	158
1571	19	87.919	131	149
405	20	87.861	152	173
2722	21	87.817	173	197
2372		87.817	173	197
179	23	87.778	158	180
NA27	24	87.640	234	267
940	25	87.293	158	181
1071	26	87.266	233	267
03		87.266	233	267
947	28	87.097	27	31
731s		87.097	27	31
1400	30	87.050	121	139
274	31	86.957	140	161
711	32	86.875	139	160
...				

Table 90: Pre-genealogical coherence with W1=032 using **IGNTP data** (top 32 rows). Witnesses highlighted in grey are not present in the ECM data.

W2	NR	PERC1	EQ	PASS
𝔓108	1	100.000	21	21
𝔓90	2	93.750	60	64
𝔓52		93.750	30	32
A	4	90.789	276	304
1071	5	88.924	281	316
03		88.924	281	316
1571S	7	88.889	136	153
33	8	88.608	280	316
019	9	87.975	278	316
1321	10	87.937	277	315
0109	11	87.368	83	95
FΠ	12	87.342	276	316
397		87.342	276	316
044		87.342	276	316
1293	15	87.097	27	31
865	16	87.025	275	316
033		87.025	275	316
02		87.025	275	316
01		87.025	275	316
04	20	86.905	219	252
333	21	86.901	272	313
0290	22	86.873	225	259
F1	23	86.709	274	316
226		86.709	274	316
1561		86.709	274	316
MT	26	86.392	273	316
597		86.392	273	316
35		86.392	273	316
18		86.392	273	316
1320		86.392	273	316
1242		86.392	273	316
07		86.392	273	316
037		86.392	273	316
...				

Table 91: Pre-genealogical coherence with W1=032 using ***ECM data*** *(top 33 rows). Witnesses highlighted in grey are not present in the IGNTP data.*

These examples follow the expected pattern of roughly similar coherence percentages, but much lower ranks in the IGNTP data.

Tables of genealogical coherence data (potential ancestors) are easier to compare. Consider W1=04: Table 92, IGNTP; and Table 93 ECM. In these short tables there seems to be no problem – in the IGNTP data there are only four potential ancestors (A, 03, NA27, 1803) and in the ECM only two (A, 03) – and the ranks and coherence percentage of the only shared witness 03 is basically the same in both tables (rank two, ~93%).[18]

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
2726		-	100.000	13	13				
𝔓59		-	100.000	10	10				
2290		-	100.000	9	9				
𝔓52		-	100.000	3	3				
1143		-	95.918	94	98	1	1		2
A	1		93.056	201	216	12		1	2
03	2		92.544	211	228	9	2	2	4
NA27	3		92.105	210	228	12	1	3	2
2307		-	92.035	104	113	3	3	3	
1803	4		91.379	53	58	2	1	2	
96		-	91.176	124	136	3	3	5	1
019		-	90.789	207	228	8	8	3	2
2517		-	87.234	82	94	4	4	2	2
2676		-	77.000	77	100	6	6	9	2

*Table 92: Potential ancestors for W1=04 (**IGNTP data**).*
Witnesses highlighted in grey are not present in the ECM data.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
𝔓52		-	100.000	11	11				
𝔓90		-	100.000	3	3				
A	1		94.583	227	240	12		1	
𝔓59		-	93.333	14	15			1	
03	2		92.857	234	252	11		6	1
019		-	91.667	231	252	8	8	5	

*Table 93: Potential ancestors for W1=04 (**ECM data**)*
Witnesses highlighted in grey are not present in the IGNTP data.

---

[18] Note that 'A' is technically a different witness in each dataset, as it is a different critical text.

A longer table, however, does highlight the problem. Consider the genealogical coherence tables for 032: Table 94, IGNTP; and Table 95, ECM. Here the IGNTP includes nine manuscripts that rank higher as potential ancestors for 032 than any witness included in the ECM.[19] This is expected, as the ECM has excluded many witnesses because of their similarity to those that were included. Thus, for example, 019 is rank four (89%) in the ECM data, yet rank eighteen (86%) in the IGNTP data. Now the rank is significantly different; a change that is more than enough to suggest multiple emergence if a connectivity setting of, say, ten ("Average") were to be used to draw a textual flow diagram. Yet the coherence percentage is less than 3% different – and assuming that people like round numbers (multiples of five) then the textual flow diagrams based on coherence percentage would be unchanged for this relationship.

---

[19] Plus NA27 and its A text.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
𝔓108		-	100.000	10	10				
𝔓52		-	100.000	9	9				
1803	1		94.828	55	58	2		1	
1143		-	93.220	110	118	4	4		
A	2		90.763	226	249	20		3	
2307	3		89.062	114	128	7	3	4	
1349	4		88.889	120	135	6	5	3	1
475s		-	88.889	24	27	1	1	1	
96	5		88.000	154	175	8	6	7	
2717	6		87.975	139	158	7	6	5	1
1571		-	87.919	131	149	7	7	3	1
2372	7		87.817	173	197	9	8	4	3
NA27	8		87.640	234	267	23	1	9	
1071	9		87.266	233	267	15	10	8	1
03			87.266	233	267	18	2	11	3
1400	11		87.050	121	139	6	3	8	1
274	12		86.957	140	161	8	6	7	
2632	13		86.700	176	203	10	9	7	1
33	14		86.692	228	263	14	11	8	2
2398	15		86.667	143	165	7	6	8	1
233	16		86.466	230	266	13	9	11	3
1816	17		86.364	228	264	13	10	11	2
930		-	86.344	196	227	11	11	8	1
1558		-	86.207	200	232	11	11	8	2
019	18		86.142	230	267	18	7	10	2
2524	19		85.932	226	263	13	12	10	2
904	20		85.827	109	127	7	6	5	
669	21		85.768	229	267	13	11	12	2
235		-	85.768	229	267	12	12	12	2
1566			85.768	229	267	13	11	13	1
139			85.768	229	267	13	11	12	2
1006			85.768	229	267	16	11	9	2
044			85.768	229	267	15	10	11	2

...

*Table 94: Potential ancestors (abridged) for W1=032 (**IGNTP data**). Witnesses highlighted in grey are not present in the ECM data.*

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
𝔓108		-	100.000	21	21				
𝔓90	1		93.750	60	64	3	1		
𝔓52		-	93.750	30	32	1	1		
A	2		90.789	276	304	23		4	1
1071		-	88.924	281	316	12	12	8	3
03	3		88.924	281	316	22	1	10	2
019	4		87.975	278	316	17	8	11	2
1321		-	87.937	277	315	13	13	10	2
0109	5		87.368	83	95	6	4	2	
1293		-	87.097	27	31	1	1	2	
865		-	87.025	275	316	14	14	10	3
033		-	87.025	275	316	14	14	10	3
04	6		86.905	219	252	13	8	11	1
05S	7		83.810	176	210	11	10	8	5
087	8		80.952	34	42	2	1	4	1
𝔓59	9		68.750	11	16	2		2	1

*Table 95: Potential ancestors for W1=032 (**ECM data**).*
Witnesses highlighted in grey are not present in the IGNTP data.

Textual Flow

If the coherence tables can be hard to compare, then what about textual flow diagrams? Let us begin with a simple example: John 18:2/24-36. There are ten readings in the ECM. Reading *a* is the rarely attested text chosen by NA27 and by Parker for his initial text: *συνηχθη ιησους εκει μετα των μαθητων αυτου*. In the ECM this is attested by only 01, 019, 033 and 865. 03 reads the same, but misses the *εκει*. 𝔓66 misses the *αυτου* – although the firsthand corrector notices the omission and adds it.[20] All other readings include the definite article before *ιησους*. In the IGNTP data this reading is labelled *2*.[21] Only one additional witness is included, 2591. 03 has its own reading (*6*) again. The reading attested by 𝔓66 in the ECM is included in the IGNTP data, but only 798 is shown as attesting it (a witness not included in

---

[20] Due to 𝔓66*'s many unique readings it is preserved as a witness, even when 𝔓66C* changes it. Other witnesses are represented by their firsthand corrector where it exists.

[21] This follows the convention from Aland, *Text Und Textwert Der Griechischen Handschriften Des Neuen Testaments*. There "1" = Majority; "2" = "NA"; "1/2" = where 1 and 2 agree; "3", "4", ... are the remaining readings.

the ECM). Somewhat surprisingly, 𝔓66 is shown as lacunose here in the IGNTP data. It is designated reading number 98 in the Münster database, indicating that the textual evidence does not allow unambiguous support for a single reading.[22] The IGNTP transcription for the two lines in question is as follows:

> [κι]ς συνηχθ[η ις εκει μετα των]
>
> [μα]θητων ο̣ [ουν ιουδας λαβων][23]

Presumably the ECM transcribers measured the space needed for the additional article and decided that it would not fit, so therefore 𝔓66 attests the *a* reading, without the final word which the scribe then added by correction. They did mark it as 𝔓66*V, i.e. *ut videtur* (“apparently”). The earlier IGNTP transcribers or collators must have been more circumspect and withheld judgement. This decision, and the many other such small, individual decisions, will change the coherence data and thus can change the textual flow.

But, moving on from 𝔓66, let us consider the textual flow diagrams of reading *a* or *2* shown in Figure 76 and Figure 77 respectively. In this example we have two textual flow diagrams that differ in the only non-trivial sub-tree: 019 is the ancestor of 033 and 865 in Figure 76 yet in Figure 77 865 descends from 033 and neither descend from 019. Note that the convention of absolute coherence equating to rank 499 is inappropriate for the IGNTP data, since 2591 needs to descend from its 970th ranked potential ancestor (A) to allow perfect coherence. It is straightforward to state that 2591 arrived at this reading independently from the rest of this group (i.e. by multiple emergence). Perhaps a more complicated example is required before more similarities will appear in the diagrams.

---

[22] This is the equivalent of Münster’s “zw” reading and “99” equates to “zz”, meaning truly lacunose.

[23] These are lines 15 and 16 of folio 125.
See http://www.iohannes.com/XML/transcriptions/greek/04_P66.xml#K-B04K18 [accessed 14/11/2017].

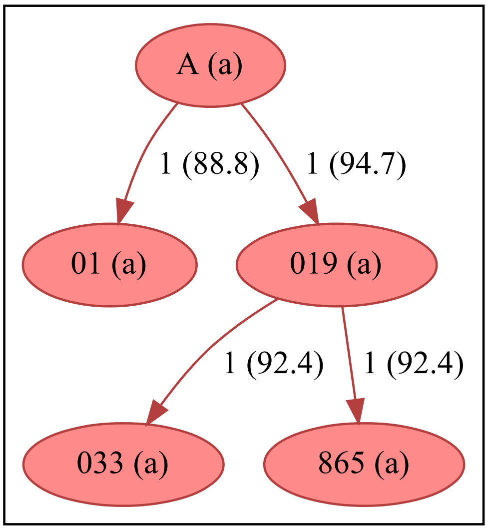


Figure 76: John 18:2/24-36 reading a (ECM data) with absolute coherence

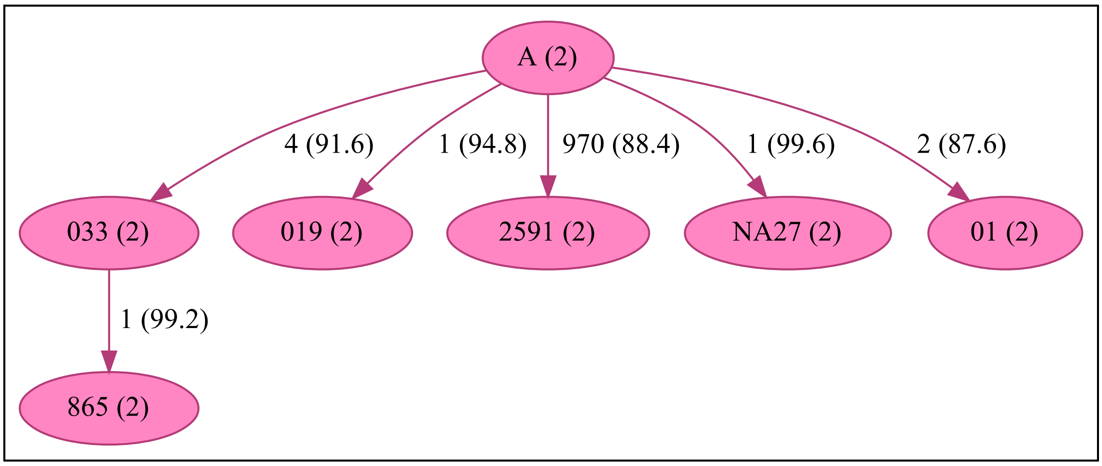


Figure 77: John 18:2/24-36 reading 2 (IGNTP data) with absolute coherence

Consider John 18:4/18-22, reading *a* or *2* (εξηλθεν και λεγει), shown in Figure 78 (ECM) and Figure 79 (IGNTP).[24] Note that Family 1 (F1 in the ECM diagram) is represented by 1582, 1, 209, 205, 2886, 2713 and 884 in the IGNTP diagram, which are greyed out for convenience.[25]

---

[24] The other main readings being εξελθων ειπεν and εξελθων λεγει.

[25] It is also very interesting to compare this diagram with the "Family 1" section in Chapter 5, "Phylogenetics of John 18". The shape of the Family 1 tree, with the additional "members", is very similar.

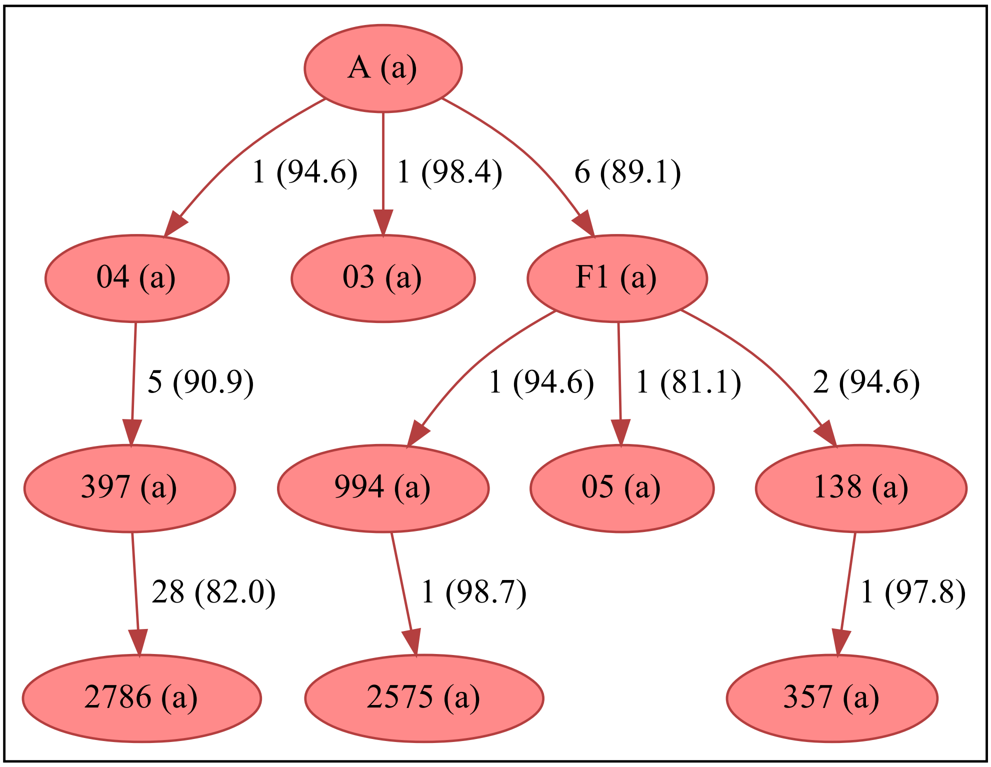


Figure 78: John 18:4/18-22 reading a with absolute coherence (ECM data)

It is clear that these two diagrams have a lot in common, notably: ECM: A → F1 → 994 → 2575, IGNTP: A → 1582 → 994 → 2575; and ECM: A → F1 → 138 → 357, IGNTP: A → 1582 → 994 → 138 → 357. Now 05 descends from F1 in the ECM diagram, but from A in the IGNTP. It is not surprising that some manuscripts move up to descend directly from A – it is often the highest ranked potential ancestor for early witnesses (e.g. 05). The final major feature in the ECM data (A → 04 → 397 → 2786), however, is totally missing in the IGNTP diagram. Here 397 descends by itself from A, and 2786 is at the end of the following chain: A → 1582 → 2702 → 2786. How can 2786's position be explained? The rank is the key. The ECM needs 2786 to descend from its twenty-eighth ranked potential ancestor, and the IGNTP its 754th! It is clear that the fact of 2786 attesting this reading is a co-incidence, and this is a case of multiple emergence.

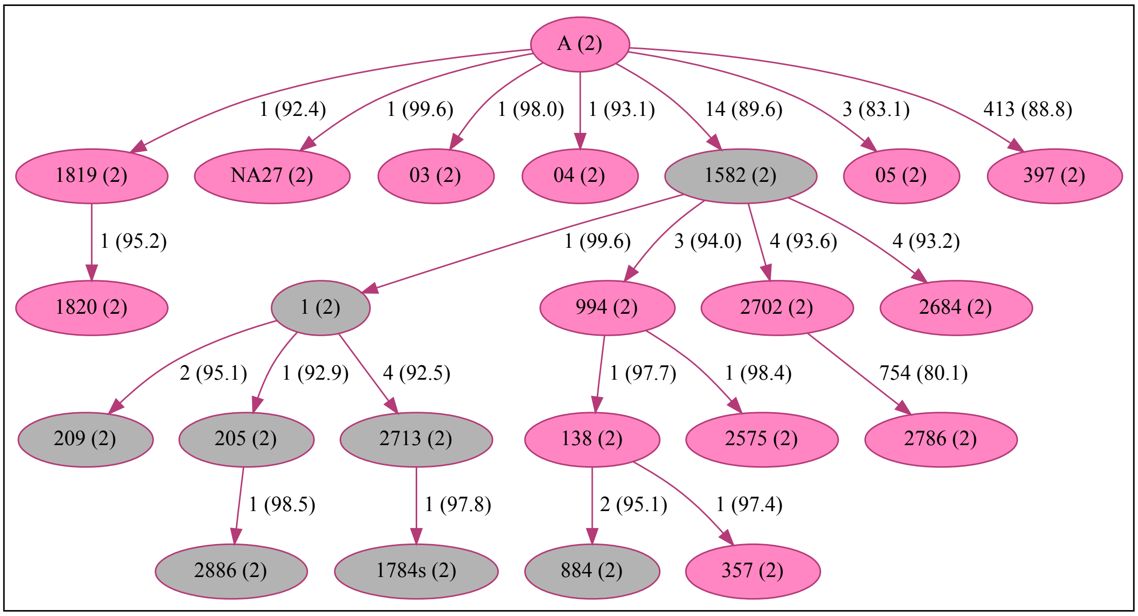


Figure 79: John 18:4/18-22 reading 2 with absolute coherence (IGNTP data). Family 1 manuscripts are greyed out.

But what about 397? Why does it not descend from 04? The surprising answer is that, in the IGNTP data, 04 is 397's 415th highest ranked potential ancestor (87.3%).[26] In the ECM data 04 is 397's fifth ranked potential ancestor (90.9%) – see Table 96. Even if the coherence percentage were the same as in the IGNTP's data (87.3%) then the rank would only be thirteen – which might easily be accepted for a connective variant in a textual flow diagram (even if it would be the lowest ranked potential ancestor of all for 397). Note that the different of 3% coherence in the ECM data (in this case) equates to a difference in rank of eight – but the equivalent 3% in the IGNTP data equates to only fifty-three – out of 423. This underlines the fact, argued by this chapter, that ranking (alone) is not an objective measure.

---

[26] See Appendix 5 for the genealogical coherence data for W1=397.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
𝔓90		-	93.750	60	64	2	2		
𝔓59		-	93.750	15	16			1	
1071	1		93.354	295	316	9	6	6	
33	2		92.722	293	316	10	7	5	1
0109		-	92.632	88	95	3	3	1	
1321	3		92.381	291	315	10	8	6	
213		-	92.089	291	316	9	9	7	
865	4		91.456	289	316	12	9	6	
04	5		90.873	229	252	16	3	4	
033	6		90.823	287	316	13	10	6	
0290		-	90.347	234	259	8	8	9	
1571S		-	90.196	138	153	6	6	3	
A	7		90.132	274	304	26		3	1
044	8		89.873	284	316	13	11	7	1
02	9		89.241	282	316	13	11	8	2
019			89.241	282	316	19	7	7	1
03	11		88.608	280	316	25	1	8	2
032	12		87.342	276	316	13	11	12	4
087		-	85.714	36	42	1	1	2	2
𝔓108		-	85.714	18	21	1	1	1	

Table 96: Potential ancestors for W1=397 using ***ECM data***

In other words, creating a textual flow diagram on ranking alone involves *implicitly* choosing different metrics for each relationship within the diagram. This implicit metric involves an interplay between the following factors (at least): First consider the total number of potential ancestors. A witness with few potential ancestors will necessarily only have "high-ranked" ancestors – using the rank alone as a metric. A witness with many potential ancestors may have mostly (apparently) low ranked ancestors. Secondly consider the total number of witnesses in the data set (or the number excluded from it). An eleventh ranked ancestor with 88.6% coherence (such as 03 in the ECM data) cannot be considered even broadly the same as the thirteenth ranked ancestor with 93.6% (such as 926 in the IGNTP data). Remove many of the witnesses and the IGNTP data becomes the ECM (roughly speaking). The witnesses have not changed, but the results would.

This John 18 experiment has shown that the rank of a potential ancestor is too easily affected by factors outside the textual data itself. If all witnesses shared a character (for example they all had a similar number of closely related other witnesses) then the problems identified here would be lessened in effect – for example choosing only majority text witnesses. But it is a mistake to expect to be able to compare the rankings of potential ancestors of witnesses with few close relatives – such as 𝔓66, 01 or 05 – with those of majority text witnesses such as 35 or 226.

Conclusion

This chapter leads directly to two main conclusions: First, textual flow diagrams themselves are not sufficient to determine whether a reading is a case of multiple emergence or not (this is not a new thing to say); and secondly (and new) textual flow diagrams should show the rank and the percentage, thus providing the reader with the necessary information to identify places that require further investigation.

When interpreting textual flow diagrams the scholar needs to remember that a diagram drawn with a connectivity (rank) setting will be more likely to show multiple emergence for later variant readings and to show perfect coherence for earlier ones. Conversely, a diagram drawn with a coherence (percentage) setting will tend to show multiple emergence for earlier readings and perfect coherence for later ones. It would be valuable if a maximum rank *and* a minimum coherence percentage could both be specified when creating a textual flow diagram – then multiple emergence would be indicated in both situations (i.e. not hidden).

A final thought is that this conclusions of this chapter hold true only for manuscript traditions where many manuscripts have been lost. Where all manuscripts survive, then using the rank probably raises no problems.

Chapter 8: The CBGM: The Strength of Textual Flow

The Strength of Textual Flow is Invisible

Several scholars have raised concerns about the various strengths of flow found in textual flow diagrams. For example, Jongkind argues that the CBGM's textual flow diagram for 1 John 4:4/6a contains eight (out of seventy-three) suspect relationships, many of which occupy key locations in the textual flow diagram.[1]

The key problem here with textual flow diagrams is that it is not clear how strong any given textual flow is. Consider John 1:28/4-8. The majority of witnesses attest either reading *a*, εν βηθανια εγενετο, or reading *d*, εν βηθαβαρα εγενετο.[2] The textual flow diagram (drawn showing the percentage coherence and rank of the ancestor, as recommended in Chapter 7) in Figure 80 has a variety of strengths of textual flow, but this is invisible to the reader. Note that in the INTF's "Genealogical Queries" user interface it is possible to show the coherence table, and thus the potential ancestors can be made visible (along with the data necessary to calculate the strengths of textual flow) – but the user needs to choose to seek out this information. This textual flow diagram is of interest as it shows several core relationships often found forming a spine in the textual flow diagrams for John, namely: F1 → FΠ → 18 → 35. These relationships become so familiar to the CBGM user (in this data) that they are unconsciously accepted as being "normal" and thus in some sense "the truth".

---

[1] See Jongkind, 'On the Weighing and Counting of Variants', 13.

[2] Intriguingly, 04's corrector changed the text from reading *a* to *d*. Witnesses for reading *a*: MT, 𝔓75, 02, 03, 04, 07, 011, 019, 028, 032S, 033, 037, 038, 044, 045, 063, 0211, 109, 157, 168, 173, 213, 295, 333, 357, 377, 382, 397, 430, 544, 579, 597, 732, 792, 807, 841, 865, 892, 992, 1009, 1010, 1014, 1071, 1241, 1242, 1253, 1293, 1319, 1320, 1321, 1344, 1424, 1463, 1654, 1788, 1797, 2192, 2223, 2561, 2575, 2585, 2615, 2680, 2718, 2768, 2786. Witnesses for reading *d*: FΠ, F1, F13, 04C1, 029, 083, 0141, 18, 33, 35, 226, 821, 1093, 1128, 1230, 1546, 1561, 1571, 2106, 2411.

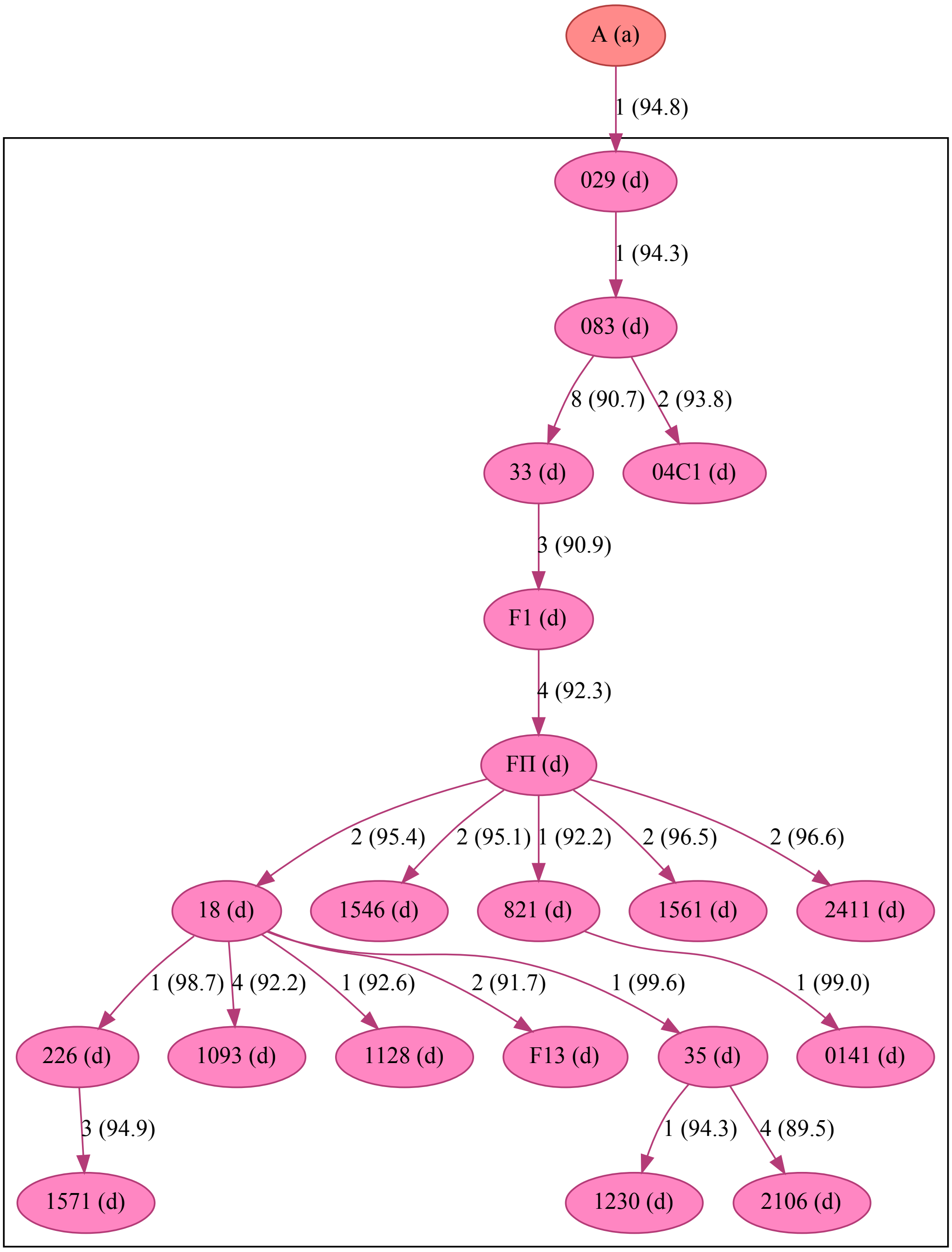


Figure 80: Textual flow diagram for John 1:28/4-8 (reading d) with absolute connectivity

Visible Strength

The strength of textual flow in this spine is actually rather weak, and should cause the scholar to doubt some relationships in the diagram – but to find this out they must turn to genealogical coherence tables, which is time consuming and relatively opaque. If instead the strengths were shown on the diagram, then the scholar would immediately be alerted to any doubtful relationships. Consider Figure 81 where dotted lines represent very weak textual flow (five or less), dashed lines represent weak textual flow (from six to twenty-five) and solid lines stronger textual flow (above twenty-five). Now it is clear that the spine is made up of weak flow, but the strengths still cannot be precisely quantified. Note that a weak textual flow can exist between two witnesses that are strongly related. Indeed, this is likely as very closely related witnesses will have few differences and so the number of prior and posterior readings must be low.

So now consider Figure 82. Here the number of prior and posterior readings is shown underneath the existing relationship label (rank and coherence) as [prior/posterior]. Consider 18 → 35 which we can now see as [11/9]. This means these very similar manuscripts (99.6% coherence) differ in only twenty variant units in John; in eleven cases 18 is prior, yet in nine cases 35 is prior. This implies that 18 and 35 are much more like cousins than parent and child. They represent, together, a small assortment of readings in the majority text of their time. Because 18 and 35 are so similar, a change in one variant unit could force this relationship to become undirected, and two changes could reverse the flow. This matters because textual flow diagrams are used in the second pass of creating local stemmata. The user needs to be alerted to the fact that this relationship should not be used to determine the direction of a (local) genealogical relationship between variant readings. This would be a kind of forced feedback loop, where just the presence of a weak flow is used as an argument to make that very flow stronger (by modifying local stemmata to agree with the initially weak flow). This will be discussed further in Chapter 9. A weak textual flow between very closely related witnesses can be used positively, however, to state that there is likely to be a close relationship between their readings – even if the direction of that relationship cannot be identified from the textual flow data alone.

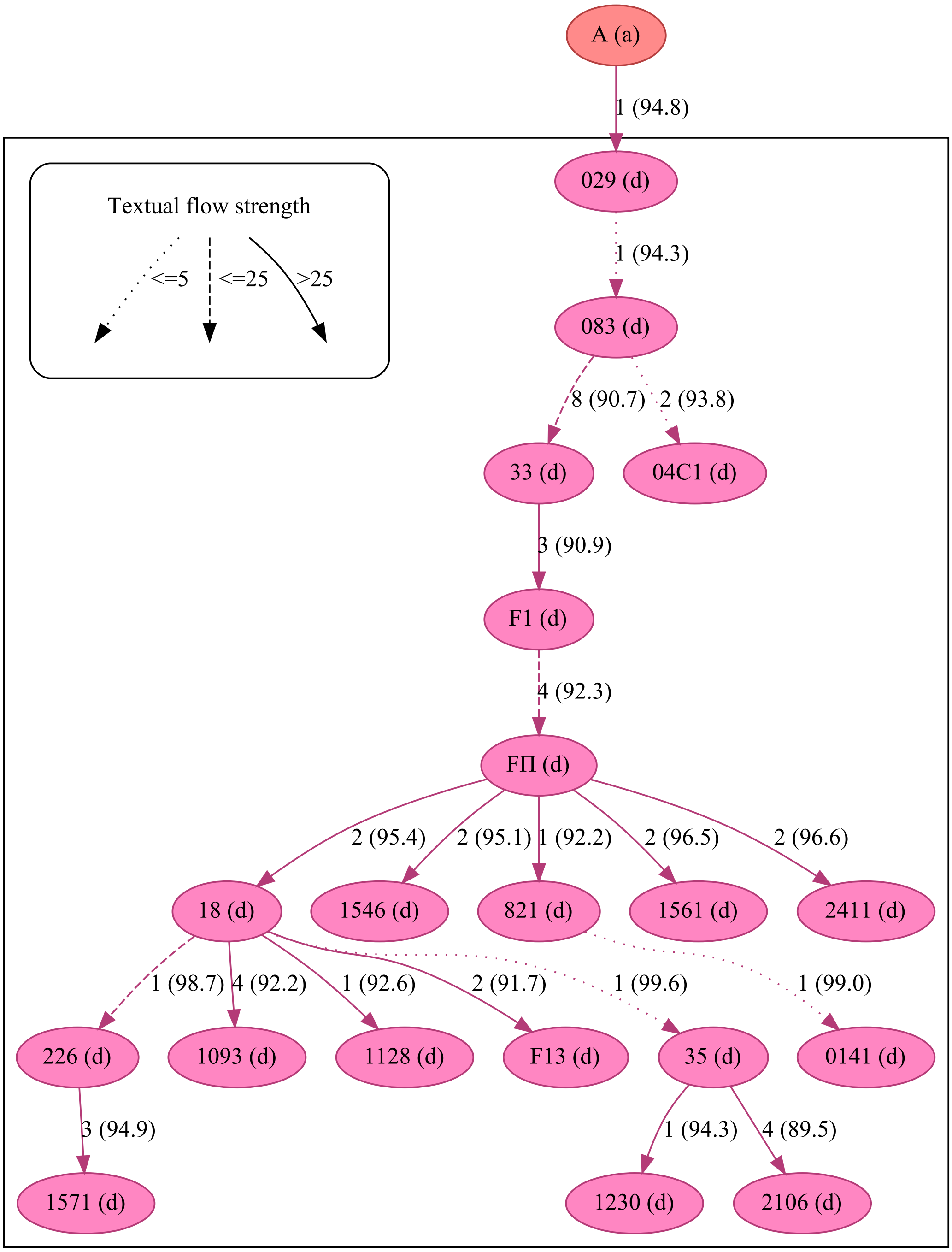


Figure 81: Textual flow diagram for John 1:28/4-8 (reading d) with absolute connectivity. Line styles represent strength of textual flow.

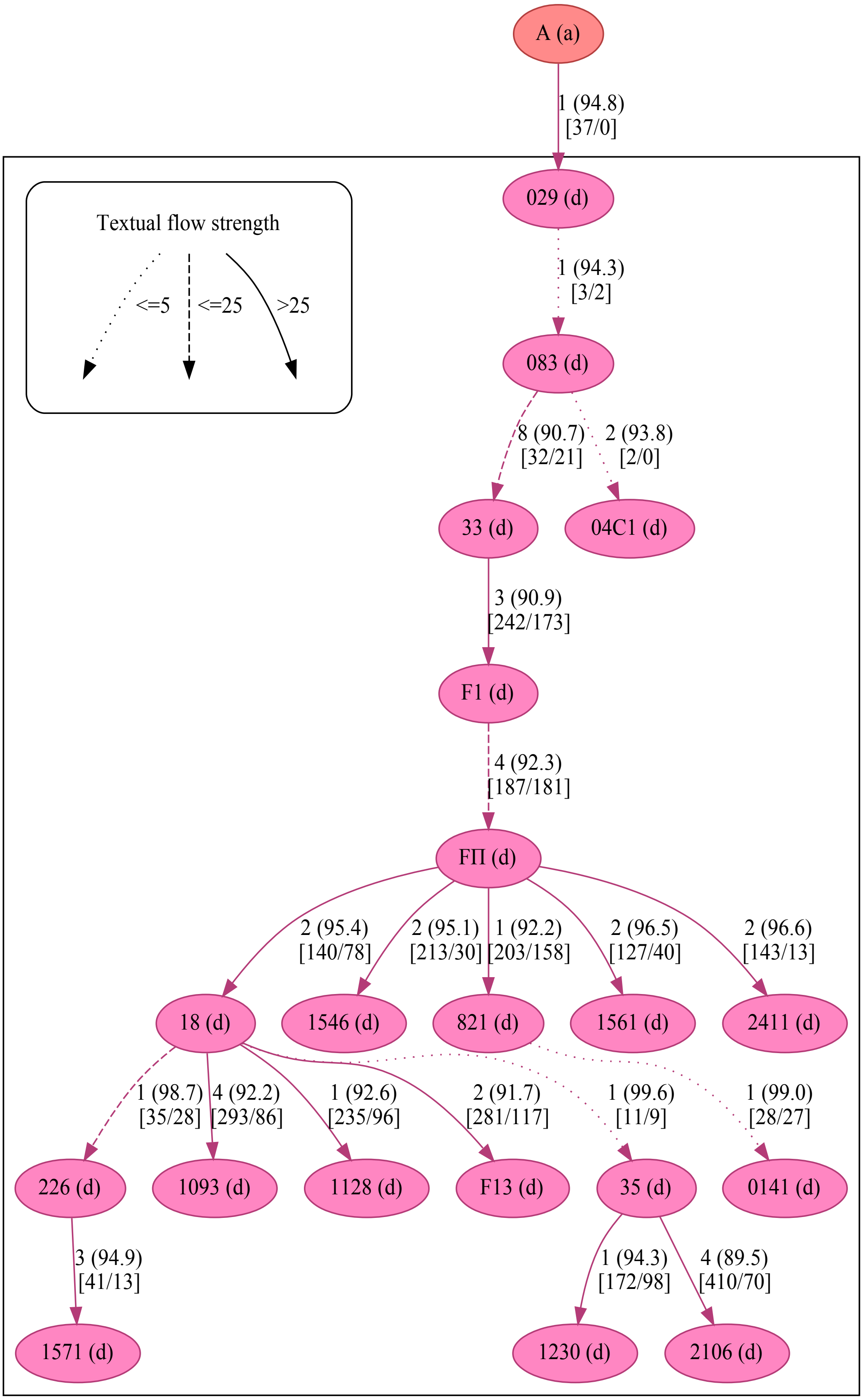


Figure 82: Textual flow diagram for John 1:28/4-8 (reading d) with absolute connectivity. Prior/posterior counts are shown in square brackets.

Now this diagram looks too cluttered, and so perhaps the line styles alone might suffice. But having the actual data shown in place is of great use. So perhaps it should be left to the individual scholar's preferences – but in all cases some visualisation of the strength must be represented.[3]

Minimum Strength

A further advance could be to set a minimum textual flow strength for a textual flow diagram. In other words, textual flow weaker than the threshold would result in the relationship being treated as undirected. Figure 83 shows the equivalent textual flow diagram to Figure 82 but with a minimum strength of five. So, for example, 083 can no longer descend from 029, but now descends from A – indicating multiple emergence of the reading. This highlights something about multiple emergence – it is easily affected by small changes. Multiple emergence can sometimes be shown because of a genuinely undirected genealogical relationship – where just one change to a local stemma could provide a suitable potential ancestor. Setting a minimum strength just widens the group of "undirected" relationships and so makes this scenario more likely. In this diagram we find threefold multiple emergence where before there was perfect coherence. This would be, of course, a kind of false multiple emergence – as who would genuinely argue that two very closely related witnesses sharing a reading is evidence of multiple emergence? So while this diagram would mitigate against the problem of the forced feedback loop described above, it increases the chance of wrongful identification of multiple emergence (which already exists in normal textual flow diagrams for witnesses with undirected relationships). It could provide valuable information, nevertheless, in a situation where closely related witnesses (with weak textual flow) differ in

---

[3] Gurry wrote a paragraph suggesting a similar solution to this problem, that of indicating the relative strengths of textual flow by changing the opacity of the lines connecting witnesses. See Gurry, *A Critical Examination of the Coherence-Based Genealogical Method*, 214.

attestation in a particular variant unit. In any case, it is important to remember that textual flow diagrams are always artificial, and do not claim to show historical manuscript relationships.

A further notable difference between Figure 82 and Figure 83 is that 35 no longer descends from 18, but is now its sibling. Setting a minimum strength is likely to have this flattening effect, as close parent/child relationships become siblings due to both witnesses having the same top-ranked potential ancestor (where before one had it as rank two with the now-sibling witness as rank one). So which is better? Each diagram is emphasising something about the data: Figure 82 shows that there is perfect coherence in reading *d*, made up of close witnesses. Figure 83 shows that the relationships are so close that the textual flow is too weak to rely on for defining/redefining local stemmata. Neither diagram is incorrect, but neither shows the whole truth either. See Figure 84 for a further flattened diagram with minimum strength ten.

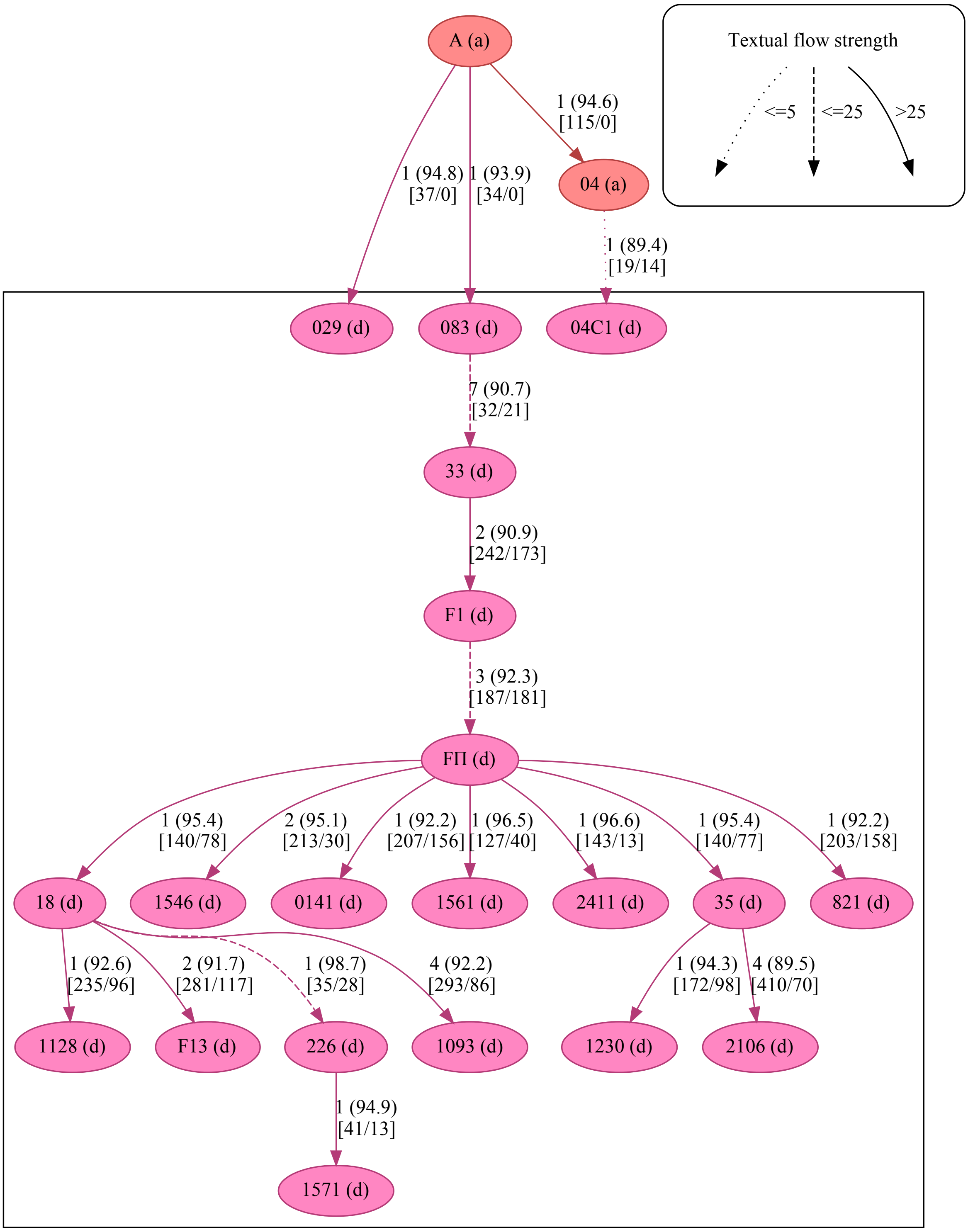


Figure 83: Textual flow diagram for John 1:28/4-8 (reading d) with absolute connectivity and minimum flow strength 5. Prior/posterior counts are shown in square brackets.

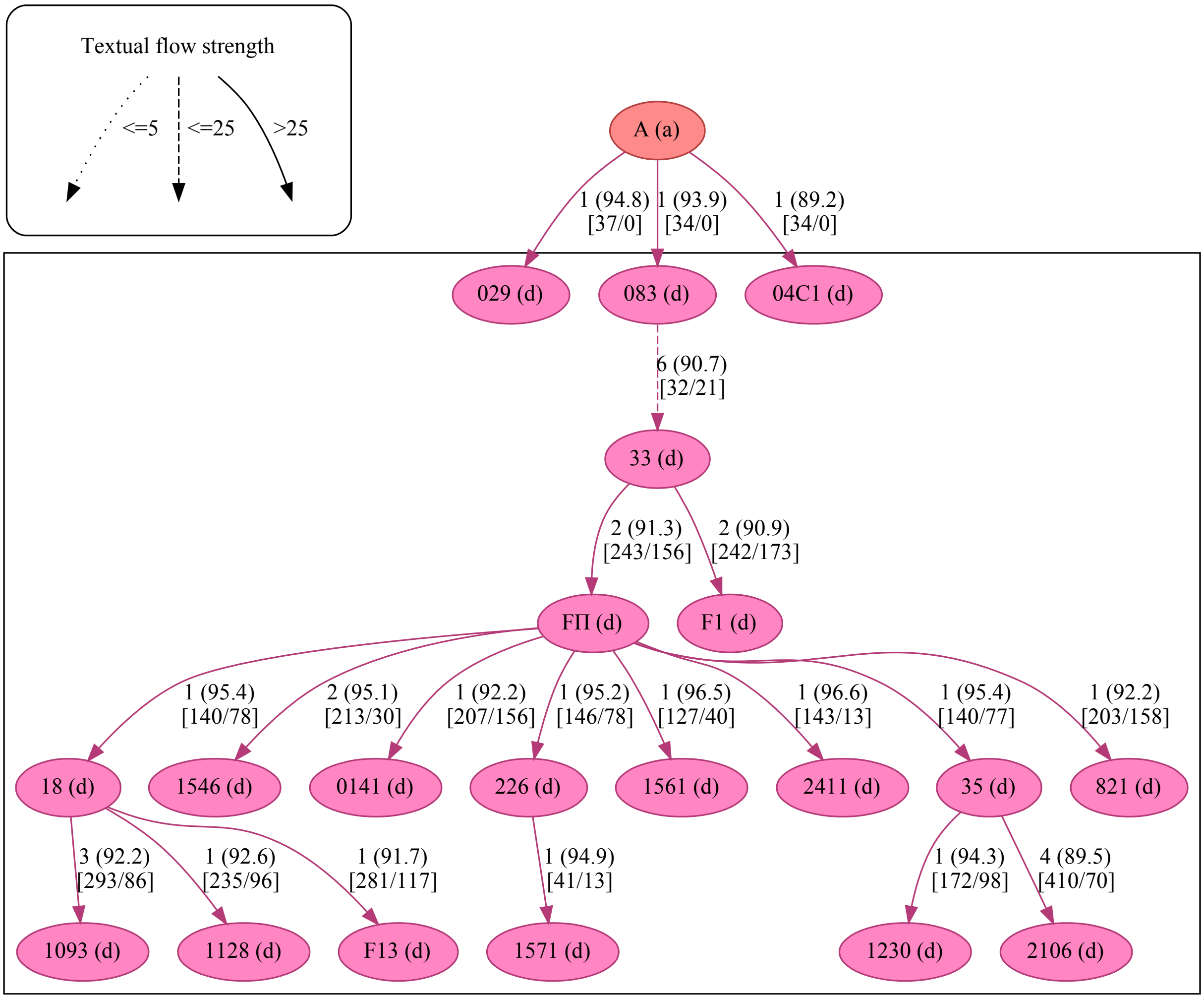


Figure 84: Textual flow diagram for John 1:28/4-8 (reading d) with absolute connectivity and minimum flow strength 10. Prior/posterior counts are shown in square brackets.

While there are obvious advantages to this method of setting a minimum strength for textual flow diagrams (and it would be particularly useful if the scholar could dynamically vary it via the user interface), it will also have the unfortunate effect of driving the diagrams further away from reality. Take, for example, 0141 and 821 – which in Chapter 5 were shown to be exemplar and copy.[4] In the textual flow diagram (Figure 82) 0141 is shown descending from 821 with twenty-seven prior variants compared to 821's twenty-eight – i.e. a strength of one. This is a real manuscript relationship that is hidden when the minimum strength is set – and this is to be expected as very close witnesses will no longer be considered as potential

---

[4] See Taylor Farnes, 'Scribal Habits in Selected New Testament Manuscripts' for a detailed examination of this relationship.

ancestors. Interestingly, a further feature of textual flow diagrams is highlighted here: The real relationship would be 0141 → 821 – but in Figure 82 *it is inverted*. This inversion is commonplace and will be discussed in more detail in Chapter 9.

Undirected Relationships

Consider the genealogical coherence table for FΠ, shown in Table 97. This might normally be referred to as the potential ancestors for FΠ, but not all the rows correspond to potential ancestors in the strict sense. For example the first row, 02, shows an undirected relationship where each of 02 and FΠ have 113 prior readings (with respect to each other) – highlighted by the "-" in the "D" column. In this table, I have also highlighted the witnesses where there is a directed relationship but the strength of textual flow is less than ten (so to correspond to the textual flow diagram in Figure 84). These rows have a "w" in the "D" column, indicating weak textual flow, and are treated as undirected. In Figure 84 these witnesses would also not count as potential ancestors.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL	READING
02		-	94.583	4994	5280	113	113	52	8	a
892		w	93.345	3212	3441	94	88	29	18	a
044	1		93.002	5595	6016	223	107	73	18	a
0290		-	92.871	495	533	12	12	10	4	
086		w	92.342	410	444	17	16	1		
F1		w	92.281	5547	6011	187	181	71	25	d
33	2		91.312	5381	5893	243	156	82	31	d
A	3		90.658	5376	5930	476		53	25	a
033	4		90.581	3895	4300	171	145	65	24	a
083	5		89.984	575	639	34	21	8	1	d
04	6		89.980	2263	2515	146	61	30	15	a
019	7		89.504	5270	5888	357	128	94	39	a
019C	8		89.235	315	353	24	6	8		
03	9		89.221	5306	5947	448	42	122	29	a
070		w	88.768	735	828	41	33	12	7	
𝔓75	10		88.743	3508	3953	261	96	71	17	
029	11		88.503	739	835	56	22	16	2	d

04C1	12		87.335	331	379	27	10	6	5	d
032	13		87.162	3965	4549	241	206	98	39	
𝔓66C*	14		87.078	1934	2221	112	98	56	21	
𝔓60		w	86.364	380	440	26	20	8	6	
01Cca	15		86.207	2200	2552	144	125	55	28	
01Ccb2		-	85.549	296	346	18	18	7	7	e

Table 97: Genealogical Coherence for W1=FΠ showing readings for John 1:28/4-8

Excluding weak or undirected relationships increases the confidence that the textual flow diagram is unlikely to change (as local stemmata are revised) and is therefore of great use during that process. But, as was said above, this moves the diagram further away from reality. What if we now move the bar in the other direction, and allow even undirected relationships in textual flow diagrams? These *potential* potential ancestors create a textual flow diagram where very close witnesses are shown as related (which must surely be the case). These undirected or weak textual flows (between witnesses that are likely to be strongly related) are shown with a bi-directional arrow and highlighted in red, as in Figure 85 (where they also show as having rank zero, because they are not true potential ancestors).

This diagram is lengthened, where Figure 84 was flattened. Several relationships on the spine are now shown with this bi-directional red arrow, showing that they could easily be reversed on a small number of changes to local stemmata. It also shows three groups of witnesses whose exact place in the textual flow is uncertain, but the location of the group itself seems unlikely to change, namely {029, 083, 04C1, F1, FΠ, 33}, {821, 0141, 18, 226, 35} and {2411, 1571, 1546, 1128, 1230}. This diagram could even be changed to make these groups explicit, as in Figure 86. This version of the diagram emphasises the fact that, in this textual flow diagram, there are groups of closely related witnesses where the exact flow is uncertain, but these groups are connected to one another (and a few individual witnesses) by stronger textual flows.

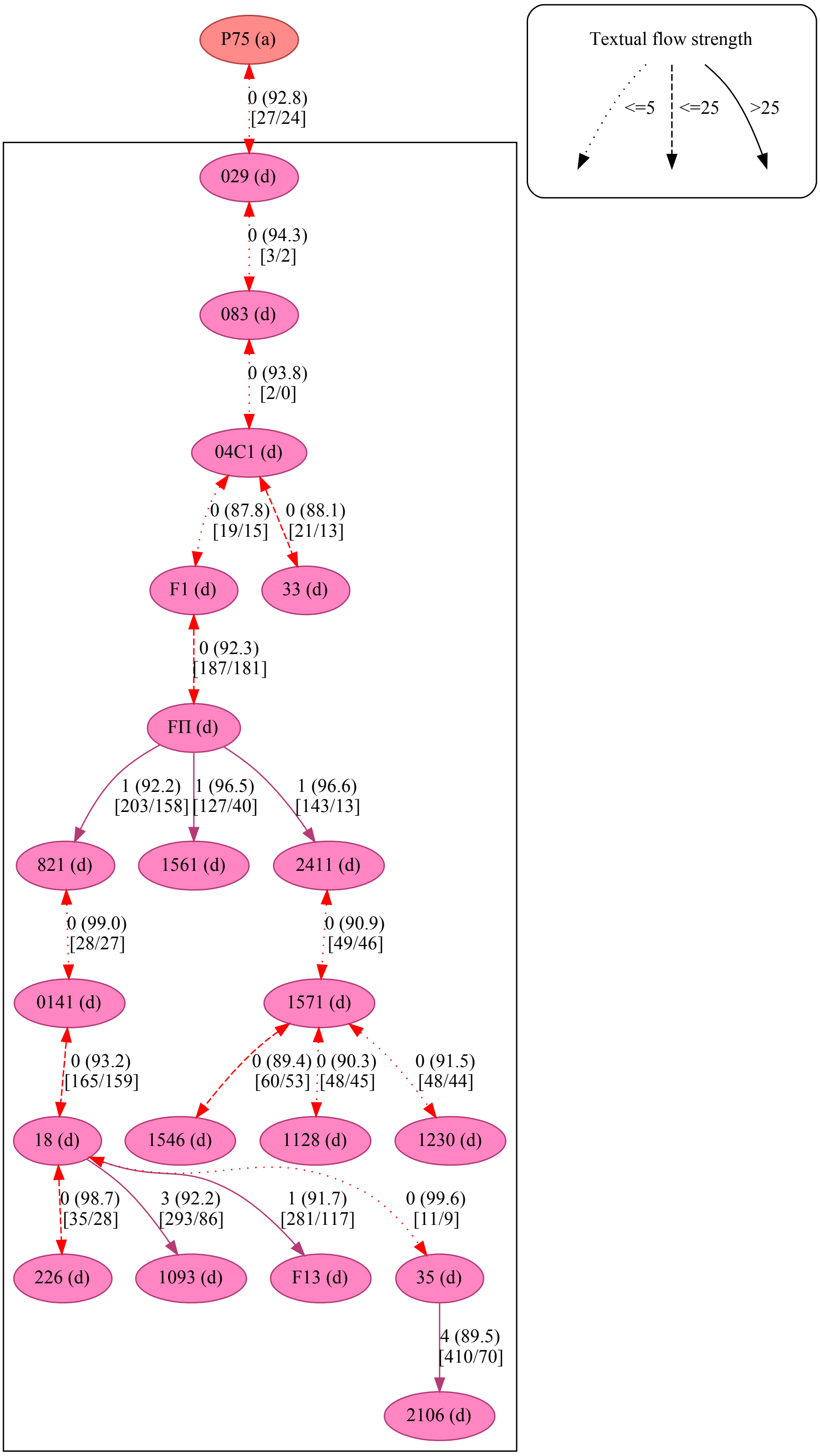


Figure 85: Textual flow diagram for John 1:28/4-8 (reading d) with absolute connectivity; textual flow strength <= 10 treated as undirected; and allowing undirected relationships.

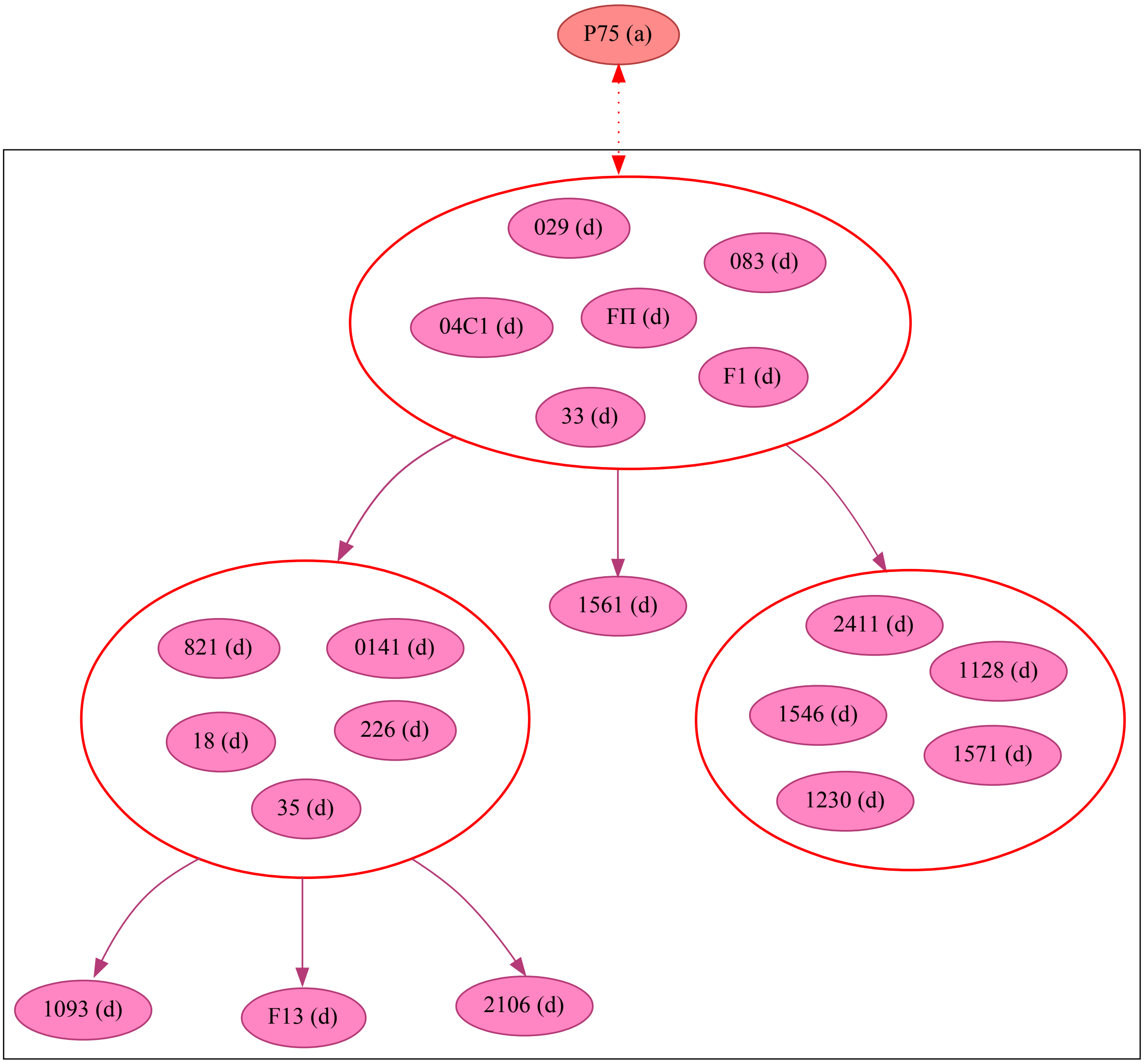


Figure 86: Textual flow diagram for John 1:28/4-8 (reading d) with absolute connectivity; textual flow strength <= 10 treated as undirected; and showing undirected relationships in groups.

For a scholar interested in finding out how the witnesses are genuinely related to one another, these modified diagrams could be of great use (when considered together). They are, however, of little interest to someone whose goal is to define an initial text, where the actual relationships between witnesses are relatively unimportant compared to identifying the abstract flow of the text (and thus its beginning).

Conclusion

The obvious conclusion of this chapter is that the scholar needs to be able to apply different criteria when creating textual flow diagrams, and needs to be aware of their benefits and disadvantages:

1. A minimum strength could be used to create a textual flow that is less susceptible to change and is therefore useful for refining local stemmata.
2. Weak relationships should be highlighted to prevent the scholar accidentally relying on them.
3. Undirected (or weak) textual flows could be allowed, but highlighted (for example with the double-headed red arrow) to create a diagram less likely to contain confusing or misleading ancestral relationships, while being less useful for refining local stemmata.
4. These settings need to be able to be dynamically modified by the user.

Chapter 9: The CBGM: Further Criticisms and Improvements

Introduction

This chapter discusses several further potential problems with the CBGM, or areas where it may be improved. The goal, as for the entire thesis, is to determine whether the CBGM should be considered methodologically sound, or not.

The Formula for Pre-genealogical Coherence

The CBGM calculates pre-genealogical coherence by comparing agreements across all variant units. Consider the simple example data in Table 98. The formula for pre-genealogical coherence is simply the number of variant units where two witnesses agree (EQ) divided by the number where they are both extant (PASS), expressed as a percentage. Table 99 shows the pre-genealogical coherence values for witness α.

Witness	VU1	VU2	VU3	VU4	VU5
α	This is	a simple	example	of the CBGM’s pre-genealogical coherence	formula
β	Here is	a simple	example	of the CBGM’s pre-genealogical coherence	formula
γ	Here is	a simple	demo	of the pre-genealogical coherence	method
δ	Here is	an	example	of the pre-genealogical coherence	formula

Table 98: Example variant units

W2	NR	PERC1	EQ	PASS
β	1	80%	4	5
δ	2	40%	2	5
γ	3	20%	1	5

Table 99: Pre-genealogical coherence for W1=α

The problem is that the pre-genealogical coherence is entirely at the mercy of the collation, and therefore also at the mercy of the extant tradition – yet is hailed as if it were objective.[1] In 'On the Weighing and Counting of Variants: The Coherence Based Genealogical Method, Potential Ancestors, and Statistical Significance', Jongkind gives an example of two variant units that he argues should be counted as only one. Therefore, he concludes, the data is subjective and not solid enough to provide a base for quantitative analysis.[2] The artificial example in Table 98 amplifies the problem, as only a small number of variant units are present and therefore any change to the collation would cause a significant corresponding change to the coherence values. If, for example, VU1 and VU2 were merged into a single variant unit then the pre-genealogical coherence percentage of witnesses α and γ would be reduced to zero.

A simple replacement algorithm for calculating pre-genealogical coherence would be to use the Levenshtein distance, an algorithm devised in 1965 by Vladimir Levenshtein to measure the similarity between two strings.[3] In this case, the pre-genealogical coherence for α would be as shown in Table 100.[4] The crucial point here is that this formula is collation-independent.

W2	NR	PERC1
β	1	94.52%
δ	2	82.70%
γ	3	75.56%

Table 100: Pre-genealogical coherence for W1=α using the Levenshtein distance

These figures are more intuitively correct, as "This is a simple example of the CBGM's pre-genealogical coherence formula" (α) and "Here is a simple demo of the pre-

---

[1] Consider, for example, Aland et al., *ECM IV 2nd ed. (Catholic Letters)*, 35*. "[the CBGM] confronts the editors with the objective data of pre-genealogical coherence"

[2] Jongkind, 'On the Weighing and Counting of Variants', 6.

[3] See Levenshtein, 'Binary Codes Capable of Correcting Deletions, Insertions, and Reversals'; Atallah and Fox, *Algorithms and Theory of Computation Handbook*, 14–4; 14–35. This is the method used in Stripey (see Chapter 4).

[4] I have removed the EQ and PASS columns as they are no longer relevant for calculating the pre-genealogical coherence.

genealogical coherence method" (γ) are much more than 20% similar (agreeing in 7 words out of 10 or 11). It should be noted that the percentages produced by the two algorithms are not comparable – and the Levenshtein method is likely to produce higher values (a value of less than 90% might be rare).

Let us consider a real example based on Parker's ECM data for John, as of 20 June 2017. Table 101 shows the pre-genealogical coherence with W1=019 using the CBGM's algorithm to calculate the percentage, and Table 102 using the Levenshtein distance.

W2	NR	PERC1	EQ	PASS
019C	1	97.486	349	358
A	2	94.991	5575	5869
03	3	93.314	5485	5878
04	4	92.898	2250	2422
33	5	92.459	5370	5808
086	6	92.342	410	444
044	7	92.265	5463	5921
𝔓75	8	92.236	3671	3980
083	9	91.862	587	639
029	10	91.844	777	846
070	11	91.766	769	838
033	12	91.552	3858	4214
892	13	90.746	3128	3447
213	14	90.592	4757	5251
397	15	90.308	5255	5819
0290	16	89.811	476	530
1321	17	89.764	5288	5891
821	18	89.723	5247	5848
017C	19	89.702	331	369
0141	20	89.648	5248	5854
F1	21	89.578	5286	5901
FΠ	22	89.504	5270	5888
597	23	89.480	5282	5903
799	24	89.412	5202	5818
032	25	89.396	3996	4470
054	26	89.370	950	1063
1071	27	89.289	5277	5910
𝔓60	28	89.269	391	438
02	29	89.005	4590	5157
249	30	88.994	3008	3380
01Cca	31	88.982	2221	2496
865	32	88.934	4396	4943
869	33	88.895	1721	1936
18	34	88.887	5239	5894
35	35	88.874	5256	5914
022	36	88.871	3338	3756
157	37	88.834	5227	5884
333	38	88.833	5107	5749
226	39	88.825	5246	5906
𝔓66C*	40	88.793	1965	2213
MT	41	88.700	5228	5894
1320		88.700	5228	5894
317	43	88.603	2690	3036
2786	44	88.592	5234	5908
1546C	45	88.589	427	482
045	46	88.458	5227	5909
118S	47	88.434	1101	1245
028	48	88.364	5217	5904
1561	49	88.254	5207	5900
1571S	50	88.231	3366	3815
07	51	88.200	5210	5907
037	52	88.199	5082	5762
1014	53	88.081	4848	5504
994	54	88.050	4966	5640
07C	55	87.982	571	649
1293	56	87.957	4353	4949
1010	57	87.947	4889	5559
011	58	87.937	4928	5604
𝔓66	59	87.818	4664	5311
𝔓45	60	87.814	454	517
1230	61	87.795	5136	5850
1571	62	87.749	1103	1257
138	63	87.728	4854	5533
2561	64	87.709	4674	5329
992	65	87.703	5178	5904
2615	66	87.701	4628	5277
109C	67	87.621	361	412
357	68	87.603	5116	5840
032S	69	87.586	896	1023
2411	70	87.500	4998	5712
038	71	87.489	5175	5915
1241	72	87.487	5118	5850
1242	73	87.460	5140	5877
807	74	87.432	5127	5864
2575	75	87.404	5024	5748
1463	76	87.253	5127	5876
0211	77	87.226	5135	5887
377	78	87.198	5129	5882
2790	79	87.106	2432	2792
1654	80	87.041	5125	5888
1788	81	86.941	5073	5835
2561C3	82	86.926	1343	1545
430	83	86.904	1805	2077
0233	84	86.880	1980	2279
2768	85	86.878	5118	5891
2680	86	86.833	5124	5901
2223	87	86.759	5111	5891
1319	88	86.655	5091	5875
1582C1	89	86.646	571	659
1128	90	86.627	4949	5713
109	91	86.619	5088	5874
1797C	92	86.589	523	604
037C	93	86.579	329	380
1546	94	86.452	5105	5905
544	95	86.375	5059	5857
063	96	86.321	915	1060
F13	97	86.314	5102	5911
05S	98	86.166	436	506
1797	99	86.086	5061	5879
1344C	100	86.018	283	329
04C1		86.018	283	329
01	102	85.978	5028	5848
1009	103	85.939	5024	5846
295	104	85.923	4950	5761
2766	105	85.869	4454	5187
213S	106	85.855	522	608
892S	107	85.634	1675	1956
892SC	108	85.569	421	492
579	109	85.537	4749	5552
01Ccb2	110	85.465	294	344
1788C	111	85.408	398	466
1424	112	85.368	5029	5891
1344	113	85.360	4268	5000
1093	114	85.337	5011	5872
799C	115	85.249	445	522
732	116	85.127	4911	5769
04C2	117	85.068	564	663
1029	118	85.004	4943	5815
1278C1	119	84.976	526	619
841	120	84.947	4949	5826
892S*	121	84.879	421	496
2766C	122	84.848	364	429
1242C	123	84.564	630	745
226C	124	84.275	343	407
2585	125	84.155	2613	3105
2192	126	84.100	4808	5717
1253	127	83.840	2262	2698
382	128	83.797	3672	4382
2718	129	83.523	4334	5189
168	130	83.423	4489	5381
2192C	131	83.156	780	938
168C	132	82.934	554	668
173	133	82.576	3033	3673
2106	134	82.503	4720	5721
2561C1	135	82.319	568	690
792	136	80.720	4689	5809
05	137	80.484	3827	4755

Table 101: Pre-genealogical coherence with W1=019 (CBGM algorithm)

W2	NR	LEV	EQ	PASS
019C	1	99.761%	349	358
A	2	99.194%	5575	5869
03	3	98.916%	5485	5878
070	4	98.796%	769	838
083	5	98.780%	587	639
086	6	98.770%	410	444
033	7	98.624%	3858	4214
044	8	98.560%	5463	5921
04	9	98.544%	2250	2422
𝔓75	10	98.489%	3671	3980
33	11	98.402%	5370	5808
397	12	98.312%	5255	5819
0290	13	98.233%	476	530
213	14	98.131%	4757	5251
P60	15	98.121%	391	438
029	16	98.106%	777	846
037C	17	98.075%	329	380
054	18	98.058%	950	1063
032	19	98.035%	3996	4470
799	20	98.030%	5202	5818
𝔓66C*	21	98.023%	1965	2213
01Cca	22	97.954%	2221	2496
1321	23	97.931%	5288	5891
02	24	97.926%	4590	5157
118S	25	97.874%	1101	1245
0141	26	97.850%	5248	5854
022	27	97.824%	3338	3756
821	28	97.814%	5247	5848
037	29	97.773%	5082	5762
FΠ	30	97.721%	5270	5888
017C	31	97.720%	331	369
333	32	97.710%	5107	5749
𝔓45	33	97.691%	454	517
1071	34	97.687%	5277	5910
F1	35	97.675%	5286	5901
317	36	97.655%	2690	3036
2561C3	37	97.654%	1343	1545
𝔓66	38	97.652%	4666	5311
04C1	39	97.640%	283	329
038	40	97.630%	5175	5915
05S	41	97.619%	436	506
597	42	97.607%	5282	5903
869	43	97.602%	1721	1936
1546C	44	97.599%	427	482
157	45	97.577%	5227	5884
2561	46	97.548%	4677	5329
1571	47	97.517%	1103	1257
0211	48	97.495%	5135	5887
892SC	49	97.481%	421	492
249	50	97.472%	3008	3380
1582C1	51	97.438%	571	659
892S*	52	97.437%	421	496
1797C	53	97.403%	523	604
18	54	97.402%	5239	5894
35	55	97.392%	5256	5914
892	56	97.386%	3128	3447
1242	57	97.382%	5140	5877
063	58	97.360%	915	1060
892S	59	97.353%	1675	1956
2768	60	97.352%	5118	5891
1230	61	97.346%	5136	5850
045	62	97.335%	5227	5909
799C	63	97.333%	445	522
226	64	97.325%	5246	5906
1241	65	97.290%	5119	5850
028	66	97.283%	5217	5904
07	67	97.282%	5210	5907
1320	68	97.240%	5228	5894
MT	69	97.240%	5228	5894
2786	70	97.227%	5235	5908
1561	71	97.202%	5207	5900
213S	72	97.196%	522	608
1128	73	97.188%	4950	5713
1242C	74	97.182%	630	745
07C	75	97.140%	571	649
1014	76	97.132%	4849	5504
992	77	97.112%	5178	5904
1010	78	97.108%	4889	5559
2411	79	97.103%	4998	5712
807	80	97.076%	5127	5864
011	81	97.042%	4928	5604
1546	82	97.000%	5105	5905
1293	83	96.989%	4353	4949
04C2	84	96.970%	564	663
1424	85	96.944%	5029	5891
377	86	96.925%	5129	5882
01	87	96.918%	5028	5848
2561C1	88	96.916%	568	690
1463	89	96.907%	5127	5876
1654	90	96.899%	5125	5888
F13	91	96.894%	5102	5911
2680	92	96.894%	5124	5901
1788	93	96.891%	5073	5835
430	94	96.874%	1805	2077
032S	95	96.864%	896	1023
109	96	96.852%	5088	5874
2223	97	96.839%	5111	5891
994	98	96.807%	4966	5640
1319	99	96.803%	5091	5875
1797	100	96.783%	5061	5879
1571S	101	96.779%	3366	3815
2615	102	96.750%	4628	5277
1029	103	96.732%	4943	5815
01Ccb2	104	96.729%	294	344
109C	105	96.726%	361	412
357	106	96.718%	5116	5840
841	107	96.715%	4949	5826
0233	108	96.713%	1980	2279
168C	109	96.687%	554	668
2192	110	96.648%	4808	5717
1009	111	96.633%	5024	5846
138	112	96.588%	4854	5533
2766	113	96.584%	4454	5187
732	114	96.554%	4911	5769
544	115	96.488%	5059	5857
168	116	96.466%	4489	5381
2192C	117	96.449%	780	938
1093	118	96.439%	5011	5872
579	119	96.377%	4749	5552
2790	120	96.284%	2432	2792
2766C	121	96.103%	364	429
865	122	96.036%	4396	4943
295	123	96.018%	4950	5761
382	124	95.935%	3672	4382
1253	125	95.842%	2262	2698
1344C	126	95.840%	283	329
2575	127	95.830%	5024	5748
1344	128	95.805%	4268	5000
2106	129	95.587%	4720	5721
2585	130	95.405%	2613	3105
1278C1	131	95.402%	526	619
792	132	95.304%	4689	5809
1788C	133	94.932%	398	466
05	134	94.810%	3827	4755
226C	135	94.660%	343	407
173	136	94.306%	3033	3673
2718	137	90.994%	4334	5189

Table 102: Pre-genealogical coherence with W1=019 (Levenshtein algorithm)

The Formula for Pre-genealogical Coherence

It is perhaps clearer to consider the top rows of each table side by side, as in Table 103 and Table 104 which show the top thirty-five rows of each table.

W2	NR	PERC1	EQ	PASS
019C	1	97.486	349	358
A	2	94.991	5575	5869
03	3	93.314	5485	5878
04	4	92.898	2250	2422
33	5	92.459	5370	5808
086	6	92.342	410	444
044	7	92.265	5463	5921
𝔓75	8	92.236	3671	3980
083	9	91.862	587	639
029	10	91.844	777	846
070	11	91.766	769	838
033	12	91.552	3858	4214
892	13	90.746	3128	3447
213	14	90.592	4757	5251
397	15	90.308	5255	5819
0290	16	89.811	476	530
1321	17	89.764	5288	5891
821	18	89.723	5247	5848
017C	19	89.702	331	369
0141	20	89.648	5248	5854
F1	21	89.578	5286	5901
FΠ	22	89.504	5270	5888
597	23	89.480	5282	5903
799	24	89.412	5202	5818
032	25	89.396	3996	4470
054	26	89.370	950	1063
1071	27	89.289	5277	5910
𝔓60	28	89.269	391	438
02	29	89.005	4590	5157
249	30	88.994	3008	3380
01Cca	31	88.982	2221	2496
865	32	88.934	4396	4943
869	33	88.895	1721	1936
18	34	88.887	5239	5894
35	35	88.874	5256	5914

Table 103: Pre-genealogical coherence with W1=019 (CBGM algorithm) – top 35 rows

W2	NR	LEV	EQ	PASS
019C	1	99.761%	349	358
A	2	99.194%	5575	5869
03	3	98.916%	5485	5878
070	4	98.796%	769	838
083	5	98.780%	587	639
086	6	98.770%	410	444
033	7	98.624%	3858	4214
044	8	98.560%	5463	5921
04	9	98.544%	2250	2422
𝔓75	10	98.489%	3671	3980
33	11	98.402%	5370	5808
397	12	98.312%	5255	5819
0290	13	98.233%	476	530
213	14	98.131%	4757	5251
𝔓60	15	98.121%	391	438
029	16	98.106%	777	846
037C	17	98.075%	329	380
054	18	98.058%	950	1063
032	19	98.035%	3996	4470
799	20	98.030%	5202	5818
𝔓66C*	21	98.023%	1965	2213
01Cca	22	97.954%	2221	2496
1321	23	97.931%	5288	5891
02	24	97.926%	4590	5157
118S	25	97.874%	1101	1245
0141	26	97.850%	5248	5854
022	27	97.824%	3338	3756
821	28	97.814%	5247	5848
037	29	97.773%	5082	5762
FΠ	30	97.721%	5270	5888
017C	31	97.720%	331	369
333	32	97.710%	5107	5749
𝔓45	33	97.691%	454	517
1071	34	97.687%	5277	5910
F1	35	97.675%	5286	5901

Table 104: Pre-genealogical coherence with W1=019 (Levenshtein algorithm) – top 35 rows

While the top three rows show the same witnesses in each table, the rest is considerably different – and so it may be safely assumed that each new collation would yield a different set of standard pre-genealogical coherence data. The Levenshtein distance is an appropriate algorithm for this task and presents a result that has a much better claim to being objective than does the CBGM's own algorithm.

Textual Flow is Backwards

The general textual flow moves from the initial text, through older forms of the text towards the majority text and its variations. This direction is "correct", at least according to the text-critical approach taken by the ECM's editors. One wonders what would happen if instead a majority text priority approach were to be adopted. Even if this general flow is as expected, in microcosm the textual flow is often the opposite of what might be expected.

Two manuscripts are included in the ECM of John that are now known to be exemplar and copy – namely 0141 and 821.[5] It might be expected, therefore, that the CBGM would show this relationship appropriately. However, the CBGM will generally show 0141 as descending from 821. Why is this? 821 is 0141's best potential ancestor and, as is shown in Table 105, they only differ in sixty-nine variant units. 821 is prior in twenty-eight units, and 0141 prior in twenty-seven. It is safe to assume that sometimes the scribe made mistakes (creating new posterior readings) and sometimes corrected mistakes in the exemplar (returning to existing prior readings). If he had made two extra mistakes then the textual flow would be inverted and 0141 would become the ancestor. Any scribe who corrects more mistakes than he makes will create an inverted textual flow – since it is unlikely that any witness will have survived that is more closely related to the manuscript than its copy is, and thus the copy becomes the best potential ancestor for the exemplar.[6]

---

[5] See Chapter 5 and Taylor Farnes, 'Scribal Habits in Selected New Testament Manuscripts'.

[6] Of course, using the concept of correcting and making mistakes assumes that the local stemmata in these passages are "correct" and can be relied upon.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
821	1		98.924	6341	6410	28	27	12	2

Table 105: Best potential ancestor with W1=0141

This is a perfect example of the importance of distinguishing between manuscript and text. The manuscript 821 was copied from 0141. But the text in 821 is (only very slightly) prior to that of its real-world exemplar. The CBGM deals in texts, and is correct to show what it does. But it will mislead those who read “821” and “0141” to mean the manuscripts with those names – as the CBGM is making no direct claim about the manuscripts themselves.[7] The convention to label the texts with manuscript identifiers has caused significant misunderstandings and criticisms of the CBGM. Sadly, no better solution presents itself other than to keep reiterating that the CBGM deals in texts and not manuscripts.

Forced Feedback Loop and Weighing versus Counting

Chapter 8 mentioned the danger of relying on weak textual flow when making decisions about local stemmata – as these decisions will subsequently affect the textual flow. In other words using a weak textual flow as a deciding factor in a troublesome local stemma will serve to strengthen that textual flow. This then presents a stronger textual flow, when in reality the weight of evidence is no different. This kind of forced feedback loop is a significant risk in numerical, iterative methods such as the CBGM. In Chapter 8 the solution proposed is to highlight weak textual flow on textual flow diagrams thus alerting the user that certain flows should not be relied upon for making future textual decisions. It is not just about very weak textual flow, however.

The iterative part of the CBGM is the process of revising the local stemmata in the light of the current state of the genealogical coherence, and so the textual flow. The aim is to

---

[7] Gurry describes Jongkind’s unpublished concerns with such reversed relationships due to contamination in Gurry, *A Critical Examination of the Coherence-Based Genealogical Method*, 151ff.

define a complete set of local stemmata that holds true with reference to the genealogical coherence data (at the time). Imagine a collation with a thousand variant units. The editor is working through the units defining the local stemmata. First the apparently unambiguous local stemmata are defined: say, 700 of them. Now consider an undefined local stemma. Two readings have seemingly equal claim to being prior on text-critical grounds. Each reading is attested by only a handful of witnesses. The textual flow is strongly in favour – say strength fifty or more – of one decision. And so the decision becomes easy. There was no doubt and no worrying weak textual flow. Now imagine that this is repeated twenty more times. At each stage the textual flow has been strong, and there is no problem. But the decisions were made on the basis of that strong flow. Without it they could have gone the other way. Now imagine that they did all go the other way. Instead of a flow of strength seventy (existing fifty plus twenty new local stemmata) we have a strength of thirty – but the textual evidence has not changed.

This example is unrealistic, but it serves the purpose of highlighting a danger: one entry in the CBGM's database does not represent the same weight of textual decision as another. A decision made on strong textual grounds is more valuable than a secondary decision made on the grounds of the sum of all decisions made so far. Perhaps the scholar should define the unambiguous local stemmata, and then keep hold of the genealogical coherence data from this time to refer back to throughout the remainder of the process – as well as the up-to-date data based on all local stemmata at a given time. This would avoid the danger of allowing the feedback loop to over-strengthen a flow that, based on textual evidence alone, should be considered weak.

A second example of a potential problem caused by treating all data points as alike has been highlighted by several scholars, for example Porter and Alexanderson.[8] Their concern is that the CBGM treats every variant unit as having equal weight. Variants that are highly indicative (to use Lachmann's term) or connective (the CBGM's term) should, they argue, carry more weight than a trivial error or co-incidental agreement. This surely opens up the CBGM to the criticism that it counts variants rather than weighing them – in direct contradiction to the well known maxim. But what could it do instead?

Should an editor exclude certain variant units or readings at the collation stage? One might expect the answer "no!", as on the face of it this would distort the data – but of course that is exactly what is done. Simple error or nonsense readings and orthographic readings are regularised away in the early stages of preparing the collation, in order to remove such noise from the data. Who decides where to draw the line? Should *nomina sacra* be included or regularised? Should we keep different spellings of the same place name? What about differences in punctuation? Should ligatures be expanded to their normal characters? What about the final *ν*? The reason to exclude all such features is often purely practical. If one manuscript has accents but another does not, then they cannot be compared directly. The scribe's decision of whether to use *nomina sacra* and which ones to use is one of fashion and scribal preferences, it is thought. Spelling of place names surely follows the spelling used at the time the scribe wrote. These are all quite uncontroversial, and all such features of the text tend to be regularised in preparation for the CBGM. The point of this is to take complex and varied information and convert it into pieces of data that may safely be compared (and counted).

But could this be extended, as Porter, Alexanderson et al. would like? What about excluding all corrections in the use of articles to match contemporaneous practices? Should

---

[8] See Porter, *How We Got the New Testament*, 32; Alexanderson, 'Problems in the New Testament: Old Manuscripts and Papyri, the New Genealogical Method (CBGM) and the Editio Critica Maior (ECM)', 8.

omissions that are clearly attributable to *homeoteleuton* be regularised? Or all unintentional changes? Or instead of ignoring such variant units, perhaps they could be given less weight than a deliberate change? It would be possible, without changing the software, to test the impact of doing such things: variant units that are considered more important could simply be repeated a number of times in the collation. However, if the book has ten thousand variant units, then to define a weight for each one would take a long time, and to be done properly would need to become an iterative process in itself. The question is whether this extra effort yields enough value. Only further research could say for sure, but this would allow the variants to be genuinely weighed for the first time by the CBGM.

In some ways, this idea parallels the proposal above that the coherence data based on the unambiguous local stemmata be retained. Some data should be treated as being more important than others when it comes to making text-critical decisions based on percentages.

Correctors and Textual Flow

The CBGM does not include correctors, other than firsthand correctors which replace the firsthand of a witness. Both my implementation of the CBGM, and Parker's data, do include correctors. There are two options when creating a witness representing a corrector, and they have different benefits and disadvantages.

First, a witness can be created that represents the state of the text as the manuscript left the corrector's hand. The primary argument in favour of doing this is that it represents a real text that existed in history; a text that could have been copied into another manuscript. This witness will be very similar to the uncorrected text – perhaps with over 99% pre-genealogical coherence. Consider a manuscript such as 01, which has been systematically corrected several times over hundreds of years. Each different state of the text may well have been copied, indeed the corrections may well have been performed in preparation for making the copy. This information is lost in the standard CBGM.

Imagine the hypothetical example in Figure 87, where manuscript 2 has been corrected four times, and copies made after three of those corrections. The CBGM, by ignoring correctors, would show the textual flow as in Figure 88 - i.e. greatly simplified. Now, given that the CBGM does not claim to reconstruct the actual copying history of the manuscripts this might not seem like a problem. However, this is statistically unsatisfying, since 2Cc is a better potential ancestor for 4 than 2 is – the text is closer and agrees in more places. Indeed, if we insert a witness very similar to 2 then this new witness could be a better potential ancestor for 4 than 2 is – and so the situation would be made worse. In Figure 88 it appears that 3, 4 and 5 are siblings – where they could be shown in three different generations if the correctors were included.

The second approach with correctors is to create a witness consisting only of the corrections themselves. The character of this witness will be precisely the opposite of the first option, as it will have no readings in common with uncorrected text. In this case the focus is on the parents of the correctors – i.e. where do the corrections come from? I find this a much more satisfying situation, as understanding what a text is corrected towards can shed some light on the scribal context of the time.[9] This is the approach we have adopted in applying the CBGM to Parker's ECM data of John. In my implementation of the CBGM these fragmentary correctors do still appear as parents of real witnesses, as I do not exclude fragmentary witnesses as the CBGM proper would do. There is a key argument against this approach: a fragmentary witness with, say, fifty variant units corresponds to a very low sample size, statistically speaking – and even though a witness may agree with it in 98% of units, it may

---

[9] It is encouraging to note that this is also the approach Gurry advocates. See Gurry, *A Critical Examination of the Coherence-Based Genealogical Method*, 204.

disagree strongly with the uncorrected text in the rest of the manuscript. In other words, the fragmentary corrector witness is also misleading in nature – as it represents a text unrelated to the uncorrected text but rather to that of the exemplar used for correcting.

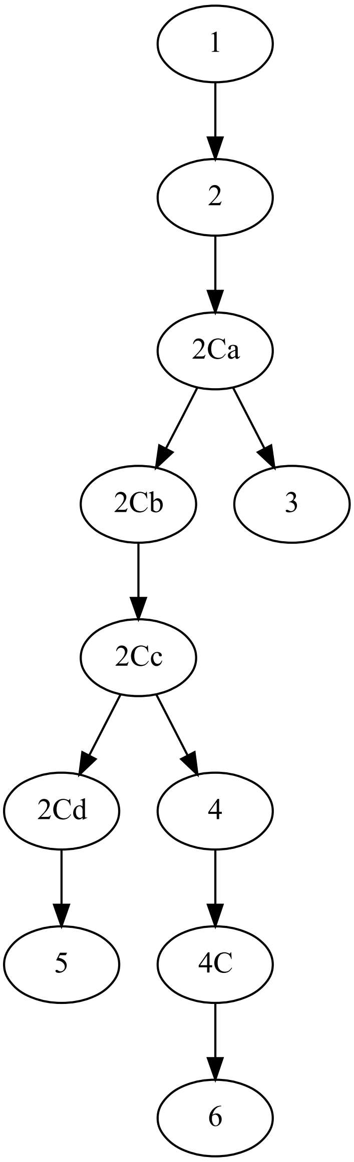


Figure 87: Example textual flow including correctors as full texts

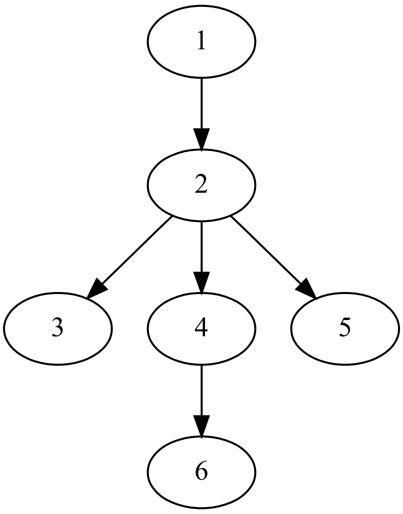


Figure 88: Example textual flow excluding correctors

As an example, imagine manuscript X was corrected against manuscript Y – creating XC. Then XC was used as the exemplar for witness Z. The full text corrector witness (the first option above) is very similar to X, and may provide a better potential ancestor for Z than X itself would. It is highly unlikely, however, that any information will be presented by the CBGM as to the precedence of the corrections. The fragmentary corrector witness (the second option) would, in all likelihood, descend from Y – showing that the corrections come from Y. It is unlikely that XC would descend from X, however (unless over 50% of variant units were corrected).

It is necessary to note that, while including correctors allows for these interesting research questions to be investigated (such as: where do XC's corrections originate?), Parker later decided that correctors should be excluded from textual flow diagrams for ECM John as they tend to add complexity and offer little help when carrying out the CBGM's primary purpose of establishing the initial text. But if, instead, a scholar wishes to explore the many states of the text and research individual manuscripts then the inclusion of correctors is a valuable tool.

Conclusion

In this chapter I have discussed several key issues regarding the CBGM. First, the algorithm used to generate pre-genealogical coherence is not objective and is entirely at the mercy of two external factors: regularisation and collation decisions made by the scholar; and the particular set of manuscripts known about or included in the database. Instead the Levenshtein distance provides an objective value that would not be affected by either of these factors.

Secondly I suggested that the coherence data based on only the unambiguous local stemmata should be retained. It should periodically be used to check that later results are correct and have not been skewed by the potential forced feedback loop that adds strength to weak textual flow and then bases decisions on that flow presuming it to be stronger than it really is.

Thirdly, and in a similar vein, there may be considerable value in modifying the CBGM to allow relative weights to be applied to variant units. However, to do so would add significant extra work to an already labour-intensive method. Further research is required to determine if weighing variants in this way would give sufficient benefit to the results.

Finally, when applying the CBGM the scholar should include all correctors if they are seeking to explore the development of the text over time. However, if the scholar's goal is simply to identify an initial text then they may safely use the texts as they left the firsthand or scriptorium (with exceptions such as 𝔓66* which includes many readings found in no other witness).

Chapter 10: The CBGM and Phylogenetics

Introduction

Earlier in this thesis, Chapter 3 set out a brief comparison of the CBGM's results with those produced by phylogenetic techniques, using a small test dataset. We are now in a position to perform this comparison in more detail using the full data from Parker's work on the ECM of John's Gospel.[1]

Phylogenetic Tree

For *MrBayes*, Parker's collation of John is all that is required. I converted this into a NEXUS file with 138 taxa (witnesses) and 6,506 characters (variant units).[2] I then ran *MrBayes* on BlueBEAR for thirty million generations with two runs, thirty-two chains and temperature 0.1, using the lessons learned in Chapters 5 and 6.

After thirty million generations the average effective sample size (ESS) was 1,858.67, the potential scale reduction factor (PSRF) was 1.019 and the average standard deviation of split frequencies (ASDSF) was 0.03591 (see Figure 89).[3] Together, these values are more than sufficient to say that the runs have converged, even though the PSRF and ASDSF have not quite reached their ideal values. The analysis took a little over 114 hours (nearly five days) and used nearly twenty-four days of CPU time.

---

[1] This analysis was performed in December 2017 using my CBGM tag 1.12, with data from 31 October 2017. The data does not represent the final version that will be used in the ECM, but it is excellent for the purposes of this chapter.

[2] See Part Two for more information about *MrBayes*, NEXUS files, and associated terminology.

[3] See Chapter 6 for more details about MCMC convergence. See also Ronquist and Deans, 'Bayesian Phylogenetics and Its Influence on Insect Systematics'; Gelman and Rubin, 'Inference from Iterative Simulation Using Multiple Sequences'.

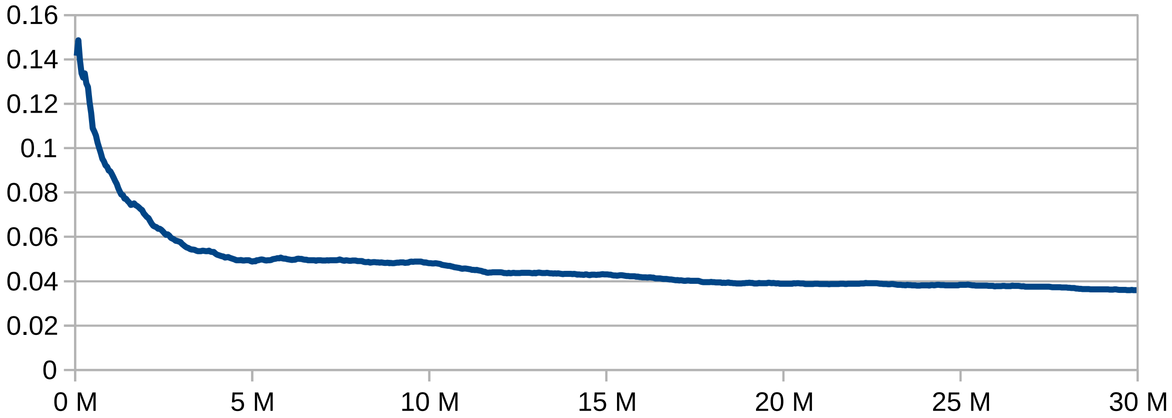


Figure 89: Average standard deviation of split frequencies (ASDSF) by generation

The phylogenetic consensus tree is shown in Figure 90. This is quite hard to read at this scale, but smaller sections of it will be discussed in detail later and shown in much larger print.

Median Joining Network

I converted the NEXUS file to an amino acid Roehl data format (RDF) file, setting the maximum fragmentation level to 15%, and then used Fluxus Engineering's *Network* software to create a median joining network.[4] The resultant network (drawn with the Graphviz NEATO algorithm) is shown in Figure 91. It will be compared to the phylogenetic tree and CBGM results below, with relevant sections of the diagram printed at a readable scale. Note that 1320 and MT have been merged into a single node by the software as they agree in all variant units considered by *Network*. This diagram is very hard to read at this scale, but is printed here to give an overall impression of its layout. A searchable, zoomable electronic copy is available at http://epapers.bham.ac.uk/3144/1/dcp_john_combined_network.neato.pdf

---

[4] The 15% max. frag. level means that any manuscript that is less than 85% extant is excluded from this network. This is necessary as the *Network* software excludes any character that is not found in all taxa – and so the analysis is limited to only those variant units found in the intersection of all witnesses. See Chapter 4 for details of the whole process from NEXUS file to usable SVG files.

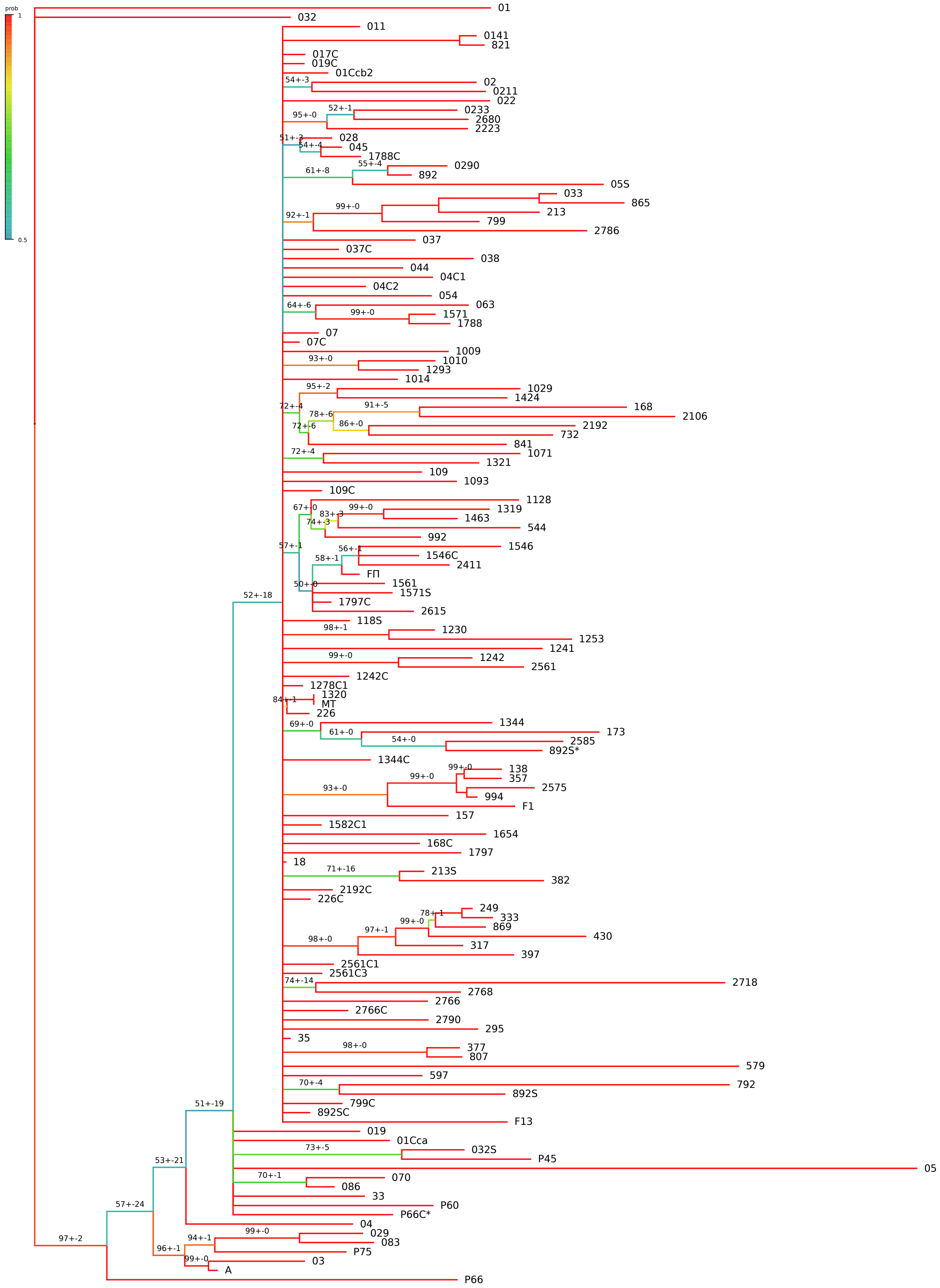


Figure 90: Phylogenetic Consensus Tree for John

Median Joining Network

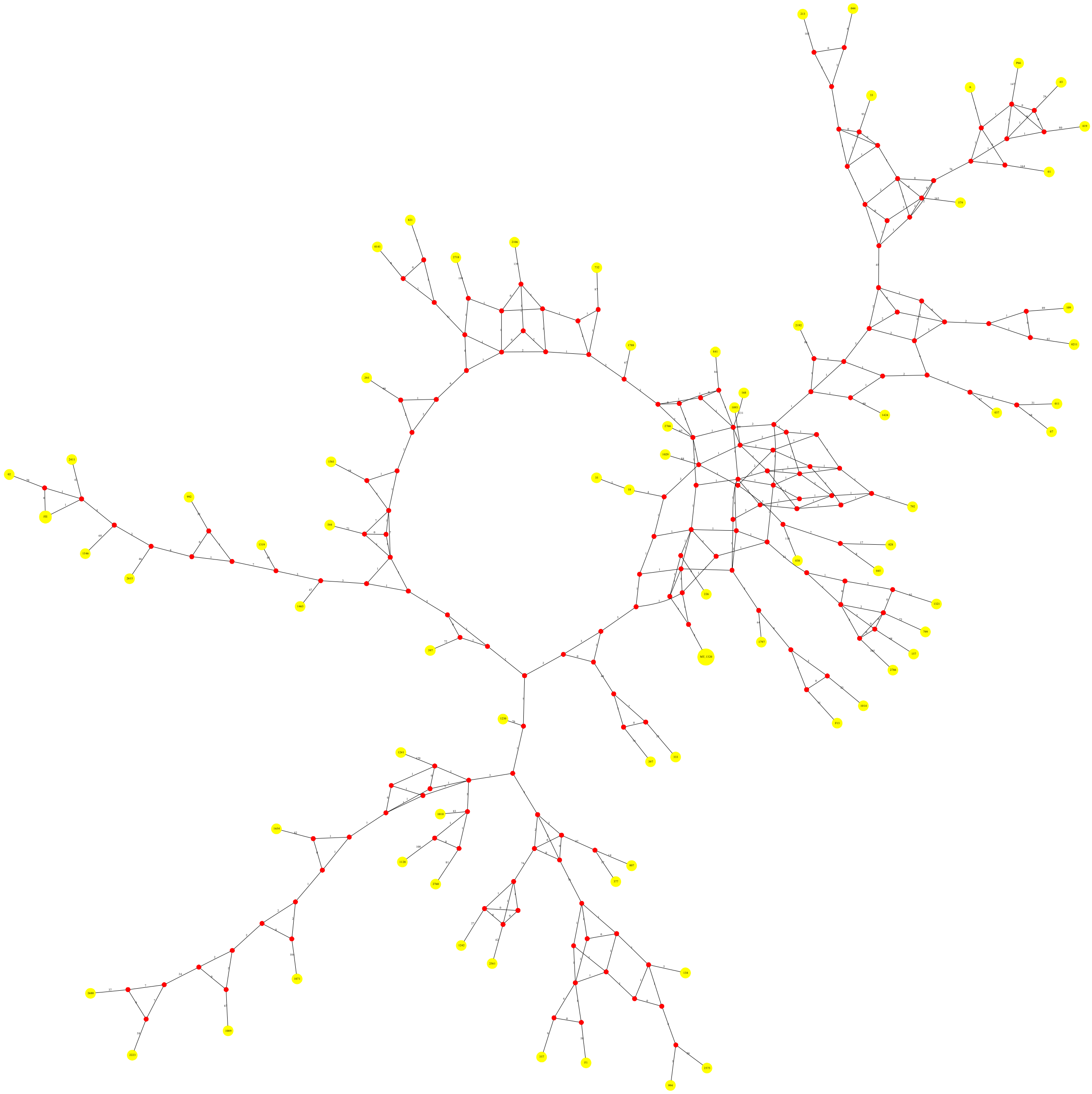


Figure 91: Median Joining Network for John (excluding witnesses <85% extant)

CBGM

Optimal Substemmata

For the CBGM I created lists of combinations of ancestors for each witness and then chose the “best” combination to create optimal substemmata.[5] Recall that the *optimal* substemma for a witness will contain the smallest number of necessary ancestors that explain all its readings (or the most possible) – either by agreement or by the ancestor’s reading being prior (henceforth ‘explained by posteriority’).[6] The CBGM’s table headings for combinations of ancestors are shown in Table 106.

Vorf	Combination of ancestors
Vorfanz	Number of ancestors in the combination
Stellen	Number of variants explained by agreement with an ancestor
Post	Number of variants explained by posterity
Fragl	Number of cases of unknown source of variant (i.e. incomplete local stemma)
Offen	Number of cases not explained by this combination
Hinweis	“Note”. Only used to highlight the combination explaining the most variants in the descendant compared with combinations which are equal in number of ancestors – shown by “<<”.

Table 106: Column header definitions for combinations of ancestors

When presented with a set of combinations of 10,000 rows or more, there are a few tricks that reduce the time required to find the best combination of ancestors. First, only consider those rows with the smallest number in ‘Offen’, i.e. not explained by the combination. Ideally this should be zero, but in this dataset it is commonly a little higher. This is a symptom of the fact that Parker has not yet completed the iterations of creating local

---

[5] Combinations created using CBGM tag 1.12, and the command “cbgm -d ${DB} combanc all --extracols –max-comb-len=10000”. See Chapter 3 for a detailed walkthrough of this process for creating optimal substemmata.

[6] This can be directly prior or a higher in the local stemma. In other words, a grandparent witness can be used, or a great grandparent etc. This is necessary due to so many witnesses being lost.

stemmata, and then considering the textual flow diagrams, coherence data etc. in order to refine the local stemmata. However, the number of unexplained readings is insignificant compared to those explained by agreement or posterity, and so this data is entirely adequate for creating a global stemma for my purposes.

In the case of 011, the data left after removing all rows with non-minimum entries in 'Offen' is shown in Table 107, and indeed there is only one suitable combination of ancestors.

Vorf	Vorfanz	Stellen	Post	Fragl	Offen	Hinweis
18, 2786, 03	3	5519	115	542	4	

Table 107: Combinations of ancestors for 011, with Offen=4

070 is a good example of a witness where the CBGM has been able to determine the best rows automatically (shown by "<<"); they are shown in Table 108. In this case the question is which of the top four rows represents the best solution (the other rows have Offen > 0).

Vorf	Vorfanz	Stellen	Post	Fragl	Offen	Hinweis
019, 33, 03, 𝔓75, 044	5	798	30	81	0	<<
04C1, 086, 019, 03, 𝔓75, 044	6	797	31	81	0	<<
019, 03, 𝔓75, 044	4	796	32	81	0	<<
03, 𝔓75, 044	3	792	36	81	0	<<
A, 019, 33	3	792	35	81	1	<<
A, 019, 044	3	792	35	81	1	<<
019, 33, 03	3	792	35	81	1	<<
A, 33, 044	3	792	35	81	1	<<
019, 03, 044	3	792	35	81	1	<<
33, 03, 044	3	792	35	81	1	<<
019, 𝔓75, 044	3	792	33	81	3	<<
019, 044, 01	3	792	32	81	4	<<

Table 108: Combinations of ancestors for 070 ("<<" rows only)

Combination {019, 33, 03, 𝔓75, 044} explains the most by agreement. {04C1, 086, 019, 03, 𝔓75, 044} is a larger set and so cannot be optimal. {019, 03, 𝔓75, 044} explains two more variant units by posterity, when compared to the first combination: John 11:51/30-34 (addition of an article) and John 9:9/27 (addition of δε). These two changes are clearly acceptable to include by posterity. Then {03, 𝔓75, 044} explains four further variant units by posterity: John 9:9/30 (omission of οτι), John 9:8/27 (addition of οτι), John 9:4/18 (ως as opposed to εως) and John 11:56/18-24 (εστηκοτες εν τω ιερω as opposed to εν τω ιερω εστηκοτες). All these changes are easy to see as multiple emergence, and so accepting the explanation of posterity means that the optimal substemma is {03, 𝔓75, 044}. For the remaining optimal substemma, I will only consider smaller combinations that explain 10 or fewer extra readings by posterity. This is simply a practical restriction.[7]

A more complex example is 0290. Table 109 shows those combinations where ‘Offen’ is zero, and the ‘<<’ marker is in the ‘Hinweis’ column. There are also eleven extra rows representing combinations of size eight (Vorfanz=8) and twenty-five of size nine – all of which explain 510 passages (Stellen) by agreement just as the two size seven combinations do at the top of Table 109. But these need not be considered as they cannot be optimal when compared to the rows of size seven. Within the rows for the same size (and the same ‘Post’), the rows are ordered by the sum of their constituent ranks – i.e. the top row of each block is better.

---

[7] See `analyse_combinations_of_ancestors.py` in https://doi.org/10.5281/zenodo.1296287 or https://github.com/edmondac/CBGM for a script that does this mechanical work, presenting the text critical questions to the user for analysis.

Vorf	Vorfanz	Stellen	Post	Fragl	Offen	Hinweis
2561C3, 0141, 33, A, 1071, 044, 04	7	510	9	42	0	<<
2561C3, 0141, 33, 1071, 044, 03, 04	7	510	9	42	0	<<
2561C3, 0141, 33, A, 1071, 044	6	509	10	42	0	<<
2561C3, 0141, 33, A, 044, 04	6	509	10	42	0	<<
2561C3, 0141, 33, 1071, 044, 03	6	509	10	42	0	<<
2561C3, 0141, 33, 044, 03, 04	6	509	10	42	0	<<
0141, 33, A, 1071, 044, 04	6	509	10	42	0	<<
0141, 33, 1071, 044, 03, 04	6	509	10	42	0	<<
2561C3, 0141, 33, A, 044	5	508	11	42	0	<<
2561C3, 0141, 33, 044, 03	5	508	11	42	0	<<
0141, 33, A, 1071, 044	5	508	11	42	0	<<
0141, 33, 1071, 044, 03	5	508	11	42	0	<<
0141, 33, A, 044	4	506	13	42	0	<<
0141, A, 865, 044	4	506	13	42	0	<<
0141, 33, 044, 03	4	506	13	42	0	<<
0141, 865, 044, 03	4	506	13	42	0	<<
0141, A, 044	3	503	16	42	0	<<
0141, 044, 03	3	503	16	42	0	<<
A, 044	2	490	29	42	0	<<

Table 109: Combinations of ancestors for 0290 (Offen=0, Vorfanz<8 and "<<")

Combination {2561C3, 0141, 33, A, 1071, 044, 04} explains the most readings by agreement of all the rows, and has lower summed rank than the other size seven row which is otherwise just as good. Size six combination {2561C3, 0141, 33, A, 1071, 044} (i.e. removing 04) explains one further variant unit by posteriority, John 18:34/2-4, where 0290 adds an article (απεκριθη ο ιησους) when compared to some of its ancestors, with others of its ancestors reading απεκρινατο ο ιησους. The best size five combination, {2561C3, 0141, 33, A, 044} (removing 1071), additionally explains John 19:34/22-24 by posteriority. Here 0290 reverses two words (ευθυς εξηλθεν) with respect to its ancestors. Removing 2561C3 yields a size four combination {0141, 33, A, 044} which requires a further two variant units to be explained by posteriority. First John 19:31/8-34 where επει παρασκευη ην is moved to the end of a phrase, and secondly John 19:27/10 where we find ιδου instead of ιδε. Now, removing 33 results in

{0141, A, 044}, which requires three more variants to be explained by posteriority: John 19:38/2-6, John 19:38/13 and John 19:39/14-22. The first of these is a simple omitted δε. The second is another added article. The third, however, sees *προς αυτον νυκτος το πρωτον* turn in to *προς τον ιησουν νυκτος το πρωτον*. This change has a considerably better claim to being connective than the others, and we will treat it as such.[8] Therefore this combination is rejected and the optimal substemmata is as shown in Figure 92.

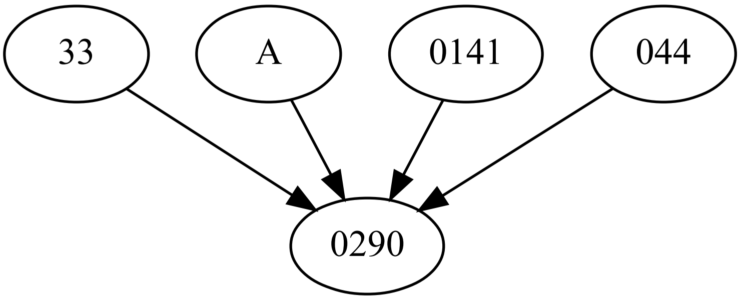


Figure 92: Optimal substemma for 0290

This example highlights an unexpected pattern, which is that as witnesses from the combinations of ancestors are removed the variant units that emerge as needing to be explained by posteriority *very often cluster together* – and indeed here they are all between John 18:34 and 19:39 and not spread throughout the Gospel. To confirm this pattern, and account for it would require further research.

For notes on all the decisions made creating the optimal substemmata for this chapter, see Appendix 6.

Global Stemma

The global stemma can now be created, and is shown in Figure 93 (left) and Figure 94 (right). This is, again, difficult to see due to its complexity and printed scale. A searchable, zoomable electronic copy is available at http://epapers.bham.ac.uk/3145/1/global_stemma.pdf

---

[8] This is actually rather less certain than I suggest here. But it is a reasonable argument, and convenient for the purpose of highlighting the effect of a connective variant in this process.

CBGM

An interesting diagram can be created by removing the initial text (A) from the global stemma. What remains is not a true global stemma, by the CBGM’s definition, but it shows the global textual flow of extant witnesses – and is therefore of significant interest (and is simpler and so easier to read). This is shown in Figure 95. Again, a searchable, zoomable electronic copy is available at http://epapers.bham.ac.uk/3146/1/global_stemma_no_A.pdf

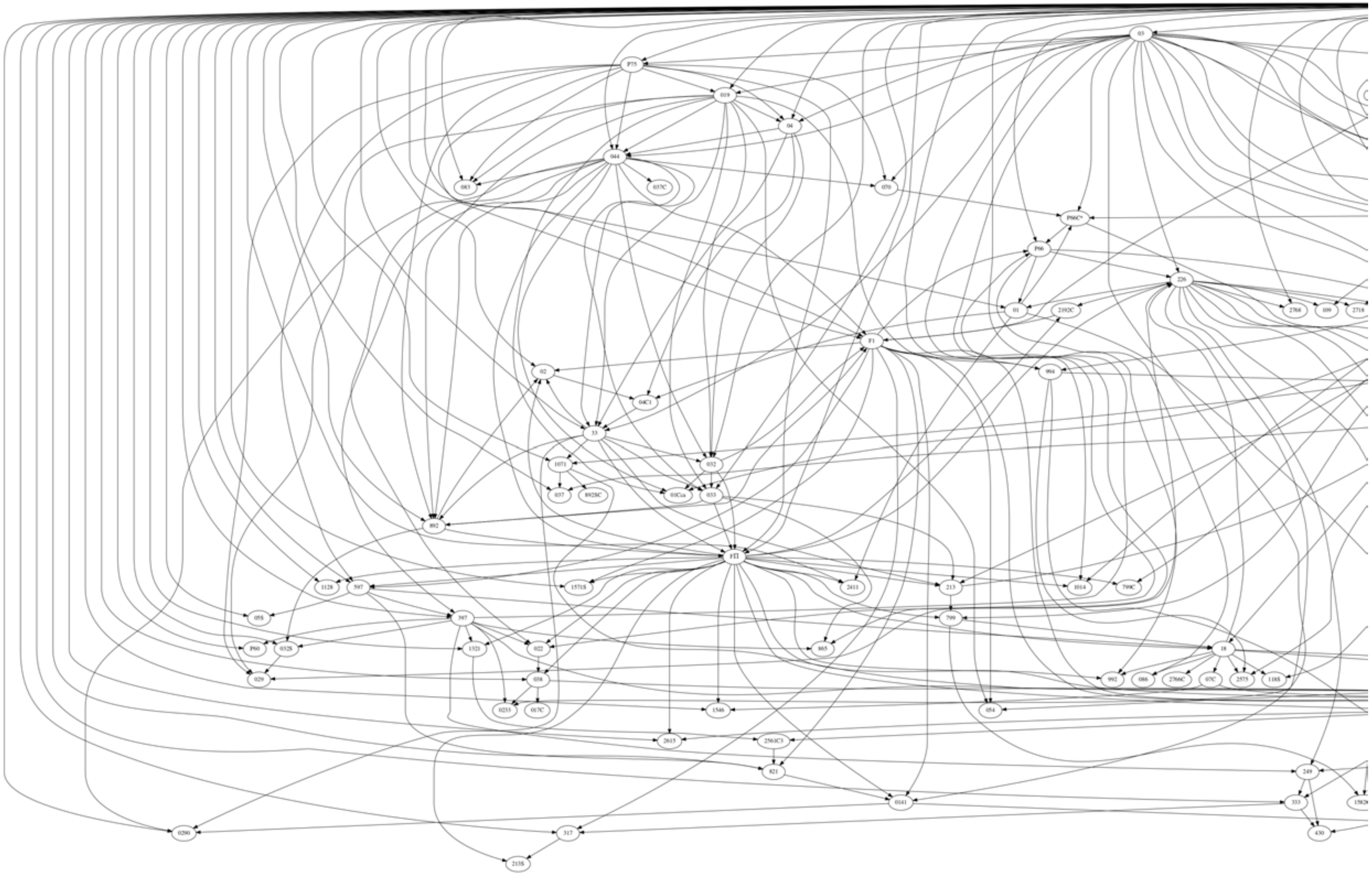


Figure 93: Global stemma for John (left)

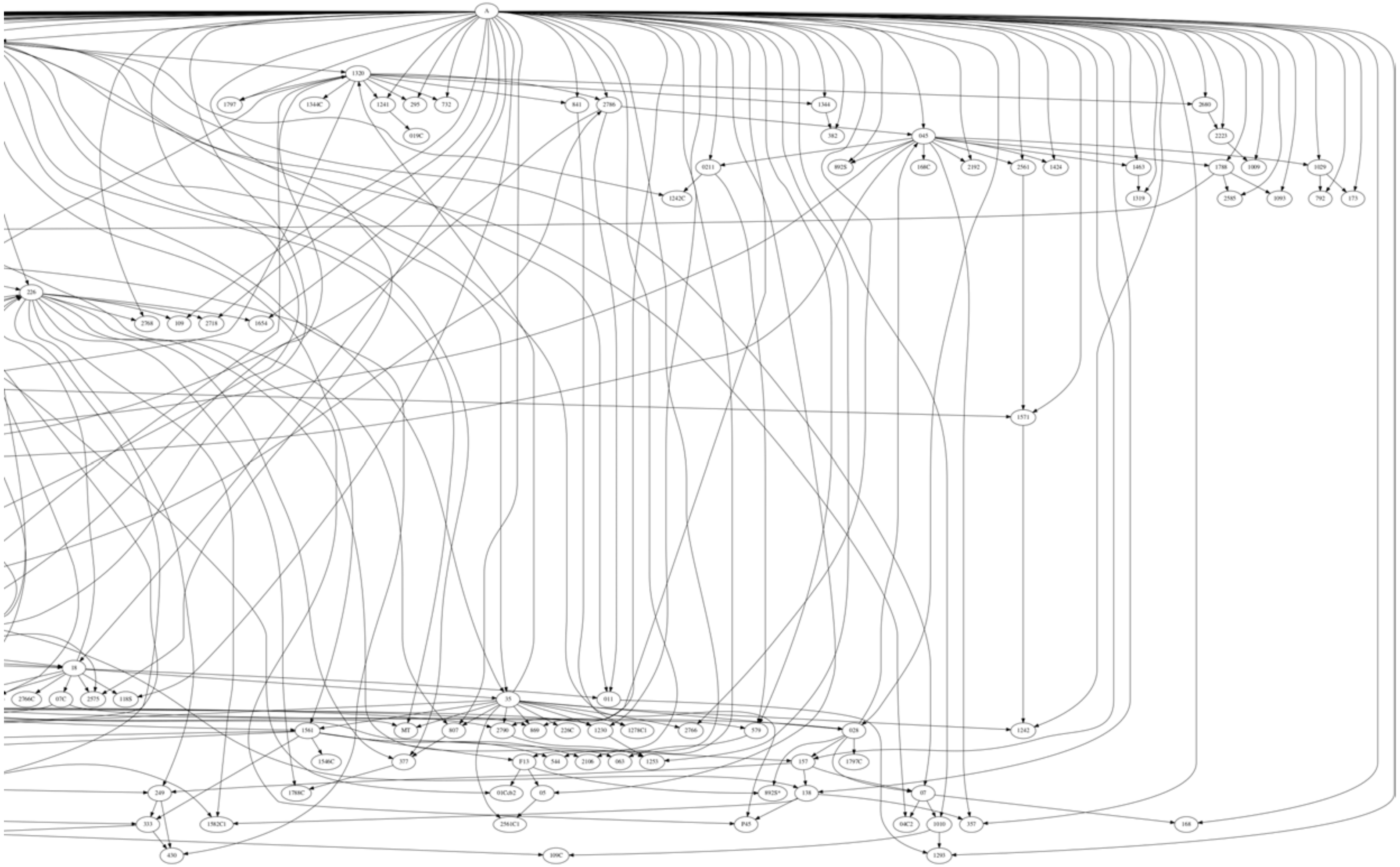


Figure 94: Global stemma for John (right)

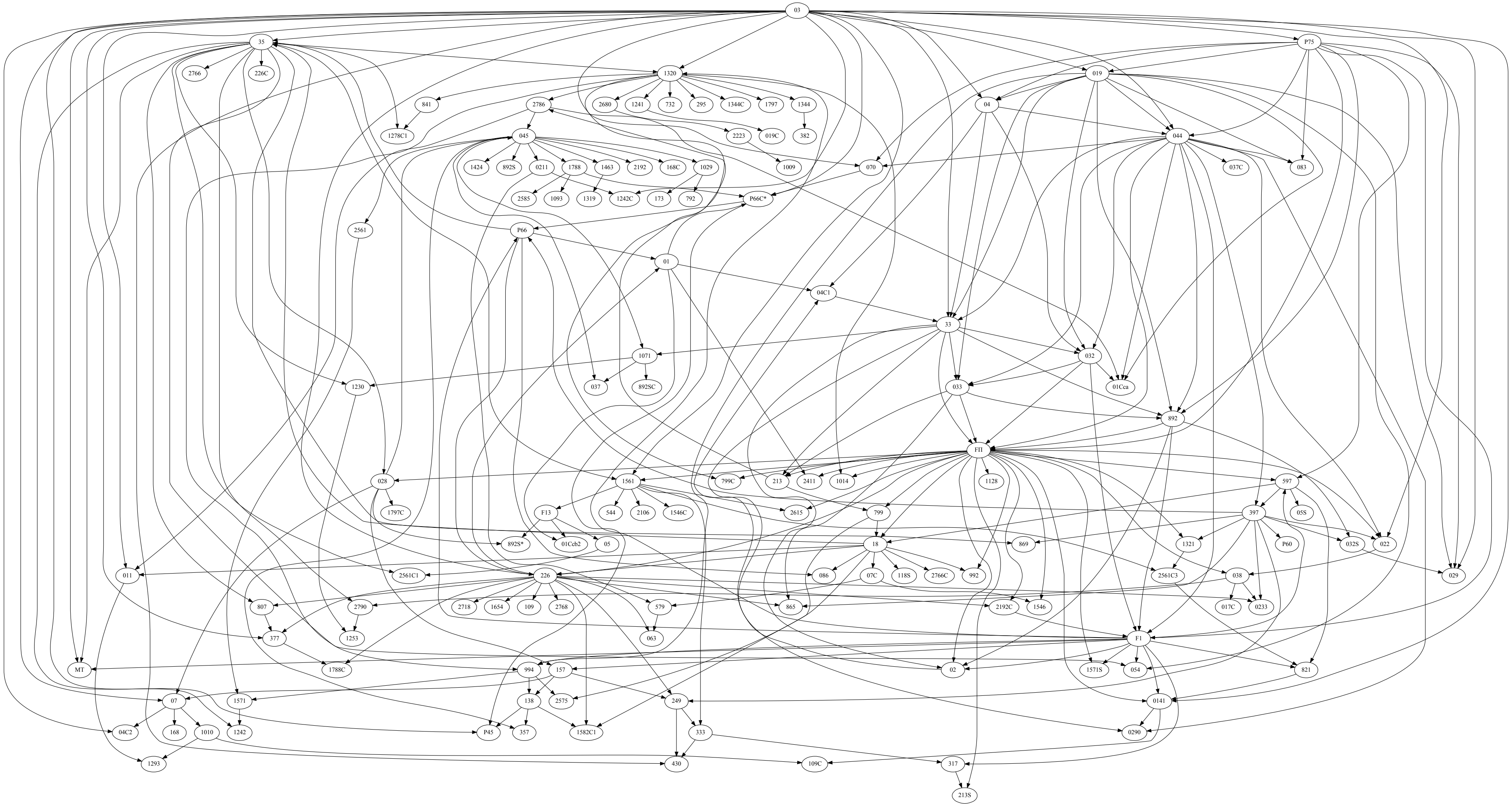


Figure 95: "Global stemma" for John, with the initial text (A) removed.

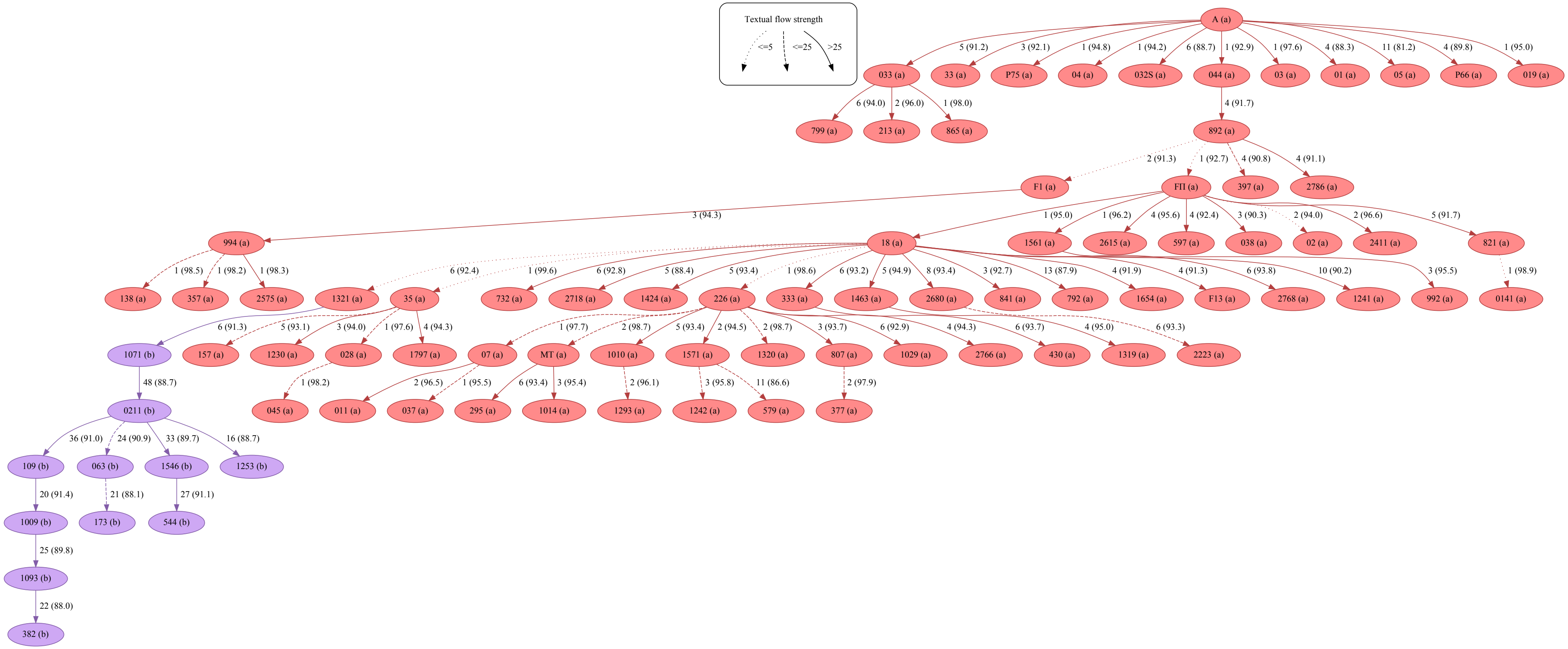


Figure 96: Textual Flow diagram for John 1:8/16 with absolute connectivity

Textual Flow Diagrams versus Global Stemmata

A note is warranted here about a frequent misuse of textual flow diagrams. Textual flow diagrams are often used in place of a global stemma, to try to identify a generic, or global, textual flow. There are, in my opinion, three reasons for this: first and most commonly, textual flow diagrams are mistakenly believed to show this information; secondly, textual flow diagrams are much simpler to interpret; and thirdly, global stemmata require so much work to create that they are frequently skipped and textual flow diagrams are substituted into their place. Consider the textual flow diagram in Figure 96 above.[9] There is an easy-to-understand flow from the top to the bottom making the diagram accessible to the reader. A problem occurs, however, when the reading *b* sub-tree is considered. It is not particularly coherent (relatively low ranks and coherence percentages), and these witnesses should not be considered as a particularly coherent sub-group in the global flow. Indeed, they are spread throughout the global stemma. They are only grouped because they all share a non-connective reading (μαρτυρησει as opposed to μαρτυρηση). These witnesses cannot, due to the nature of a textual flow diagram, have descendants attesting reading *a*. Thus the textual flow diagram cannot show even an approximation of the global textual flow for these witnesses.

It is possible, however, to gain a scholarly understanding of the prevailing textual flow by considering many textual flow diagrams. In this case the scholar will come to recognise common features such as the oft-present spine of A → 044 → 892 → FΠ → 18 → 226 → … . This spine is present in exactly this form in the global stemma – but it is one of many such routes through the graph. Textual flow diagrams are created by iterating through the witnesses and for each one taking the best parent witness that explains an attestation at this

---

[9] While this diagram is considerably easier to read than the global stemmata above, the text is still a little small. I have therefore also made a searchable, zoomable electronic copy available for this diagram at http://epapers.bham.ac.uk/3147/1/textual_flow_B04K01V08_16_c499.pdf

particular place in the text and thus forms a witness-to-witness relationship. A global stemma shows the combination of optimal substemmata – i.e. the relationships required to explain a witness in all variant units in the text. A global stemma is complex, and any attempt to simplify it will lose information. This is another point along the lines of “diagrams can be misinterpreted because they are intuitive”. Once again, the solution here is to make sure that the reader understands what a textual flow diagram is conveying, and what it can *and cannot* show.

Comparing the Diagrams

0141 and 821

It is now time to compare *MrBayes*’ phylogenetic tree, Network’s median joining network and the CBGM’s global stemma (and textual flow diagrams). Let us begin with perhaps the closest relationship in the data, 821 and its exemplar 0141.[10]

The *MrBayes* phylogenetic tree and the median joining network lend themselves to being cropped – and Figure 97 and Figure 98 show the sections of those diagrams containing 0141 and 821. The global stemma has so many crossing lines that the relevant subset of the diagram is hard to extract as an image. But, thankfully, the global stemma exists as a Graphviz DOT file, and it is easy to extract all the lines of that file that refer to one or other of 0141 and 821. This can then be drawn using the same method as textual flow diagrams, resulting in the subset of the global stemma showing 0141, 821 and their direct parents and children (Figure 99).

---

[10] See Chapter 5, and Taylor Farnes, ‘Scribal Habits in Selected New Testament Manuscripts’.

Figure 97: Section of the phylogenetic tree containing 0141 and 821

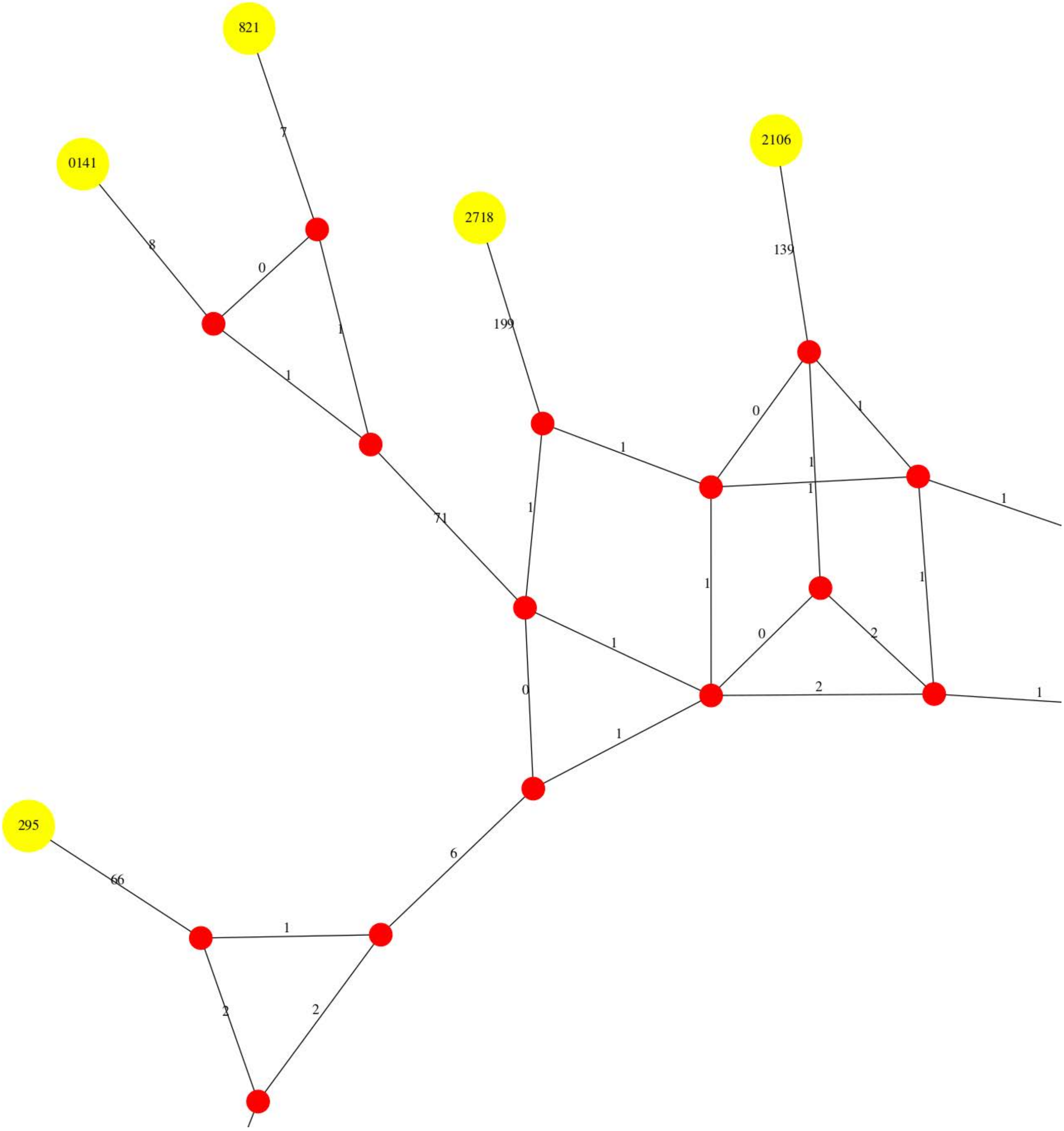


Figure 98: Section of the median joining network containing 0141 and 821

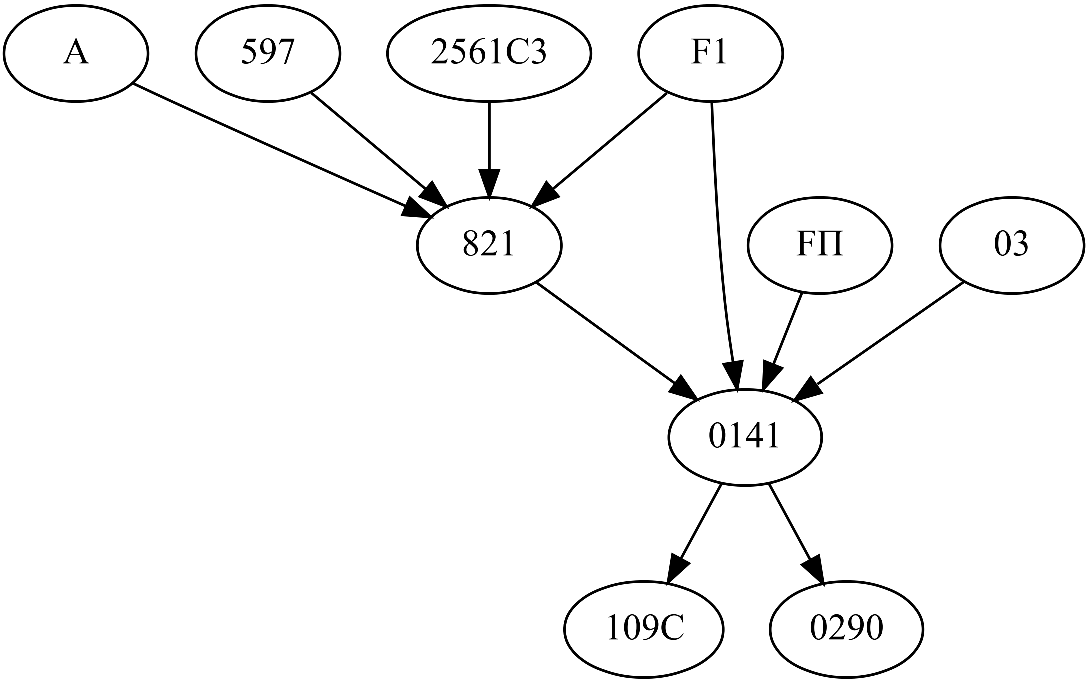


Figure 99: Subset of the global stemma showing 0141 and 821

In all three diagrams 821 and 0141 are directly related. In the phylogenetic tree they are close siblings at the end of a long branch (the nearest to a parent-child relationship that a phylogenetic tree can show). In the median joining network they are neighbours on branches of length seven (821) and eight (0141) joined to the rest of the network by two branches of length one and then a single branch of length seventy-one. Again this indicates a very close relationship and a reasonable distance from the rest of the corpus. A median joining network *can* show witnesses as parent and child – but it has not done so in this case. The subset of the global stemma shows a direct (and directed) relationship from 821 to 0141, following the usual pattern of the CBGM showing a copy as the parent of an exemplar. Of course it is actually showing that the *text* of the copy is closer to the ancestral text than that of the exemplar – which is normal if a copy has corrected some scribal errors in the exemplar but otherwise been very faithful to the exemplar's text.[11]

None of the diagrams agree which witnesses should immediately surround 0141 and

---

[11] See Chapter 9 for more information about these "backwards" textual flows.

821, but of course in the phylogenetic tree and median joining network proximity on the page is not important. Instead, left-right distance required to traverse a route from one witness to the other matters in a phylogenetic network (plus the geometry of the node hierarchy); and the sum of the branch weights (and again the geometry) matters for a median joining network. Let us look for an example with more witnesses.

The “Non-Majority” Witnesses

We will investigate a collection of “non-majority text” witnesses that appear to form a group in the diagrams. Whether or not it can be truly called a group in the text-critical sense is yet to be determined.

Phylogenetic Tree

The phylogenetic tree contains, broadly speaking, a large majority-text branch in the middle of the diagram. If that branch is removed we are left with Figure 100. Of course this is not truly the set of witnesses that do not attest the majority text, hence the quote marks. Instead we are using *MrBayes* as a starting point, and seeing if the sub-set of witnesses seen when that large branch is removed is a meaningful set that the other tools also find in the data. The nineteen witnesses are: A, 𝔓45, 𝔓60, 𝔓66, 𝔓66C*, 𝔓75, 01, 01Cca, 03, 04, 05, 019, 029, 032, 032S, 070, 083, 086 and 33. Both 𝔓66 and 𝔓66C* are included in this data, contrary to the CBGM’s standard practice of using the text as it left the firsthand (or Parker’s practice of when it left the scriptorium), because the firsthand text of 𝔓66 contains many readings that are not found in any other witness.

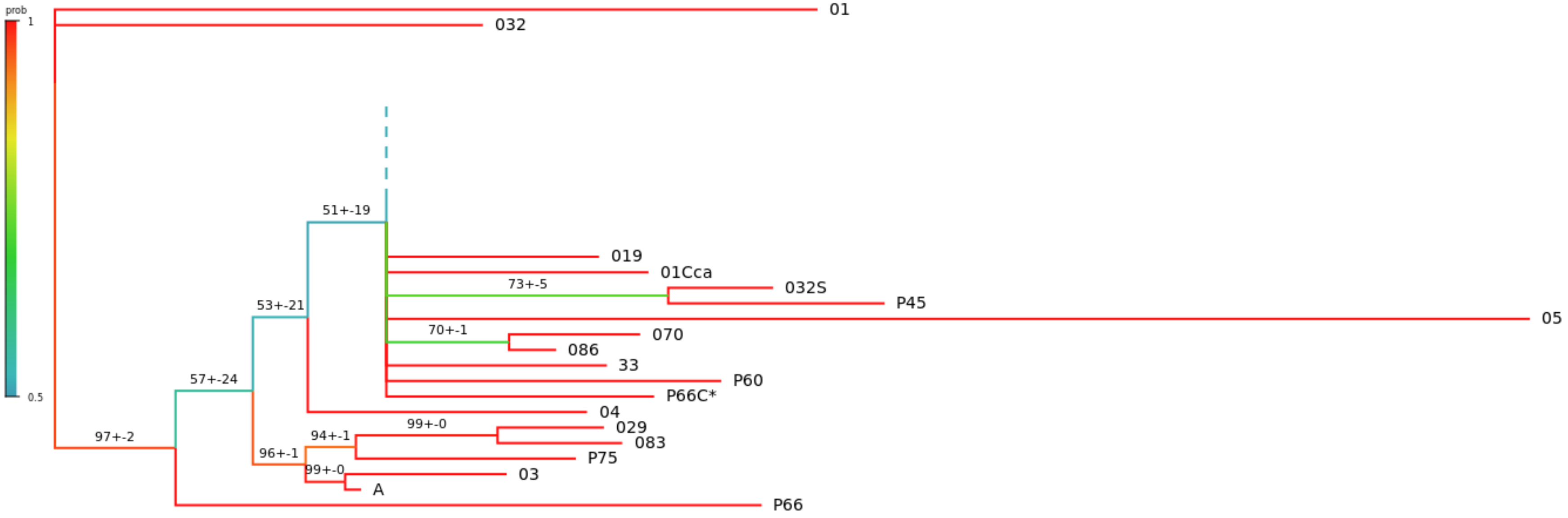


Figure 100: “Non-majority” witnesses in the phylogenetic tree

Comparing the Diagrams

Median Joining Network

The median joining network excludes 𝔓45, 𝔓60, 𝔓66C*, 𝔓75, 01Cca, 04, 05, 029, 032, 032S, 070, 083 and 086 as they are too fragmentary, leaving just A, 𝔓66, 01, 03, 019 and 33. These six witnesses appear in a distinct branch of the network, as shown in Figure 101. Here A, 𝔓66, 01, 03 and 019 are a clear group separated from the rest of the network by a branch of length 76 (henceforth "weight", as length is confusing on a diagram where the lengths do not equate to the branch weights), albeit a group where (as expected) 𝔓66 and 01 are on the end of heavyweight branches. 33 is this group's second-closest neighbour, visually speaking, (after 579) although a few other witnesses are closer when considering the sum of intermediate branch weights but traversing more median vectors (red, effectively hypothetical nodes). 044 is the closest neighbour considering branch weights. Note the branch at the bottom of Figure 101 connecting this section to the rest of the network is of weight forty-five – identifying this subset as a theoretically valid group. 213 and 579 are attached by heavy branches (weights 105 and 163 respectively) and so the core of the network is closer to this group than are these witnesses – and thus they need not be considered as potential group members.[12] Note that 33 is only moderately closer to the group than is 213, consider branch weights. There must therefore be some doubt as to 33's membership – although this doubt will be removed when considering the global stemma below.

---

[12] A similar argument would be possible for 01 and 𝔓66, which are both on heavy branches (164 and 147 respectively). However, the additional 76-weight branch connecting their sub-network to the rest of the diagram acts in their favour and they are safe to consider as group members.

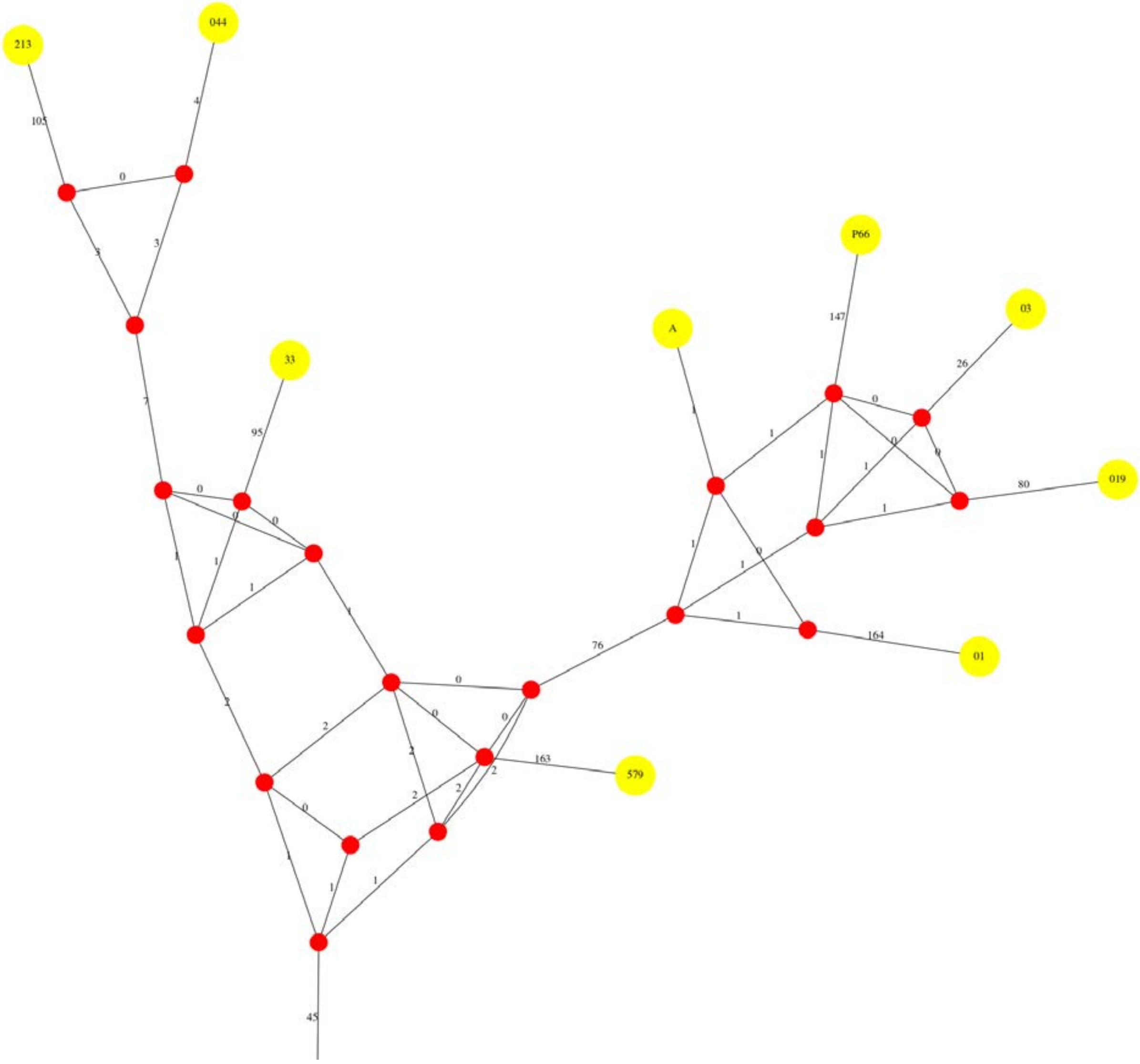


Figure 101: Top right branch of median joining network

Comparing the Diagrams

CBGM Global Stemma

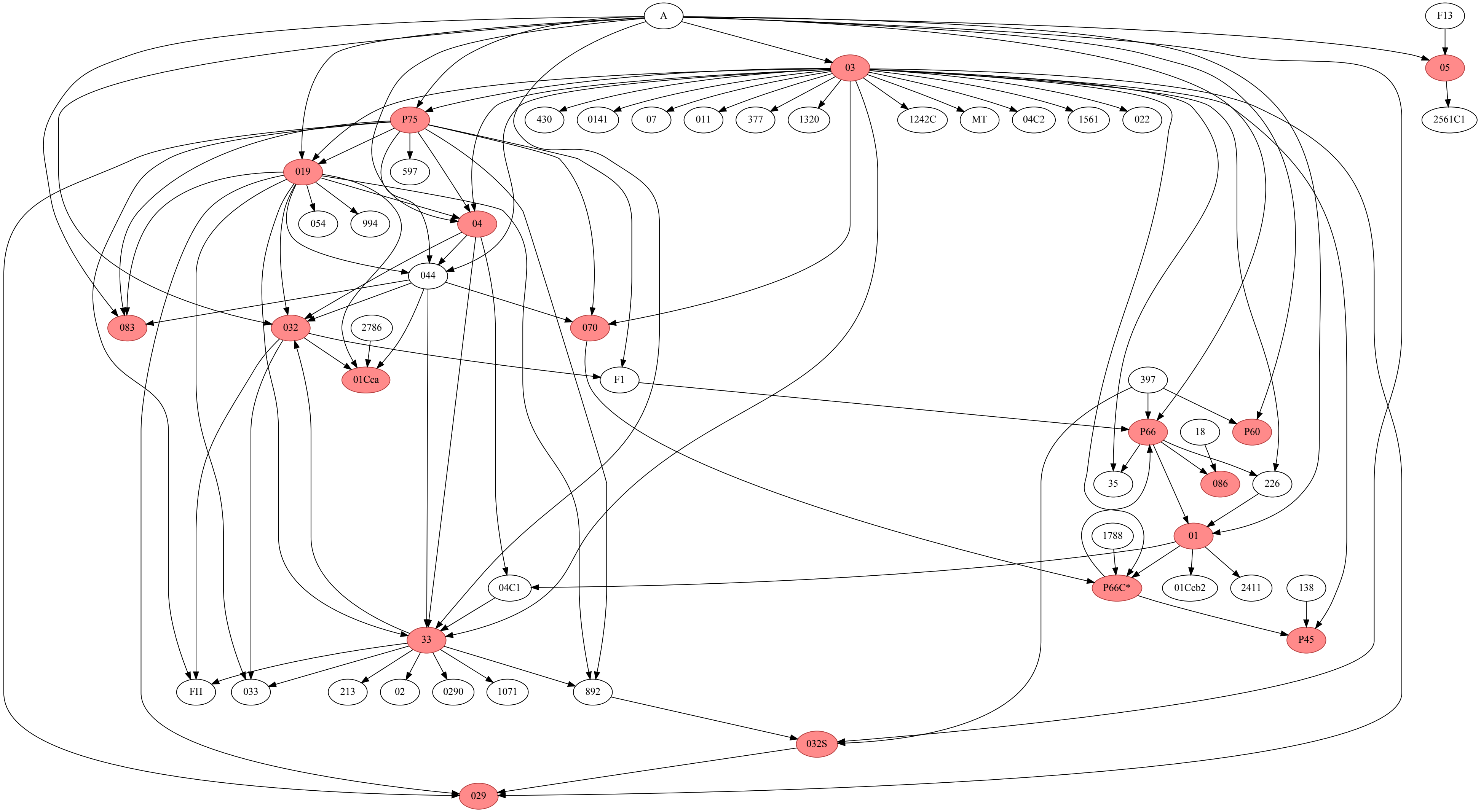
Due to the way the global stemma is drawn, subsets are difficult to identify. So instead of looking for such a subset, we will re-draw the global stemma including only those eighteen witnesses (the nineteen above excluding A) and their immediate parents or children.[13] This global stemma subset is shown in Figure 102 and the eighteen witnesses are highlighted, to make interpreting the diagram easier.

Viewing the subset of the global stemma this way it becomes clear that the eighteen witnesses do indeed form an interconnected block – with a few extra candidate members. Indeed, "viewing" is the key here: an aim of this section is to make discoveries about the data (Parker's collation of John's Gospel) by exploring the three different diagrams created using three different techniques. Note that 33 is directly linked to four other current group members (from 03, 04 and 019 and to 032), confirming its membership when compared to, say, 213 (linked only to 33) which was rejected above. Should any of the related witnesses be considered for group membership? Clearly no witness that is only attached by a single line to one of our eighteen need be considered. Similarly, witnesses that are only descended from our group can be safely ignored.[14] There are, therefore, six witnesses to consider: 044, 226, 892, F1, 04C1 and 397.

---

[13] If all of A's children are included then the diagram once again becomes too complex. Instead, links are shown for A just as every other witness – i.e. only where it is directly connected to one of the eighteen witnesses.

[14] We are looking for strong candidate members, and not just manuscripts that could be outliers in the group. Adding new members at the edge, if repeated, would eventually consume the entire global stemma into the group.


A
F13
03
05
P75
430
0141
07
011
377
1320
1242C
MT
04C2
1561
022
2561C1
019
597
054
994
04
044
083
032
2786
070
01Cca
F1
397
P66
18
P60
35
086
226
1788
01
04C1
P66C*
01Ccb2
2411
138
33
P45
FΠ
033
213
02
0290
1071
892
032S
029

Comparing the Diagrams

Candidate members

Even in the complete global stemma, 044 has only five parents: A, 𝔓75, 03, 04 and 019 – all of which are in our group. It has only one extra potential ancestor (which was not required in its optimal substemmata) as shown in Table 110. It has several children not in this group, along with the included 070, 083 and 33. 044 has a good claim to group membership based on the CBGM evidence. 044 has already been proposed as a potential group member as it is the closest witness (considering branch weights) to the group in the median joining network. In *MrBayes*' phylogenetic network, 044 is found in the "majority" branch, but not particularly close to any other witness and so no strong counter argument can be made that it should *not* be considered for group membership based on the phylogenetic tree. Therefore, from the three diagrams there are two positive votes and one abstention, and so 044 will be added to the group.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
019C		-	97.333	73	75			2	
A	1		92.918	5563	5987	355		42	27
019	2		91.182	5832	6396	239	127	170	28
04	3		90.253	2463	2729	113	68	74	11
03	4		89.983	5839	6489	325	57	236	32
𝔓75	5		89.500	3921	4381	173	110	161	16
029	6		88.653	836	943	37	29	39	2

Table 110: Potential ancestors with W1=044

In the complete global stemma, 226 has more parents outside this group than within it – including 18, 35 and FΠ, all core majority text witnesses. It only has one child in our group, 01, and this is probably due to 01's peculiar textual nature. 226 need not be considered any further.

892 is included for consideration because it has three parents in the group (𝔓75, 019

and 33) and it is a parent of 032S. 032S is a sizeable supplement, containing John 1:1 to 5:11, and so 892 cannot be brushed aside. 892 has only three additional parents in the complete global stemma: A, 033 and 044. 044 has now joined the group, and A is in the original nineteen witnesses (if not the eighteen). 892's closest potential ancestors with significant passages in common are 044, A, 33 and 04 (as shown in Table 111) all of which are group members.[15] It is even more closely related to 086 (a group member) but the relationship is undirected. It is, therefore, clearly a close relative of the group if not a member. 892 has only three further links in the global stemma: 02, F1 and FΠ are descended from 892. In other words, it has more connections (in the global stemma) within the group than outside (six connections against four) and all but one of its ancestors are group members. It seems reasonable therefore to include 892 in the group.

---

[15] There are a number of potential ancestors in the table with small numbers of passages in common ("PASS"), but their high rank could be coincidental due to the particular passages shared with 892. It is safer to consider those potential ancestors with hundreds or thousands of passages in common.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
1797C		-	100.000	67	67				
07C		-	100.000	56	56				
017C		-	100.000	31	31				
892SC		-	100.000	1	1				
892S*		-	100.000	1	1				
2192C		-	96.190	101	105	1	1	1	1
037C	1		93.182	41	44	1		1	1
019C	2		92.593	25	27	1		1	
086		-	92.077	430	467	17	17	3	
01Ccb2	3		91.837	45	49	1			3
044	4		91.660	3451	3765	143	68	83	20
A	5		90.814	3134	3451	272		30	15
33	6		90.655	3347	3692	140	105	83	17
04	7		90.300	1443	1598	75	36	32	12
033	8		90.269	1911	2117	75	62	55	14
083	9		90.015	604	671	32	26	7	2
019	10		89.559	3371	3764	192	71	110	20
029	11		89.290	842	943	50	22	27	2
𝔓66C*		-	88.976	339	381	11	11	15	5
03	12		88.472	3323	3756	244	30	142	17
𝔓75	13		88.113	3054	3466	194	84	115	19
070	14		87.630	758	865	43	29	30	5
032	15		86.558	2228	2574	116	101	107	22
04C1		-	82.000	41	50	3	3	3	
0290		-	81.818	9	11			2	
213S		-	75.000	9	12			3	
05S		-	75.000	9	12			3	
𝔓60		-	66.667	8	12			4	
892S		-	56.250	9	16			7	
118S		-	41.667	5	12			7	
054		-	41.667	5	12			7	

Table 111: Potential ancestors with W1=892

F1 has six parents: A, 𝔓75, 032, 044, 892 and 2192C. Of those, A, 𝔓75 and 032 are in our original group; and 044 and 892 have now joined it. 2192C is the intermediate step to F1 from both 226 and FΠ. F1 is only being considered for membership as it is a parent of 𝔓66 (a group member). However, like the argument for 226 and 01 above, 𝔓66's unusual textual nature presumably accounts for the need for F1 as a parent. There is no strong reason to consider F1 further, rather it is a link between this group and more "majority text" witnesses.

04C1 sits as an intermediary in the diagram between 04 and 01 and 33. In the full diagram it has only one other relationship, namely that 02 is its parent. So all but one of 04C1's links are with members of the group – and thus it seems logical to consider it further. 04C1 is in a very similar position to 044 in the phylogenetic tree as they are siblings with similar length branches and therefore the same arguments must apply – i.e. no strong evidence either for or against membership.[16] 04C1 is too fragmentary to be included in the median joining network, and so we are left with the CBGM evidence alone. Table 112 shows 04C1's potential ancestors with current group members highlighted in grey. It is clear that 04C1 certainly is no closer to the group than to witnesses outside it – indeed the frequency of group members increases as we descend the table – indicating that 04C1 has more close relatives outside the group than within. It is therefore excluded.

W2	**NR**	**D**	**PERC1**	**EQ**	**PASS**	**W1<W2**	**W1>W2**	**UNCL**	**NOREL**
168C		-	100.000	13	13				
037C		-	100.000	11	11				
2561C3		-	100.000	10	10				
226C		-	100.000	8	8				
1788C		-	100.000	8	8				
1546C		-	100.000	8	8				
2766C		-	100.000	6	6				
019C		-	100.000	6	6				
892SC		-	100.000	5	5				
109C		-	100.000	5	5				
07C		-	100.000	4	4				
799C		-	100.000	3	3				
1797C		-	100.000	2	2				
017C		-	100.000	1	1				
2192C	1		95.833	23	24	1			
213S		-	94.737	18	19			1	

---

[16] Recall that the vertical proximity of 044 and 04C1 is coincidental and therefore provides no evidence for or against relationships in this diagram. They are next to each other as siblings are sorted alphabetically!

𝔓66C*		-	92.857	26	28			1	1
1278C1		-	91.667	11	12			1	
0290		-	91.667	11	12				1
02	2		87.850	94	107	5	3	4	1
083	3		85.714	12	14	2			
063		-	85.714	12	14	1	1		
1344C		-	85.714	6	7			1	
029			85.714	6	7	1			
𝔓60		-	85.000	17	20	1	1	1	
038	5		84.545	93	110	7	6	3	1
F1		-	84.404	92	109	7	7	2	1
892		-	82.000	41	50	3	3	3	
1242	6		81.818	90	110	7	6	5	2
2585		-	81.481	22	27	2	2	1	
2786	7		80.909	89	110	8	7	4	2
044	8		80.000	88	110	10	5	5	2
579	9		79.545	70	88	8	2	4	4
033	10		79.310	69	87	6	5	5	2
2561	11		78.652	70	89	8	6	3	2
01	12		78.182	86	110	14	3	4	3
𝔓66	13		77.778	70	90	10	4	4	2
A	14		77.451	79	102	19		2	2
032	15		75.281	67	89	11	2	6	3
𝔓75	16		75.000	33	44	7	3	1	
430		-	75.000	18	24	2	2	1	1
05S			75.000	9	12	2			1
03	18		74.545	82	110	19		6	3
05	19		73.333	66	90	9	6	6	3
019	20		70.213	66	94	12	9	5	2
869		-	66.667	14	21	2	2	2	1
04	21		60.377	64	106	19	15	7	1

Table 112: Potential ancestors with W1=04C1 highlighting group members

397 is a parent of three group members: 𝔓60, 𝔓66 and 032S. Should it therefore be a group member? In the complete diagram it has three parents: A and 044 (in the group) and 597. It has six further children in the complete diagram: 022, 0233, 249, 865, 869, and 1321. In the median joining network, 397 is in a small sub-network with 333. The next closest witnesses are core majority witnesses such as MT, 1320, 226 and so on. In the phylogenetic tree it is in a sub-tree with a strong 98% branch, shown in Figure 103. It seems that 397 is

therefore closer to its children than its parents in the global stemma, and so it need not be considered further.

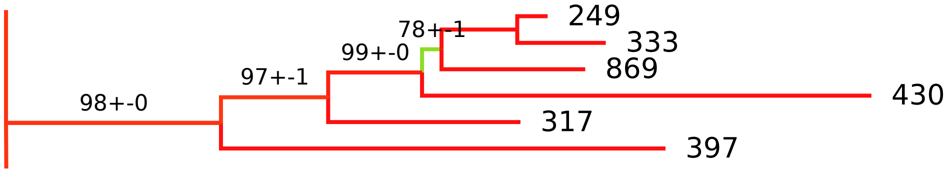


Figure 103: Sub-tree of MrBayes' phylogenetic tree showing 397

We can now iterate our process and create a global stemma subset diagram including 044 and 892 in the group. The new partial global stemma is shown in Figure 104. A risk with this kind of process is that each iteration will yield a few more witnesses to include in the group, and eventually the whole diagram is consumed. However, in this case the only witnesses suggested (visually) as being worthy of consideration are 04C1, 226, 397 and F1 – all of which were considered in the first attempt and excluded. Therefore it seems we have arrived at a stable group of witnesses: A, 𝔓45, 𝔓60, 𝔓66, 𝔓66C*, 𝔓75, 01, 01Cca, 03, 04, 05, 019, 029, 032, 032S, 044, 070, 083, 086, 33 and 892. This has served as an example of the value of exploring such manuscript data using different techniques. Further research is needed to determine what (if any) textual characteristics can be identified as group markers.

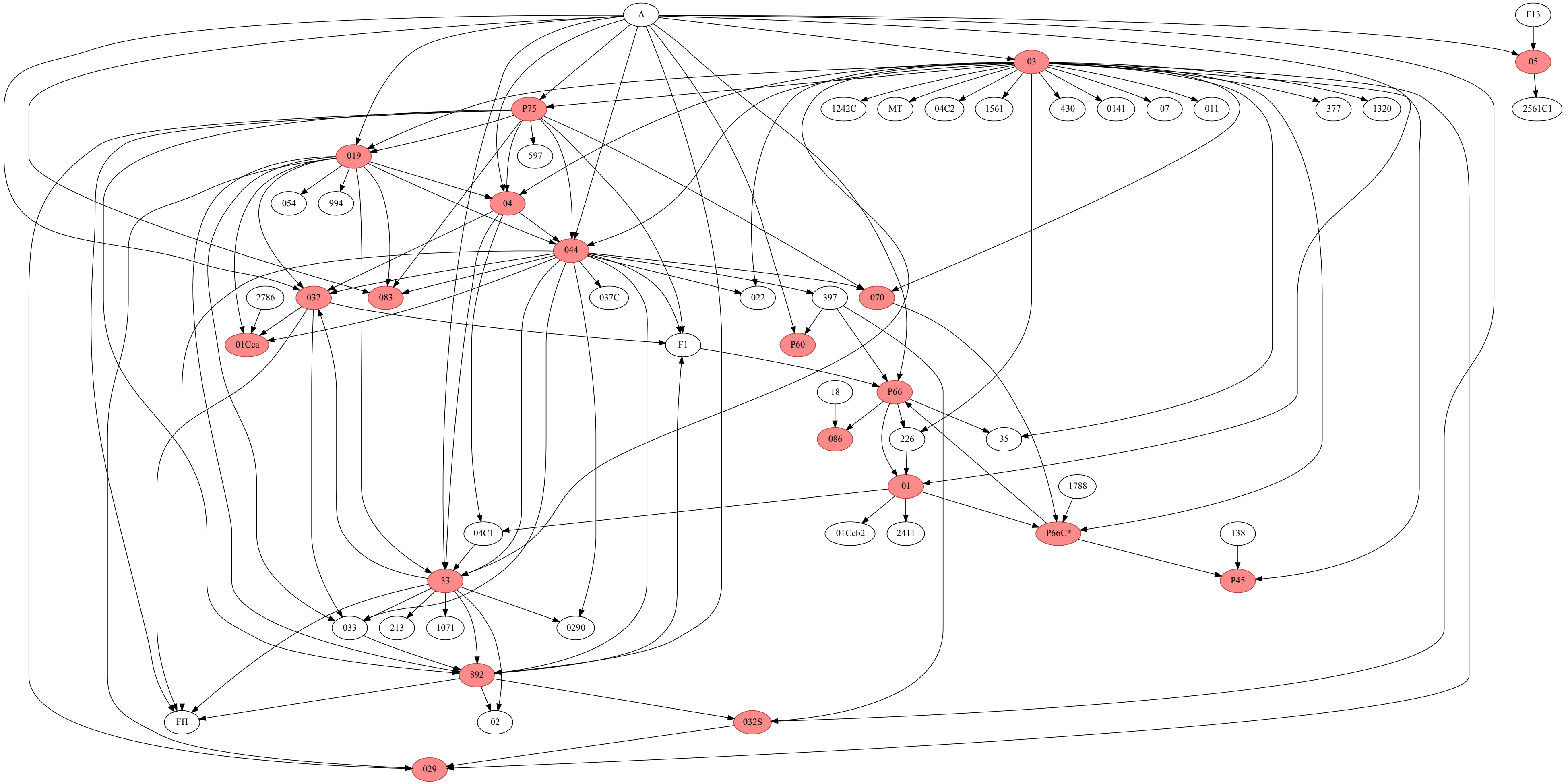


Figure 104: Subset of the global stemma for the “non-majority” witnesses (version 2)

Conclusion

This process has shown that the three diagrams created by independent methods (*MrBayes*'s phylogenetic consensus tree, *Network*'s median joining network and the CBGM's global stemma) do indeed identify similar groupings of manuscripts. This is a strong indicator that the methods are identifying structures inherent in the underlying data. Thus we can say that it is entirely appropriate to apply the CBGM to such data, with the caveats discussed in previous chapters.

As for a use for phylogenetics in textual criticism, a straightforward example would be to use such tools to identify groups of manuscripts, and also those which are more distantly related, and then to use such information to aid in the process of selecting which manuscripts to include in a critical edition. For example, phylogenetics can help to identify appropriate representatives of well known families (for example Chapter 5 showed that 826 is central to Family 13). Similarly, phylogenetics can identify witnesses that are not closely related to any other and thus it presents a strong case for their inclusion.

Chapter 11: Conclusions

In this thesis I explored and analysed the Coherence-Based Genealogical Method (CBGM) both directly and using phylogenetics. I have suggested several areas where it could be improved. The CBGM and phylogenetics are conceptually similar, while being methodologically very different, and their results are therefore effectively independent. The primary conclusion of this thesis is that the CBGM and phylogenetics can both be applied successfully to manuscript traditions. This can be seen when comparing the results of applying the CBGM with those produced by applying phylogenetic techniques to the same input data. In summary, I propose the following five changes to the CBGM:

1. Textual flow diagrams should show more information, either as numbers, colours or line-styles:
    a. Show both the rank and coherence percentage of potential ancestors to avoid the user being misled regarding multiple emergence.
    b. Highlight weak textual flows to avoid the user basing decisions on textual flows that will change with only a few modified local stemmata.
2. The user needs to be able to dynamically apply different settings to a textual flow diagram to reduce the chances of a diagram being misleading:
    a. First, the user should be able to specify both a rank and a coherence percentage as the “connectivity” setting.
    b. Secondly, the user should be able to specify the inclusion or exclusion of weak and undirected textual flows.
3. The Levenshtein distance should be used to calculate pre-genealogical coherence.

4. Coherence data based on unambiguous local stemmata should be kept and used to check later decisions.
5. The decision to include or exclude correctors should be based on the user's goal(s) when applying the CBGM.

In Part One I introduced the CBGM and reviewed the primary and secondary literature. I showed that the primary literature is, generally speaking, inaccessible to non-experts and is complicated. This was confirmed many times in my review of the secondary (and often negative) literature. I was surprised by the number of scholars who have either totally or partially rejected the CBGM. It seems that such rejections often stem from misunderstandings of the method (sometimes simply due to different definitions of terminology) or mistrust of such a radical change (or addition) to the traditional methods of textual criticism. It is to be hoped that Wasserman and Gurry's attempts to provide a more accessible introduction to the CBGM will be of assistance here. Similarly, my step-by-step example of the CBGM in Chapter 3 based on just four verses should be of significant benefit to anyone seeking to understand how to apply the CBGM. It is worth noting that fully understanding the CBGM, and applying it to even a short text, both take a very long time. The CBGM is powerful, and does what it sets out to do: it captures the textual decisions made and allows the user to see the consequences of those decisions. It therefore brings a new level of consistency to textual criticism.

Part Two showed that phylogenetic software can be successfully applied to the ECM data for John's Gospel and raised many interesting research questions by allowing manuscript relationships to be visualised effectively. Several of these relationships are worthy of further study, for example the trio 0141, 821 and 1370; or the very close pair 217 and 578. Applying phylogenetics is much less time-consuming than using the CBGM, but phylogenetics cannot create an initial text. It could be very useful as an aid in the pragmatic task of the selection (or elimination) of manuscripts for subsequent inclusion in the CBGM/ECM. For example, selection could be aided by identifying groups and appropriate representative manuscripts of such groups, or by identifying manuscripts that have no close relatives at all. Part Two also contained a brief excursus showing that *MrBayes* may be sped up eighteenfold for the IGNTP data for John 2, by setting appropriate parameters as opposed to simply using *MrBayes*' default values.

In Part Three I suggested several key improvements that could be made to the CBGM. First, Chapter 7 showed that textual flow diagrams should show both the rank of potential ancestors and the coherence percentage, thus providing the reader with the necessary information to understand whether the data implies multiple emergence or which places require further investigation before such a decision could be taken. Secondly, considering the connectivity setting used to draw a textual flow diagram for contaminated traditions where many manuscripts are also lost, I showed that a diagram based on rank as the connectivity setting will be more likely to show multiple emergence for later variant readings and to show perfect coherence for earlier ones. Conversely, a diagram based on coherence percentages will tend to show multiple emergence for earlier readings and perfect coherence for later ones. The user should be able to specify both a maximum rank *and* a minimum coherence percentage when creating a textual flow diagram – then multiple emergence would be indicated in both situations rather than being suppressed in one or the other.

In Chapter 8 I proposed a further important modification to the CBGM: the user needs to be able to apply different criteria dynamically when creating textual flow diagrams, and needs to be aware of their benefits and disadvantages: a minimum strength could be used to create a textual flow that is less susceptible to later change and is therefore useful for refining local stemmata; weak relationships should be highlighted to prevent the user accidentally relying on them; undirected (or weak) textual flows could be allowed, but highlighted to create a diagram less likely to contain confusing or misleading ancestral relationships, while being less useful for refining local stemmata.

Chapter 9 set out three main suggestions. First, the algorithm used to generate pre-genealogical coherence is entirely at the mercy of external factors and is not objective, even though it is hailed as such. It is dependent on the decisions made by the textual scholar applying the method, such as what to regularise and how to collate the text. These decisions shape the basic input data for the pre-genealogical coherence calculations. It is also at the mercy of the extant tradition, for any new variant reading found in a freshly discovered manuscript will change the collation (unless the scholar regularises it away). By contrast, the Levenshtein distance presents a result that has a much better claim to being objective than does the CBGM's own algorithm and would not be affected by either of these factors.

Secondly, the coherence data based on only the unambiguous local stemmata should be retained and periodically used to check that later results have not been skewed by the potential forced feedback loop that can create falsely strong textual flow.

Thirdly, I argued that further research should be carried out to identify whether the CBGM's algorithm should be modified to allow weights to be applied to variant units. The question is whether this would yield results valuable enough to warrant the considerable extra effort required to apply such weights to thousands of units.

Fourthly, when applying the CBGM the scholar should choose whether or not to include all correctors. All correctors should be included if the goal is to explore the development of the text over time. However, if the scholar's aim is simply to identify an initial text then they may safely use the texts as they left the firsthand or scriptorium (with exceptions such as 𝔓66* which includes many readings found in no other witness).

In Chapter 10 I compared the CBGM's results to those of independent phylogenetic methods (*MrBayes*'s phylogenetic consensus tree and *Network*'s median joining network), showing that all three methods are identifying real structures inherent in the underlying data. Therefore all three methods can be trusted as long as their output is properly interpreted. I have found phylogenetics a useful tool for discovering facts about the manuscript tradition and would recommend its wider use in textual criticism.

My own CBGM implementation has been used throughout this thesis.[1] It has allowed me to experiment with changes to the method in ways not possible just using the existing INTF implementation. There is considerable scope for further work here. In particular, two major developments would be valuable for my CBGM implementation. Its primary design concerns are to be accessible and modifiable – and it can therefore be quite slow at certain tasks. There are doubtless several areas where its speed could be improved substantially. It would also be useful to provide a graphical interface to it, allowing it to be used by researchers with less technical expertise. This is an Open Source implementation and perhaps, therefore, other scholars will choose to collaborate on it.

Wasserman and Gurry have identified a significant problem: "What scholars and student[s] most need now is a version of the CBGM that allows them to edit the local stemmata and create their own database of genealogical data."[2] The new implementation of the CBGM from the INTF and the Cologne Center for eHumanities will be very useful and contains many improvements on the previous system, but it is still something of a "black box" that can only be operated with the aid of the INTF. My implementation of the CBGM allows

---

[1] See https://doi.org/10.5281/zenodo.1296287 or https://github.com/edmondac/CBGM

[2] Wasserman and Gurry, *A New Approach to Textual Criticism*, 119.

full transparency and can be operated by anyone with access to a computer running Linux – and thus should address this problem highlighted by Wasserman and Gurry.

In the six years I have spent on this part-time PhD the scholarly landscape surrounding the CBGM has changed radically, for example the introduction of the new Cologne interface and the seminal contributions to the field by Gurry and Wasserman. It has also, in other respects, stayed much the same: The method itself has changed little, despite the continuing waves of criticism from the secondary literature. I hope that this thesis will help to convince sceptical scholars that the CBGM improves the ECM's methodological consistency and therefore the quality of its text. I also hope that my suggestions will be incorporated into the CBGM with the potential of further improving the text of the ECM.

Appendix 1: CBGM Example (Remaining Data)

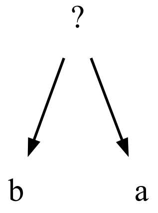
Introduction

This appendix completes the CBGM example from Chapter 3 and together they contain the complete record of the local stemmata (first pass with only pre-genealogical coherence data and second with genealogical coherence data) and optimal substemmata for John 6:21-24 (Papyri and Majuscules).

Local Stemmata

John 6:21/28-30

The two readings in John 6:21/28-30 differ only in their case. Moule states that "[t]he use of the cases after ἐπί is very fluid" and it is no surprise to see the accusative and genitive both used in this situation.[1] The genitive is more common, and indeed throughout the IGNTP transcriptions of John's gospel, the ratio of "επι της γης" to "επι την γην" is about 10:1. But this isn't enough to be confident in this variant unit, and so we will wait for genealogical coherence data. The readings are shown in Table 114.



a	της γης ...	
b	την γην	01, 0211
lac		𝔓75, 091

Table 113: Local stemma for John 6:21/28-30

---

[1] Moule, *An Idiom Book of New Testament Greek*, 49.

John 6:21/36

Here 01 reads υπηντησεν, which is a word John uses when people meet one another but here seems like a mistake by the scribe. The local stemma is shown in Table 114.

a	υπηγον	...
b	υπηντησεν	01
	lac	091

Table 114: Local stemma for John 6:21/36

John 6:22/3

In John 6:22/3, 0211 uniquely adds τε – creating the local stemma in Table 115.

a		...
b	τε	0211
	lac	091

Table 115: Local stemma for John 6:22/3

John 6:22/10

In this unit, 038 and 045 don't read the article. The normal Johannine construction in this case would be to use the article, and so it seems clear that reading *a* is prior. The local stemma is shown in Table 116.

a	ο	...
b		038, 045
	lac	091

Table 116: Local stemma for John 6:22/10

John 6:22/12

Here, 01 uniquely reads εστως instead of εστηκως. Morphologically, both are RAPMSN but εστως is considerably more common.[2] John 12:29, for example, reads "ο ουν οχλος ο εστως" in the majority of older witnesses including 01.[3] So it seems that 01 has changed this text, possibly by harmonisation. The local stemma is shown in Table 117.

a	εστηκως	...
b	εστως	01
	lac	091

Table 117: Local stemma for John 6:22/12

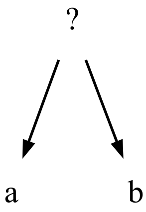
---

[2] RAPMSN = Perfect Active Participle Masculine Singular Nominative

[3] Although here also many witnesses read εστηκως.

John 6:22/40

John 6:22/40-52 was a later addition. This long addition contains a number of variant units, the first of which is John 6:22/40.[4] Here 05 and 0211 omit εκεινο. On the one hand, with the word the text is smoother and so 05 and 0211 could be witnessing to an earlier, shorter form. On the other hand, however, εκεινο sounds somewhat like a marginal note, and would therefore be later. We will wait for further evidence. The readings are shown in Table 118.



a	εκεινο ...
b	05, 0211
lac	𝔓75, 02, 03, 019, 022, 032, 044, 063, 091

Table 118: Local stemma for John 6:22/40

John 6:22/42

034 omits John 6:22/42; εις. Normally, ενεβησαν is paired with εις (as in John's two uses: John 6:17 and John 6:24). The text doesn't make sense without εις so 038 must have made a mistake in omitting it. The local stemma is shown in Table 119.

a	εις ...
b	034
lac	𝔓75, 02, 03, 019, 022, 032, 044, 063, 091

Table 119: Local stemma for John 6:22/42

---

[4] Thus making John 6:40-52 an "overlapping" variant unit.

<u>John 6:22/46</u>

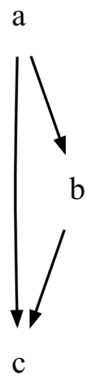
Here, 037 and 047 read ανεβησαν (ascended) rather than ενεβησαν (embarked). While ανεβησαν can be used to describe getting into a boat the vast majority of early witnesses read ενεβησαν for this concept in John's gospel, with ανεβησαν attested to almost exclusively by later witnesses.[5] So it seems clear that ανεβησαν is a later change (or error). The local stemma is shown in Table 120.

a	ενεβησαν	...
b	ανεβησαν	037, 047
	lac	𝔓75, 02, 03, 019, 022, 032, 044, 063, 091

Table 120: Local stemma for John 6:22/46

<u>John 6:22/52</u>

Here 01 replaces αυτου with του ιησου and 05 adds ιησου, making the odd phrase αυτου ιησου. 05 was later corrected to read του ιησου but is its original reading here evidence of the evolution of this improving addition? 05's reading αυτου ιησου seems to be derived from both the other readings, and so we can construct the local stemma – see Table 121



a	αυτου	...
b	του ιησου	01
c	αυτου ιησου	05
	lac	𝔓75, 02, 03, 019, 022, 032, 044, 063, 091

Table 121: Local stemma for John 6:22/52

---

[5] The passages in question are John 6:22 (this verse), John 6:24 and John 21:3.

John 6:22/61

In John 6:22/61, 02 has evidently inserted a clarifying ο ιησους, creating the local stemma in Table 122.

a		...
b	ο ιησους	02

Table 122: Local stemma for John 6:22/61

John 6:22/68-70

In this unit 034 omits ο ιησους. While the sense is fine without ο ιησους – and therefore it could be argued that those witness that include it do so as a clarification – it seems nevertheless clear that 034 has omitted it.[6] So the local stemma in Table 123 can be created.

a	ο ιησους ...	
b		034

Table 123: Local stemma for John 6:22/68-70

John 6:22/76

In John 6:22/76 the two words πλοιαριον and πλοιον roughly split the attestations. Are the boats little or not? This is very likely to be a case of assimilation to the unanimously attested πλοιαριον in John 6:22/24, since this refers to the very same boat. The local stemma is shown in Table 124.

---

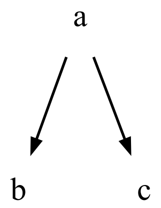
[6] Is this a situation where textual scholars will readily transgress the maxim that "witness should be weighed and not counted"? Perhaps pragmatism allows such transgressions.

a	πλοιαριον	07, 011, 013, 021, 028, 030, 034, 036, 037, 038, 045, 047, 063
b	πλοιον	𝔓75, 01, 02, 03, 05, 017, 019, 022, 032, 044, 091, 0141, 0211
	lac	031

Table 124: Local stemma for John 6:22/76

John 6:22/88

In John 6:22/88, 038 seems to improve the text by replacing *απηλθον* with *εισηλθον* and thus reflecting word 60 (*συνεισηλθε*). 01 omits the word entirely creating a further legitimate reading. But could this omission actually be a prior reading? 01's scribe has a proclivity to drop words, and so this is probably an omission. The local stemma is shown in Table 125.

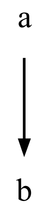


a	απηλθον	...
b	εισηλθον	038
c		01
	lac	031

Table 125: Local stemma for John 6:22/88

John 6:23/1

022 uniquely includes *και* at the start of verse 23 (reading *και αλλα δε*), and this is surely an addition. The local stemma is shown in Table 126.

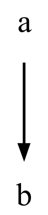


a		...
b	και	022
	lac	031

Table 126: Local stemma for John 6:23/1

John 6:23/3

In John 6:23/3 𝔓75, 03, 019 and 091 read the base text (om) but the rest of the witnesses add δε.[7] With the exception of a handful of witnesses to John 10:16, this is the only instance of “αλλα δε” in John's gospel.[8] This word is unnecessary and clearly an addition. Thus the local stemma can be constructed – see Table 127.

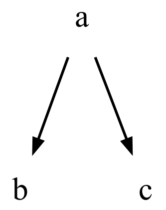


a 𝔓75, 03, 019, 091
b δε ...
lac 01, 05, 031

Table 127: Local stemma for John 6:23/3

John 6:23/12-16

032 omits John 6:23/12-16. So was the omission an earlier form and the “near the place” a later addition, or did 032's scribe shorten the text? It seems likely that 032 has simply omitted the words. 01's text “being near” is half-way regarding simplicity, but is probably a symptom of 01's scribe either making an error or simply misunderstanding a complicated text. The local stemma is shown in Table 128.



a εγγυς του τοπου ...
b εγγυς ουσης 01
c 032

Table 128: Local stemma for John 6:23/12-16

---

[7] Apart from 01 and 05 which are lacunose for this variant unit being covered by the overlapping unit John 6:23/2-6.

[8] Considering all the transcriptions prepared for the ECM version of John.

John 6:23/20-22

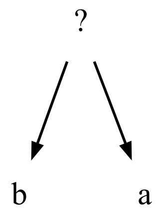
In John 6:23/20-22 01 again has an unique reading but it is not obvious whether it is prior or posterior to the rest of the tradition. The readings are shown in Table 129.

a	εφαγον τον	...
b	και εφαγον	01

Table 129: Local stemma for John 6:23/20-22

John 6:24/2-10

In John 6:24/2-10 we find yet another overlapping unit representing a unique reading of 01. It is not easy to create the local stemma, so this will be deferred. The readings are shown in Table 130.

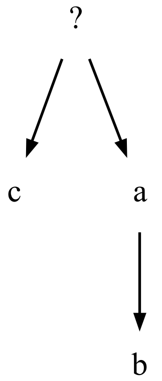


a	οτε ουν ειδεν ο οχλος	...
b	και ιδοντες	01

Table 130: Local stemma for overlapping unit John 6:24/2-10

John 6:24/6

In John 6:24/6 most witnesses read ειδεν. 013 has made a mistake and written ειπεν. 030 reads εγνω, which is valid here. But relating ειδεν to εγνω is not straightforward and so only a partial local stemma can be created (see Table 131).

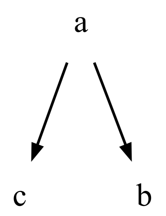


a	ειδεν	...
b	ειπεν	013
c	εγνω	030
	lac	01

Table 131: Local stemma for John 6:24/6

John 6:24/14

In John 6:24/14 01 and 038 have the article before ιησους, which seems to be a clear addition. 013 omits the word, and while not strictly necessary this does seem like a mistake. The local stemma is shown in Table 132.

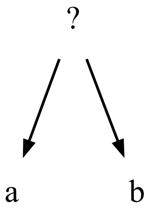


a	ιησους	...
b	ο ιησους	01, 038
c		013
	lac	031, 063

Table 132: Local stemma for John 6:24/14

John 6:24/14-20

In the overlapping variant unit John 6:24/14-20, 01 again has a unique reading – this time a matter of word order and a change of tense from *εστιν* to *ην*. It is likely (given 01's character in these verses) that genealogical coherence will be able to help decide this local stemma later. For now the readings are shown in Table 133.



a	ιησους ουκ εστιν εκει	...
b	ουκ ην εκει ο ιησους	01
lac		031

Table 133: Local stemma for John 6:24/14-20

John 6:24/28

Here we find yet another variant unit that exists purely because of 01. Here 01 omits the word and is known for shortening the text in this way. So even though 01's reading is reasonable, it is probably posterior. The local stemma is shown in Table 134.

a	αυτου	...
b		01
lac		063, 091

Table 134: Local stemma for John 6:24/28

John 6:24/31

In John 6:24/31, 030, 036 and 0211 (out of order) include the word *και*, which seems a clear addition. This creates the local stemma shown in Table 135.

a		...
b	και	030, 036, 0211
lac		05, 063, 091

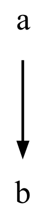
Table 135: Local stemma for John 6:24/31

John 6:24/30

John 6:24/30 contains a second example of the question of ενεβησαν versus ανεβησαν (𝔓75, 01 and 019) – with regard to a boat.[9] Now 𝔓75 and 019 are closest relatives (Table 136) with 97.143% coherence – suggesting their shared reading here is not coincidental. 01, however, is a very distant relative (last in the table with only 42.424% coherence). So either ανεβησαν is the initial text, or it has emerged independently twice. Sadly for this line of argument, it is entirely conceivable that this single-letter change could have arisen independently. So in keeping with the decision in John 6:22/46 above, we can construct the local stemma (see Table 137).

W2	NR	PERC1	EQ	PASS
019	1	97.143	34	35
03	2	94.444	34	36
091	3	88.889	16	18
022	4	88.571	31	35
044	5	86.111	31	36
032	6	85.714	30	35
0141	7	83.333	30	36
02		83.333	30	36
063	9	82.143	23	28
011	10	80.556	29	36
037	11	80.000	28	35
031		80.000	24	30
017	13	77.778	28	36
021		77.778	28	36
07		77.778	28	36
028	16	75.000	27	36
034		75.000	27	36
036		75.000	27	36
045		75.000	27	36
047		75.000	27	36
05	21	72.414	21	29
013	22	72.222	26	36
0211		72.222	26	36
030		72.222	26	36
038		72.222	26	36
01	26	42.424	14	33

Table 136: Pre-genealogical coherence with W1=𝔓75



a ενεβησαν	...
b ανεβησαν	𝔓75, 01, 019
lac	05, 063, 091

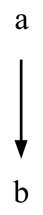
Table 137: Local stemma for John 6:24/30

---

[9] See John 6:22/46 above. Note that no witness attests ανεβησαν in both verses.

John 6:24/32

In John 6:24/32 01 and 028 omit the pronoun. This seems likely to be another example of 01's scribe dropping a word. 028 is 01's second closest relative (Table 139) and so this is probably not a case of multiple emergence. The local stemma is shown in Table 138.



a	αυτοι ...	
b		01, 028
	lac	05, 063, 091

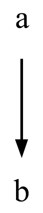
Table 138: Local stemma for John 6:24/32

W2	NR	PERC1	EQ	PASS
031	1	48.485	16	33
028	2	47.368	18	38
A	3	45.833	11	24
07	4	44.737	17	38
038		44.737	17	38
021		44.737	17	38
017		44.737	17	38
0141		44.737	17	38
013		44.737	17	38
037	10	43.243	16	37
𝔓75	11	42.424	14	33
019		42.424	14	33
045	13	42.105	16	38
036		42.105	16	38
030		42.105	16	38
011		42.105	16	38
047	17	39.474	15	38
034		39.474	15	38
0211		39.474	15	38
063	20	38.462	10	26
03	21	38.235	13	34
032	22	36.364	12	33
022		36.364	12	33
044	24	35.294	12	34
05	25	33.333	11	33
091		33.333	5	15
02	27	32.353	11	34

Table 139: Pre-genealogical coherence with W1=01

John 6:24/30-38

05's unique reading in these words is represented as an overlapping unit: John 6:24/30-38. It seems very likely that 05's attempted improvement is posterior here, although actually it seems that 05 has misunderstood which boats are being talked about. The local stemma is shown in Table 140.



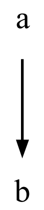
a	ενεβησαν αυτοι εις τα πλοια ...	
b	ελαβον εαυτοις πλοιαρια	05
	lac	063, 091

Table 140: Local stemma for overlapping variant unit John 6:24/30-38

Genealogical coherence and textual flow

<u>John 6:22/40</u>

This unit can be resolved by considering the likelihood of multiple emergence of the addition of εκεινο. Note that in this unit the initial text is lacunose, as it attests the omission in the overlapping variant unit John 6:22/40-52. Recall that "OL_PARENT" is a hypothetical witness where the addition in John 6:22/40-52 occurred.[10] The only viable scenario is that *b* comes from *a*, and 05 and 0211 omitted the word (see Figure 106).[11] The local stemma is shown in Table 141.



a εκεινο ...	
b	05, 0211
lac	𝔓75, 02, 03, 019, 022, 032, 044, 063, 091

Table 141: Local stemma for John 6:22/40

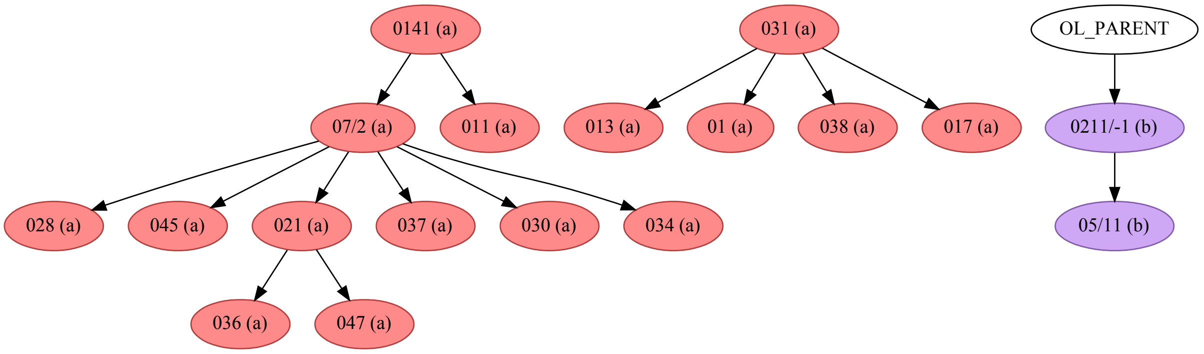


Figure 105: Textual flow diagram for John 6:22/40 with parent text b

---

[10] No extant witness can serve in this role here since the two witnesses that normally occur beneath it (0141 and 031) have undirected genealogical coherence – and thus neither can serve as ancestor to the other.

[11] If we allow absolute connectivity then 05 must descend from 0211, its 11th most likely potential ancestor and their omission is related. For lower connectivity values 05 and 0211 independently omitted the word – which is perfectly believable.

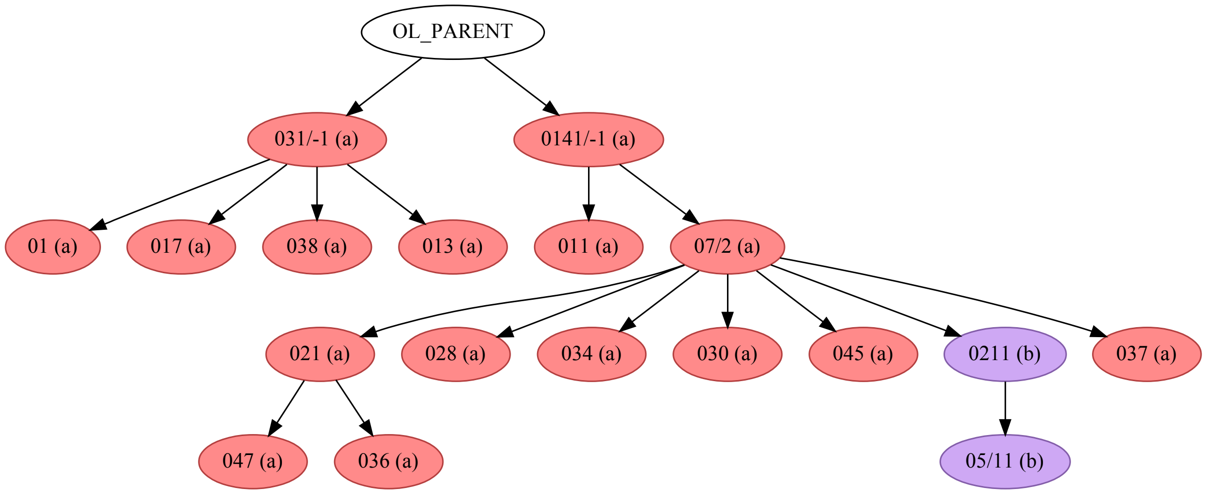
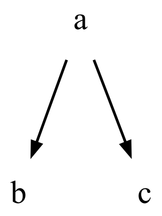


Figure 106: Textual flow diagram for John 6:22/40 with parent text a

<u>John 6:23/2-10</u>

In this unit we have two genitive absolute variations (*b* 01 and *c* 05) on the majority reading *a*. 01 and 05 are almost as distantly related as is possible in this dataset, and so it is highly unlikely that their readings are related. These genitive absolutes must be seen, therefore, as two independent improvements on the *a* reading. The local stemma is shown in Table 142.

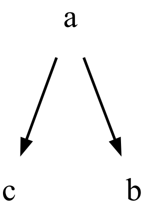


a αλλα ηλθεν πλοιαρια εκ τιβεριαδος ...
b επελθοντων ουν των πλοιων εκ τιβεριαδος 01
c αλλων πλοιαρειων ελθοντων εκ τιβεριαδος 05

Table 142: Local stemma for John 6:23/2-10

John 6:24/6

Reading *a* must surely be the initial text and so the only remaining question is that of the origin of *c*.[12] Reading *c* is more closely related to witnesses reading *a* than *b*, and indeed has only undirected coherence with the witness for reading *b* (see potential ancestors of 030 in Table 144). So the local stemma can be determined (see Table 143).



a	ειδεν	...
b	ειπεν	013
c	εγνω	030
	lac	01

Table 143: Local stemma for John 6:24/6

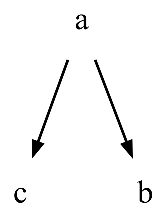
W2	NR	D	PERC1	EQ	PASS	READING	W1<W2	W1>W2	UNCL	NOREL
07	1		95.122	39	41	a	1		1	
031	2		94.286	33	35	a	1		1	
063	3		93.103	27	29	a	1		1	
045		-	92.683	38	41	a	1	1	1	
028		-	92.683	38	41	a	1	1	1	
021		-	92.683	38	41	a	1	1	1	
013		-	92.683	38	41	b	1	1	1	
011	4		92.683	38	41	a	1		2	
037		-	92.500	37	40	a	1	1	1	
017		-	87.805	36	41	a	2	2	1	
0141	5		87.805	36	41	a	2		2	1
A	6		83.333	25	30	a	5		1	
044	7		78.378	29	37	a	3		4	1
02			78.378	29	37	a	4	2	2	
022	9		77.778	28	36	a	3	2	3	
03	10		75.676	28	37	a	5	1	3	
032	11		75.000	27	36	a	3	2	4	
019			75.000	27	36	a	4	2	3	
𝔓75	13		72.222	26	36	a	5	2	3	
091		-	72.222	13	18	a	2	2	1	

Table 144: Potential ancestors with W1=030, showing their reading at John 6:24/6

---

[12] As in other places in this section, what is text-critically obvious does sometimes seem a little too much like counting witnesses. But surely a single, otherwise relatively unimportant majuscule wouldn't carry the initial text?

John 6:24/38



a	πλοια	...
b	πλοιαρια	𝔓75, 03, 05, 019, 022, 032, 044
c	πλοιον	01
	lac	063, 091

Table 145: Local stemma for John 6:24/38

This variant unit can be resolved by considering a previous unit, and following in its tracks. In John 6:23/4-10 𝔓75, 03, 032, 044 and 0141 referred to these *exact same boats* as πλοια (while here all but 𝔓75, 03, 05, 019, 022, 032 and 044 do!) Witnesses 01 and 05 were lacunose in that variant unit while 063 and 091 are lacunose here. A very small number of witnesses seem to have assimilated this to their reading in 23/4-10. In that variant unit we decided that the πλοιαρια readings were prior – and therefore here the πλοια reading must be (for most of the same witnesses to read the initial text in both places when referring to these boats.) See Table 145 for the local stemma.

Optimal Substemmata

Vorf	**Vorfanz**	**Stellen**	**Post**	**Fragl**	**Offen**	**Hinweis**
A	1	29	3	5	0	<<
A, 03	2	29	3	5	0	<<
A, 091	2	29	3	5	0	<<
A, 03, 091	3	29	3	5	0	<<
03	1	27	3	5	2	
03, 091	2	27	3	5	2	
091	1	12	2	5	18	

Table 146: Combinations of ancestors for 02

02 has three potential ancestors, and so there are seven possible combinations, as shown in Table 146. The decision in this case is easy, however, as the combination that explains the most readings by agreement is simply {A}, which explains all 02's readings.

𝔓75's combinations of ancestors are shown in Table 147. Again, this decision is straightforward. 03 explains 30 readings by agreement, but cannot explain one. A explains all 𝔓75's readings, but only 29 by agreement – and two more by posterity than 03. So {03, A} explains 30 all the readings and 30 by agreement. But so does {03, 031}. Table 148 shows that 031 is a significantly more distant relation than 03 and A, and therefore the combination {03, A} is to be preferred.

Vorf	Vorfanz	Stellen	Post	Fragl	Offen	Hinweis
03, A	2	30	2	4	0	<<
03, 031	2	30	2	4	0	<<
03, A, 031	3	30	2	4	0	<<
03	1	30	1	4	1	
A	1	29	3	4	0	<<
A, 031	2	29	3	4	0	
031	1	21	3	4	8	

Table 147: Combinations of ancestors for 𝔓75

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
019		-	97.143	34	35				1
03	1		94.444	34	36	1			1
A	2		90.625	29	32	3			
02		-	83.333	30	36	3	3		
037		-	80.000	28	35	3	3	1	
031	3		80.000	24	30	3	2	1	

Table 148: Potential ancestors with W1=𝔓75

The same classes of argument lead to these decisions: 07 {063, 0141}; 019 {03, A}; 044 {03, 𝔓75}; 091 {A}; and 0141 {044, A}.

017 has eleven potential ancestors none of which can explain all of its readings. It has twenty equally good pairs of witnesses that explain all its readings, and so {031, 0141} is chosen because the summed rank of 0131 and 0141 is higher than the summed rank of any other such pair (in this case the first and joint-second highest ranked potential ancestors).[13] Applying all these techniques provides decisions for: 011 {0141, 063}; 021 {07, 091}; 022 {044, A}; 028 {07}, 032 {044, 03}, 037 {07}, 045 {07}.

---

[13] It seems entirely reasonable to prefer the combination made up of the closest ancestors when all else is equal.

05 has 524,288 possible combinations, but the decision boils down to these four: {03, 0211}, {091, 03, 0211}, {091, 03, 047, 0211} and {091, 𝔓75, 047, 0211}. The two combinations of length four explain most by agreement but both require two eleventh ranked potential ancestors. {03, 0211} is the shortest, but requires John 6:23/26-30 (three word omission) to be explained by posteriority which is not preferable. 091 is 05's highest ranked potential ancestor, and explains this reading by agreement. The four-length combinations explain one further reading by agreement over this combination, but it is a simple one-character change (μονοι to μονον) and is fine by posteriority. As such {091, 03, 0211} is the optimal substemma.

The decision for 047 comes down to a simple choice between {021} and {021, 037}. 037 explains John 6:22/46 by agreement (ενεβησαν or ανεβησαν) but this variant is not connective and so 037 is not required. 0211 has two equally good best combinations: {036, A} or {030, A}. They are equally good by all measures we have used so far, and as such we must leave the decision open. The remaining witnesses, even though they have very large possible combinations, are all explained best by a single witness: 013 {07}; 030 {07}; 034 {07}; 036 {021}; and 038 {045}.

Appendix 2: John 2 Phylogenetic Tree from Job 920421

Appendix 3: NEXUS File of John 2

This is the contents of the NEXUS file of Parker's ECM collation of John 2 (as of the 20th May 2016), including all witnesses from the ECM that are extant in at least 25% of variant units in this chapter :

```
#NEXUS
begin data;
dimensions ntax=176 nchar=239;
format missing=? gap=- matchchar=. datatype=standard;
matrix
L663nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa?????????????????????????????????
L1096nS1W1D6 ????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaabaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa?????????????????????????????????
L17nS1W2D2 aaaaabaaacaaaaaaaaaa?aaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaadaabaaaaaa?
aaae??????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
2585 aaaaaaaaaaaaaaaaaadaaaaaabaaaaaaaaababaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaacaababaaaa?aaaoaai-
aaaaaaabaaaaaaaaaaaaaaaaaaaaaaacaaaaaaaaaaaaaaacaabaaaaaaaaa?aabaaaa?aa??
aaaaaacaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaaaaaabaaaaaabaaaaaacaaaaaa
L1096nS1W2D2 aaaaabaaacaaaaaaaaaa?aaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?
baaaaaaaaaa???????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????????????
265
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaabaaaaaaabaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaabaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
0233 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa?aaaaalaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaacaaaaaaaaaaaaaaacaabaaaadaaaa?aaaaaaa?aa??
aaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaba
0141
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagaaaaaaaaaaaaaaaaaaaaaaaaaaaaac
caabaaaaaaabaaaaabaacaabaaaaaaaaaa?aaaaaaa?aaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
2372 aaaaaaaaaaaaaaaaaaaa?aaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaacaabaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
1689 aaaaaaaaabaaaaaaaaaa?aaaaaaaaaaaaaaabaaaaaaaaaaafaaaaaaaaaabaaaaaaaaaaaaaaa?a?babaaaa?aaa?
aaaaacaaaabaaaaaaaaaaaaaaaaaaaaaaacaaabaaaaaaaaaaacaabaaaaaaaaa?aabaaaa?aa??aaaaaacaaaaaaaaaaaaaaaaaaaaaabaae-
aaaaaaaaaaaaaaaaabaaaabbaaaaaaaaaaaaa
1561 aaaaaaaaaaaaaaaaaaaa?aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaaeaaiaaaaaab-
baaaabaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
992 aaaaaaaaaaaaaaaaaaaa?aaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?aabd-
baaacabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaaaaaaaaab-
baaaaabaaaaaaaaaaaaa
1 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaabaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaaaaaaaacaaab-
baaaaaaaaaaaaaaaaaaaacbaabaaabaaaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
L335nS1W2D2 aaaaabaaacaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaacaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaaa?
aaa???????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
1582 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaabaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaaaaaaaacaaab-
baaaaaaaaaaaaaaaaaaaacbaabaaabaaaaaaaaaaaacaaaaaaacaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
333 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaa?a?baaaaaa?aaa?aa-
haaaaaaabaaaaaaaaaaaaaaaaaaaaajaaabaaaaaaaaaaaaaaaacaabaaaaaaaaa?aaaaaaa?aaa?
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
173 aaaaaaaaabaaaaaaaaaaaaaaabaaaaaaaaafabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaababaaaaaaaaebaicacaaab-
baaaabacaabaaaabaaaaabaaacaaabaaaaaaaaaaacaabaaaaaaaaa?aaaaaaa?aaa?aaaaaacaaaaaaaaaaaaaaaaaaaabaaaaaabaebbacaaacaaa-
faaaaabadcdccecedcdd
983 baaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaafaaaaaaaaaabaaaaaaaaaaaaaaa?a?babaaaa?aaa?
aaaaacaaaabaaaaaaaaaaaaaaaaaaaaaaacaaabaaaaaaaaaaacaabaaaaaaaaa?aabaaaa?aa??aaaaaacaaaaaaaaaaaaaaaaaaaaaabaae-
aaaaaaaaaaaaaaaaabaaaabbaaaaaaaaaaaaa
063 aaaaabaaacaaaaaab?gaaeaaabaaaaaaaaaaab?aaaaaaaaaabaabaaaaaaaaabaabaaaaaaaaaaaaaaaabaaaaaaaaaapaaafaaaaab-
baaaaaaaaaaaaaaaaaaaaaaababaaaaaaaaaaaaacaaaaaa???????????????????????????????????????????????c?aa???
aaaaaaaaaaa??????????????????????
544 aaaaaaaaaaaaaaaaaaaa?aaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aabaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaabaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L387nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaabaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa?????????????????????????????????
788 aaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaafaaaaaaaaaabaaaaaaaaaaaaaaa?a?babaaaa?aaa?aaaaaaaaabaaaa-
faaaaaaaaaaaaaaaaacaaabaaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaabaaaab-
baaaaaaaaaaaaa
1010 aaaaaaaaaaaaaaacaaaaaaaabaaaaaaaaaeabaaaaaaaaaaabaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaabaaaaaaaaeaaedaaaaabbaaad-
baaaaaaaaaaaaaacaabbaaabaaaaaaaabaabaaaaaaadaaaa?aaaaaaa?da??
aaaaaaaaabaaaabaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaacaaaaaaaaaaaab
33
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacabaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaad
```

```
caabaaabaabaaaaaaaacaaaaabadaaaaaaabaabaaaaa?aaaabacaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaaaaaaacaaaaabaaaaaaaaaaaaa
357 aaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?ca??aaaaaaaaaaaaaaaaaaaa?
aaaaaabaaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
118 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaabaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaacaaaabaaaaaaaaaaaaaaaaaaaacbaabaaabaaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
821
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagaaaaaaabaaaaaaaaaaaaaaaaaaaac
caabaaaaaabaaaaabaacaabaaaaaaaaa?aaaaaaa?aaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaababaaaaaaaaaabbaaaaabaaaaaaaaaaaaa
L60nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaacaaaacaaaaabaaaaaa?aaaaaba?aa??
aaaabaaaaaaaaaaaaaaaaaaaaaaabaagaaa??????????????????????????????????
L1073nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
L253nS1W1D5 ?????????????????????????????????????????????????????????????????????????????????????aaagaaaaecbaaaa-
faaaaaaaaaaaaaaacaaacaaaaaaaaaaaaaaacaabaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
1253 aaaaaaaaaaabaaabaaaa?aaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaa?aabaaaaaa?aaa?
aaaaaaaaaabaaaaaagaaaaaadaaaaaaaeaabaaaaaabaaaaaaaacaaaaaaadaaaa?aaaaaaa?a???
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L329nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????aaigaaaaacbaaag-
baaaaaaaaaaaaaaaacaacaaaaaaaaaaadaaacaabaaaadaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
1571 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaabaaaaaaaaaaacaaaaaaaaaaabaaaaaaaaaaaaaaa?aabaaaaaa?aaa?aaabaaaaab-
baaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L704nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????aaagaaaaecbaaaa-
faaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
1210 aaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaa?aaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaacaaabaaabaaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
044 aaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa?
aaaaaaaaaaaaaaaadaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaa?aaaaaaa?aaa?
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
1788 aaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaacaaaaaaaaaaabaaaaaaaaaaaaaaa?aabaaaaaaaaaaa?aaaaaaaab-
baaaaaaaaaaaaacaaaaaaaaaabaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
01Cca ??????????????????aaaaaaaaaaaaaa?????????aaaaaaa???????????????aaaaaaaaaaaaaabababaaaaaae????????????????
aaaaaaaaaaacaahaaac???????????????????????????????????????????????????aaaaaa???????????????????????????????
ababaaabaaaaaaaaaaaab
083 ??????????????????????????????????????????????????????????????????????aaaaaaaa?aa?aaaaaaaaaaaaadaae-
aaaaaaaaaaaaaaaaaaaaaaaakcaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaa?
aacaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaabaaaaadaaaaaaaaaaaba
2193 aaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaabaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaaaaaacaaaaacaaab-
baaaaaaaaaaaaaaaaaaaaabaabaaabaaaaaaaaaaacaabaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
209 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaabaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaacaaaabaaaaaaaaaaaaaaaaaaaacbaabaaabaaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
597 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaacaaaaaaaaaaaaaaaaaaabaaabaaaaaaacaaabaaaaabaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaababaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaacaaaaaa
L638nS1W2D2 aaaaabaaacaaaaaaaaaha?aaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaa?a?baaaaaa?
aaa???????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
P75
aaabaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacaaaaaaaaaaaaaaadaaaaaaaaaaaababaaaaaaaaaaaaac
caabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaadaaaaaaaaaaaaaaaaaa
22 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaacaaabaaabaaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
865 aaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaacaaaaaaaaaaaaaaaaaabacaaaaaaaae-
aaaaaabaaabaaaaaaaaaaaaaaaaaaaaacaabaaabaabaaaaabaaaaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
037 aaaaaaaaaaaaaaaaaajaaaaaaaaaaaacaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaabaaaaaaaaaaae-
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaabacaaaaaaaaaaaa?aaaaaaaaaa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
295 aaaaaaaaaaaaaaaabaaa?aaaabaaaaaaaaaaaaaaaacaaaaaabaaaaaaaaaaaabaaaaaaaa?a?aacaa?a?baaaaaa?aaae-
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaaa?
aaaaabcaaaaabaaaaaacacaaaa
1319 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaabaaaaaaaaaaagaaaaaaabaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aacaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaabaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L1076nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaacaaaaaaaaaaaa???afaabaaaaaaaaaaaaaaaacaaaaaa?a????eaa?aaaa?aa??aaaaaaaaaaaaaaaa???a???????
cbaaa?????????????????????????????????????
1321 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaa?a?babaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaaabaaaabbaaaaaaaaaaaaa
L770nS1W2D2 aaaaabaaacaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaa?baaaaaa?
aaa???????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
02
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaaabaaaagaaaaaaaaaaaaaaa
aaabaaabaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaaaaa??aaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaeaaaaabaaaaaaaaaaaaba
826 aaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaafaaaaaaaaaabaaaaaaaaaaaaaaa?a?babaaaa?aaa?aaaaaaaaabaaaa-
faaaaaaaaaaaaaaaaaaaacaaabaaaaaaaaaaaacaaaaaaadaaaa?aabaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaeaaaaaaaaaaaaaaaaabaaaab-
baaaaaaaaaaaaa
L1692nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaadaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
1093 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaadabaaaaaaaaaaabaaaaaaaaaacaaaaaaaaaaaaaaa?a?baaaaaa?aaa?aaaeaaaaab-
baaaaaaaaaaaaaaabaaaabaaacaaabaabaaaaabaacaabaaaaaaaaa?aaaaaaa?aa??
```

```
aaaaaacaabaaaaaaaaaaaaaaaaaabaaaaaaabaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
041 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa?
baaaaaaaaaaaaabaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L1091nS1W2D2 aaaaabaaacaaaaaaaaaa?aaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaa?a?baaaaaa?
aaa??????????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
226 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
543 aaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaafaaaaaaaaaaabaaaaaaaaaaaaaaaaaa?a?babaaaa?aaa?aaaaaaaaaabaaaa-
faaaaaaaaaaaaaaaaaacaaabaaaaaaaaaaaaacaaaaaaadaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaadaaa??????????????????????????????????
2768 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaa?a?baaaaaa?aaa?aaaab-
caaaabaaaaaaaaaaaaaabaaaaaaaabaaabaaaaaaaaaaacaabaaaaaaaaaa?aabaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaae-
aaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L735nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aaagacaaecbaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaacaaaaaaaaaaa?aaaaaba?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
L704nS1W2D2 aaaaabaaacaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaabaaaaaa?
aaae?????????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
L32nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaaaaaaaaaaaaaaaaaaaaaaaabaabaaabaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
2575 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaa?aabaaaaaa?aaae-
aaaaaaaadabaaaaaaaaaaaaaaaaaaaaaadaabaaaaaaaaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaabaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
L640nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaabaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaacaaaaaaadaaaacaaeaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
0162 ?????????????????????????????????????????????????????????????????????????????????????????aaa?aaddaaaaaaaaaaaaaaaaaaaaae-
aaaaaccaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa?aaaaaaaaaaaaaaaaa??a?????????a??aaa???????????????????????????????????
1344 aaaaaaaaaaaaaaaabaaa?aaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaababaaaa?aaaeaaaaaaaaaaabaaaae-
afaaaaaaaaaaaaaaaaacaaaaaaaaaaaaaaaacaabaaaaabaaa?aaaaaaa?aa??
aaaaaabaaaaaaaaaaaaaaaaaaaaadaaaaaaaaaaaaaaaaaaaacaaaaabaaaaaaaaaaaaa
L770nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
L1075nS1W2D2 aaaaabaaacaaaaaaaaha?aaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaa?a?baaaaaa?
aaa??????????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
1079
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaabaaaaaaabaaaaaaaaaaaaaaaaaaaaa
aaabaaaaaaabaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaaaaaaaaabbaaaaabaaaaaaaaaaaaa
2411 aaaaaaaaaaaaaaaaaaaa?aaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaacaa?a?baaaaaa?
aaaaaabaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaabaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaaaaaaaaabacaaabaaaaaaaaaaaaa
L425nS1W2D2 aaaaabaaacaaaaaaaaaaaaaaaabaaaaaaaaacabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaabaaaaaa?
aaae?????????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
18 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L141nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaabbaaaaaa??????????????????????????????????
L5nS1W2D2
aaaaabaaacaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaabaaaaaaaaaae????????????????????????????????
??????????????????????????????????????????????????????????????????????????????????????????????????????????????????
L60nS1W2D2
aaaaabaaacaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaababaaaaaaaaae????????????????????????????????
??????????????????????????????????????????????????????????????????????????????????????????????????????????????????
2718 aaaaaaaaaaaaaaaaaaaa?aaaaaaaaaaaaagabaaaaaaaaaadaaaaaaaaaaabaaaaaaaaaaaaaaaaaa?a?babaaaa?aaa?aacdaaaaaabaaaa-
faaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaacaabaaaaaaaaaa?aadaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaaabaaaab-
baaaaaaaaaaaaa
038
aaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaa
aaaaaaabaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaaaaa??aaaaaacaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
2786 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaacaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?aaaaaaaaaabaaad-
baaaaaaaaaaaaaaaaaabaaaaaaaaaaaaabaacaaaaaaabaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L1086nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aabgaaaaecbaaaabaaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
L1000nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaabacaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaacaaaaaaaaaaa?aaaaaba?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
L638nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaacaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
2886 aaaaaaaaaaaaaaaabaaa?aaaabaaaaaaaaaabaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaabacaaaabaaaaaaaaaaaaaaaaaaaacbaabaaabaaaaaaaaaaaacaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
045 aaaaaaaaaaaaaaaaabaaacaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaa?a?baaaaaaaaaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaacaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaa
L1075nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????aal-
gaaaaecbaaaabaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaacaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
884 ?????????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????????????????aa?aaaa?aa??aaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
2192 aaaaaabaaacdcccdedfa?aaaabaaaaaaaaaaaabaaaaaaaaaaaiaaaaaaaaaaabaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaacadaaaaaaaabaaabaabaaaaaaaacaabaaaacaaaa?aabaaaa?aa??
abaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaacaabcaaaabbaaaaaaaaaaaaa
397 aaaaaaaaaaaaaaaaaaaa?aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?aaaaaaa?
```

```
aaabaagaaaaaaabaaaaaaaaaaaaaaaaaaaaaaccaabaaaaaaaaaaaaabaacaaaaaaadbaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
1293 aaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaeabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaeaajdaaaaabbaaad-
baaaaaaaaaaaaaaacaabbaaabaabaaaaabaabaaaaaaadaaaa?aaaaaaa?aa??
aaaaaaaaabaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
L1552nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaabaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaabaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
1128 ????????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
??????????????????????????????????????????????aa??aaaaaa????aaaaaaaaaaaaaabaaaaaaaaabaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L1091nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaacaaaaabaaaaaa??????????????????????????????????
L847nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaabaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
L252nS1W2D2 aaaaabaaacabaaaaaaaa?aaaabaaaaaaaaaacabaaaabaaaaaabaaaaaaaaaaabaaaaaaaaaaaaacaa?a?
baaaaaaaaaah?????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????????????
157 aaaaaaaaaaaaaaaaabaaa?aaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaabaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaacaaabaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
0211 aaaaaaaaaaabaaaaabaaaaaaaaaaaaaaabaaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaae-
aaidaaaaaabaaaabaaaaaaaaaaaaaaaaaaabaaaaaabbaaaaaaacaaaaaaaaaaaa?aaaaaba?aa??
aaaabaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L425nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
L1076nS1W2D2 aaaaabaaacaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaaa?
aaae?????????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
346 aaaaaaaaababaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaiaaaaaaaaaaabaaaaaaaaaaaaaaaaaaababaaaa?aaa?aaaaaaaaaabaaaa-
faaaaaaaaaaaaaaaaaaacaaabaaaaaaaaaaaacaaaaaaadaaaa?aabaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaeaaaaaaaaaaaaaaaaabaaaab-
baaaaaaaaaaaaa
1654 aaaaaaaaaaaaaaaaabaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaabaaabaaaaabaaaaaaaaaaaaa
430 aaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaaabaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?aabaaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaacaaaaaaaaaaaa?aabaaaa?aaa?
aaaaaaaaaaaaaaaaaaabaaaaaaaabaaaaaabaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
213 aaaaaaaaaaaaaaaaaaaaaaaabbaaaaaaaaaaaaaaaaaaaaaadaaaaaabaaaacaaaaaaaaaaaaaaa?aabaaaaaa?aaa?
aaaaaabaaabaaaaaabaaaaaaaaaaaaacaabaaabaabaaaaaaaadaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaa
807 aaaaaaaaaaaaaaaaabaaa?aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?babaaaa?aaa?
aaaaacaaaabaaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaacaabaaaaaaaaaa?aaaaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaa-
faaaaaaaaaaabaaaabbaaaaaaaaaaaaa
017 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaae-
aabaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaacaaaacaaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
799 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
69 aaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaafaaaaaaaaaaabaaaaaaaaaaaaaaaaaa?babaaaa?aaaeaaaaaaaaaaabaaaa-
faaaaaaaaaaaaaaaaaadaaabaabaaaaaaaaacaabaaaadaaaa?aabaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaaaaabaaeaaaaadaaaaaaaaaaabaaaab-
baaaaaaaaaaaaa
basetext
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
1463 aaaaaaaaaaaaaaaaaaaia?aaaabaaaaaaaaaaaabaaaaaaaaaaagaaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?aa-
haaaaaaabaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaaaaaaaaaabaaaaababaaaaaaaaaaa
2766 aaaaaaaaaaaaaaaaaaaba?aaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?babaaaa?
aaaaaaaaacaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaabaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabbaaaaabaaaaaacaaaaaa
032S aaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaadbaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaeaa-
facaaaaacaaaaaaaaaaaaaaaaaccaabaaabaaaaaababaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaacaaaaaaaabaaaabaaaaaaaaabaaabaaaaaaaaaaaaaaab-
baaaaabaaaaaaaaaaaaa
L32nS1W2D2 aaaaabaaacaaaaaaaaaaa?aaaabaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?
aaae?????????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
L329nS1W2D2 aaaaabaaacaaaaaaaaaaaaaaaaaaaaaaaaacabaaaaaaaaaaaahaaaaaaaaaaaaaaaaaaaaaaaaaaaa?
aaaaaaaaaaa??????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????????????
1219 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?
baaaaaaaaaaaaabaaaaaaabaaaaaaaaaaaaaaaaaaaaacaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
1230 aaaaaaaaaaabaaabaaaa?aaabaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?
aaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaacaaaaaaaaaaaaaaacaaaaaaaaaaaaaaaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L847nS1W2D2
aaaaabaaacaaaaaaaaaaaaaaabaaaaaaabaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaae????????????????????????????????
??????????????????????????????????????????????????????????????????????????????????????????????????????????????????
1241 aaaaaaaaaaaaaaaaaaaa?aaaaaaaaaaaaaaabaaaaaaaaaaadaaaaaaaaaaaaaaaaaaaaaaaaacaaaaababaaaa?
aaagaakaadaaaabaaaabaaaaaaaaaaaaaaacaaabaaabaaaaaaabaaacaaaaaaadaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaaabaaabaaaaaaaaaaaaaaaaaaa
L663nS1W2D2 aaaaabaaacaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?
aaa??????????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
033 aaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaacaaaaaaaaaaaaaaaaaaabaaaaaaaaaae-
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaccaabaaabaaaaaaaabaaaaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
P66 aaaaaaaaaaaaaaaabacaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaamaadaabaaaaaaaaae-
baaaaaaaaaaaabaacaaaaaaabaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaab
aaaaaa
109 aaaaaaaaabaaaaaaaaaaaaaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaabaabaaaaaaaaaaaaaadaa?a?baaaaaaaaaae-
aaaaaaaaaabaaaaaaaaaaaaaadaaaaaaaaacaaabaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?ab??
```

```
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
1009 aaaaaaaaaaaaaaaaaaeaaaaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaababaaaaaaaaae-
aaaaaaaaaabaaaaaaaaaaaaadaaaaaaiaaabaaabaaaaaaaaaaacaabaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaagaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
03 aaaaaaaaabaaaaaaaaaa?aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa?
aaaaaaaaaaaaaadaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
acaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
35 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
382 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?aababaaaaaabcdaaae-
acaaaabaaaabaaaaaabaaaaaaaaaaaacaaaaaacaaaaaaaacaaaaaaadaabaaaabaaba?ba?
aaaaabaaaaabaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
1242 aaaaaaaaaaaaaaaaaaaa?baaabaaaaaaaaaaaaaaaaaaaaaaaacaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?aaaaaaaaabbaaa-
faaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaaacaabbabaab-
baaaaaaaaaaaaa
1320 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L1552nS1W2D2
aaaaabaaacaaaaaaaaaaaaaaaabaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaaae????????????????????????????????
???????????????????????????????????????????????????????????????????????????????????????????????????????????????????
138 ??????????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????????????????????????aa??aaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaaaa
L1086nS1W2D2
aaaaabaaacaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaaae????????????????????????????????
???????????????????????????????????????????????????????????????????????????????????????????????????????????????????
1192 aaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaacaaabaaabaaaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
579 aaaaaaaaaaaaaaaabaaaaaaaaabaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaabaaababaaaaaaaaaaaaaaaaaaaaaaaakaaiaaaaaaa?
baaaaaaaaaaaaaaaaaaafcaabaaaaaaaaaaaaaaaaaaaaaaadaaaa?
aabaaaaaaaaaaaaaaaaaaaaaaaaaaaaababaaaaabaaaaaaaaaaaaaaaaaabadaaaaabaaaaaaaaaaaaa
1424 aaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaaaaaaagaaadaaaaab-
baaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaba?aa??
aaaabaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaabaaaaaaabaaaaabaaaaaaaaaaaaa
011 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaababaaaaaaaaaa?
aaaaacaaaabaaaabaaaaaaaaaaaaaaaacbaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaaaaa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaabbaaaaaaaaaaaaa
07 aaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaae-
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaabbaaaaaaaaaaaaa
2561 aaaaaaaaaaaaaaaacaaaaaaaabaaaaaaaaaaaaaaaaaaaaaabacaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?aaaaaaaaab-
baaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaa
994 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?aabaaaaaa?aaae-
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
1278C ???????????????????????aabaaaaaaaaaaaaaaaaa????????????????????????????????????????????????????????????????????????????
aaaaaaaaaabaaaaaaaaaaa???????????aaaaaa?aaaaaaa?????????????????????aaaaaaaaaaaabaaaaaa??????????????????????aaaaaaaaaaaa
028 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
1029 aaaaaaaabbaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaa?baaaaaaaaaaeaaidaaaaab-
baaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
abaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaabaaaaaaaaaaaaabaaabaaaaaaaaa
L1073nS1W2D2 aaaaabaaacaaaaaaaaaa?
aaaabaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaae?????????????????????????????????????????????????????
??????????????????????????????????????????????????????????????????????????????????????????????
2223 aaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaaaaaaa-
haaaaaaaaaabaaaaaaaaaaaaaaaaaaaacaaacaaaaaabaaaaaaaacaaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaacaaaaabaaaaaaaaabaaa
L252nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????????????????????dcaib-
caaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaabaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
L253nS1W1D3
aaaaabaaacaaaaaaaaaaaacaaabaaaaaaaacbhbccbcbdaaaaaajaaaaaaaaaaabaaaaaaaaaaaaaaaaababaaaaaaaaaa????????????????????????????????
???????????????????????????????????????????????????????????????????????????????????????????????????????????????????
019
aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagaaaaaaabaaaabaaaaaaaaaaaaaaac
caabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
L640nS1W2D2 aaaaabaaacaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?
aaa???????????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
?????????????????????????
2713 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaagaaaaaaaaaaaaabaaaaaaaaaaaaaaaabaaaaaaaaaaa?a?babaaaa?aaa?
aaaaacaaaabaaaaaaaaaaaaaaaaaaaacaaabaaabaaaaaaaaaaaacaaaaaaadaaaa?aabaaaa?aaa?
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
13 aaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaiaaaaaaaaaaabaaaaaaaaaaaaaaaa?aababaaaa?aaa?aabaaaaaaabaaaa-
faaaaaaaaaaaaaaaaaacaaabaaaaaaaaaaaacaaaaaaadaaaa?aabaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaeaaaaaaaaaaaaaaaabaaaab-
baaaaaaaaaaaaa
377 aaaaaaaaaaaaaaaabaaa?aaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?babaaaa?aaa?
aadaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaacaabaaaaaaaaaa?aaaaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaa-
faaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L387nS1W2D2
aaaaabaaacaaaaaaaahaaaaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaaabaaaaaaaaaae????????????????????????????????
???????????????????????????????????????????????????????????????????????????????????????????????????????????????????
022 ????????????????????????????????????????????
aaaaakaaaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaaaaaaaaaaabaaaabaaaaaaaaaaaaaaaacaaabaaaaaaaaaaaaaaaacaaaaaaadaaaaaaabaaaaaaa?
aaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L141nS1W2D2
aaaaabaaacaaaaabaagaaaaaabaaaaaaaaaaabaabaaaaaaaaabaaaaaaaaaaabbaaaaaaaaaaaacaaaaabaaaaaaaaaae????????????????????????????????
???????????????????????????????????????????????????????????????????????????????????????????????????????????????????
2615 aaaaaaaaaaaaaaaaaaaa?aaaaaaaaaaaaagaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aakaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaabaaaaaaaaacaaabaaaaabaaaaaaaaaaaaa
565 aaaaaaaaaaaaaaaaaaaa?aaaabaaaaaaaaaaabaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaaaaaaaacaaab-
```

```
baaaaaaaaaaaaaaaaaaaacbaabaaabaaaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaab-
baaaaabaaaaaaaaaaaaa
1546 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaababaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaae-
aabaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaabaaaaaaaaaacaabbaaaaabaaaaaacaaaaaa
841 aaaaaaaaaaaaaaaacaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaabaacaaaaabaaaaaa?aaae-
aaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaabaaaaaaaaaaaabaacaabaaaaaaaaaa?aaaaaba?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaabaacaaaaaaaaaaaaaaaabbaaaaabaaaaaaaaaaaaa
828 aaaaaaaaabaaaaaaaaaaacaaaaaaaaaaaaaabaaaaaaaaaaafaaaaaaaaaabaaaaaaaaaaaaaaaaaababaaaaaaaa?aaaaaaaaaabaaaa-
faaaaaaaaacaaaaaaaaacaaabaaaaaaaaaaacaabaaaadaaaa?aabaaaaaaa??aaaaaaaaaaaaaaaaaaaaaaaaaaabaaeaaaaaaaaaaaaaaaaabaaaab-
baaaaaaaaaaaaa
L17nS1W1D6 ????????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaacaaaaaaaaaaaa?aaaaaba?aa??
aaaabaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
124 aaaaaaaaabaaaaaaaaaaaaaaaabaaaaaaaaababaaaaaaaaaaafaaaaaaaaaaaaaaaaaaaaaaaaaaaa?babaaaaaaaaeaaadacaaaabaaaa-
faaaaaaaaaaaaaaaaaacaaabaaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aaa?aaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaabaaaab-
baaaaaaaaaaaaa
P66C ??????????aaaaaaaaaaaa???????????????????????????????????????????aaaaaaaaaaaaaaaaaaaaaaaaalaaaaaaaaaaaaaaba????????
aaaaaaccaaa????????????????????????????????????aaaaaaaaaaaaaa??????aaaaaaaaaaaaaaaaaa?????????????????????aaaaaaaaaaa
L735nS1W2D2
aaaaabaaacaaaaaabaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaaaabaaaaaaaaaae??????????????????????????????
??????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
2680 aaaaaaaaabaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaafaaaaaaaaaacaaaaaaaaaaaaaaaaa?baaaaaa?
aaaaaaaaaaaaaabaaaaaabaaaaaaaaaaaaacaaacaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??aaaaaaaaaaaaaaaaaaaaaaaaaaaaabaae-
aaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaaa
050 ???????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
??????????????????abaacaaaaaaaaaaaa?aaaaaaa?aaa?aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabbaaaaaaaaaaaaaaaaaaa
792 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaabaaaabaaaaaaaababaaaaaaaaaabaaaaaa?aaae-
aaaaaaaaaabaaadaadaaaaaabaaaaaaaaaacaaaaaabaaaaaaaaaaaaaaaaaaaaa?aabaaba?aa??
aaaabaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaabaaaabbaaaaaaaaaaaaa
01 aaaaaaaaaaaaaaaaaaaabdaaabaaaaaaaaaadaaaaaaaaaaaaahbaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaababaaababjaahaaaaaaaaaaaacaaeaaaaaaaaaab-
daaeaaacaaaaaaaaaaaaaaaaaaaaaaaaaaaa?aabaaaa?aa??aabaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaabaaabaaaaaaaaaaaab
0127 ?????????????a?a??aaaaaaaaaaaaaaaaadaaaaaaaaaaaacaaaaaaaaaabbaaaaaaaaaaaaaaaaaa?aaaaaaa????
q??????????????????????????????????????????????????????????????????????????????????????????????????????????????????????
??????????????????????
L335nS1W1D6 ?????????????????????????????????????????????????????????????????????????????????????????aamgaaaaecbaaaa-
faaaaaaaaaaaaaaaaaaacaaabaaaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
1278 aaaaaaaaaaaaaaaaaaaa?aaaa?aaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaabaaa?aaaaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaabaaaaaaaaaaaaa
L5nS1W1D6 ??????????????????????????????????????????????????????????????????????????????????????????????????????
aaagaaaaecbaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aaa?
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaa??????????????????????????????????
732 aabcbacbcacdcccdedkbcfbcddcaaacceaahac?bcbabcbbcbmdcdcccdddadbdcdcaaaaaaaaaaaa?aaaabaaaa?aaa?aaaac-
caaaabaaaadaaaaaaaaaadaaaagaaabaaabaaaaaaaaaaaaaaabaaaadaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaacaaacaaaaabaaaaaaaaaaaaaaaacaaaaaaaabbaaaaaaaabaaaa
024 ????????????????????????????????????????????????????????????????????????????????????????????????????????????
aaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaaaaa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaabaaaaaba????a???????
1797 aaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaadaabaaaaaaaaaaaeaai-
aaaaaaabaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaacabaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaabaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaacaaaaaaabaaaaabaaaaaaaaaaaaa
892 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaadaaaaaaaaaaaaadaaaaaaaaaaaaaaaaaaaaaaaaaaa?a?baaaaaa?aaa?
aaaaaaaaaabaaaaaaaaaaaaaaaaaaaacaaabaaaaaaaaaaaaaaaacaaaaaaaaaaaa?aaaaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaabaaabaaaaabaaaaaaaaaaaaa
2106 aaaaaaaaaaaaaaaabagaaaaaaaaaaaaaaaaaabaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaa?aaaaaaa?aabababaaaa?aaa?
aaaaaaaababaaaadaaaaaaaaaaaaaaaeaaabaaabaabaaaaaaaaacaabaaaadaaaa?aaaaaaa?aa??aaaaaaaaaaaabaaaaabaaaaaaaaabaaaaaaaahacbdcbadb-
cdhceccdfcdcdccecedcdd
1071 aaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaeaaaaaaaaaaaaaaaaaaaaaaaaaaaaaababaaaaaaaaae-
aadaaaaaaabaaaabaaaaaaaaaaaaaaaaaaabaaabaaaaaaaabaacaabaaaaaaaaaa?aaaaaaa?aa??
aaaabaaaaaaaaaaaaaaaaaaaaaaabaaaaaabaaaaaaaaabaaaaaaaaabaaaaaaaaaaaaa
168 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaababaaaaaaaaaaaabaaaaaaaaaaabaaaaaaaaaaaaaaaa?aabaaaaaa?
aaaaaaaaaaaaaabaaaaaaaaaaaaaaaaaaaaaaaacaaaaaabaaaaaaaacaaaaaaaaaaaa?aabaaaa?aa??
aaaaaaaaaaaaaaaaaaaaaaaaaaaabaaaaaaaaaaaaaaaaacaabaaaaaabaaaaaaaaaaaaa
;
END;
```

Appendix 4: Coherence Data for Chapter 7

The following genealogical coherence tables constitute supporting data for Chapter 7.

Genealogical coherence with W1=032

W2	**NR**	**D**	**PERC1**	**EQ**	**PASS**	**W1<W2**	**W1>W2**	**UNCL**	**NOREL**
017C		-	97.561	40	41				1
892SC	1		95.652	66	69	2	1		
01Ccb2		-	92.727	51	55	1	1	1	1
037C		-	92.424	61	66	1	1	2	1
A	2		91.237	4144	4542	322		67	9
019C	3		90.385	47	52	2	1	2	
029	4		89.281	683	765	34	15	32	1
𝔓75	5		88.381	2708	3064	145	70	131	10
03	6		88.196	4386	4973	305	35	235	12
019	7		87.808	4278	4872	265	107	203	19
04	8		87.737	1760	2006	109	46	83	8
33	9		87.247	4262	4885	215	175	213	20
044	10		86.924	4321	4971	231	185	205	29
0290	11		84.848	476	561	29	27	26	3
1571	12		82.564	161	195	14	9	7	4
086		-	50.000	6	12			6	
032S		-	46.667	7	15			8	
083		-	46.154	6	13			7	

Table 149: Genealogical coherence with W1=032

Genealogical coherence with W1=249

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
017C		-	100.000	24	24				
892SC		-	100.000	17	17				
869	1		97.286	2079	2137	23	20	11	4
07C		-	96.491	55	57	1	1		
1797C	2		96.429	54	56	2			
019C		-	96.000	24	25			1	
01Ccb2	3		95.238	40	42	1			1
037C		-	94.286	33	35			2	
2192C	4		94.175	97	103	3	1	1	1
397	5		92.648	3415	3686	117	74	63	17
MT	6		92.431	3407	3686	98	88	77	16
1320			92.431	3407	3686	98	88	77	16
226	8		92.274	3416	3702	105	84	79	18
35	9		92.220	3414	3702	105	89	77	17
18	10		92.151	3393	3682	106	90	76	17
892	11		91.576	2620	2861	110	69	45	17
1571	12		91.512	690	754	29	27	7	1
028	13		91.493	3377	3691	108	101	88	17
045	14		91.455	3382	3698	110	108	84	14
063	15		91.388	573	627	21	19	13	1
07	16		91.320	3377	3698	110	106	89	16
1561	17		91.188	3363	3688	111	103	90	21
0290		-	90.909	10	11			1	
FΠ	18		90.869	3334	3669	143	80	97	15
821	19		90.842	3323	3658	136	89	92	18
0141	20		90.822	3325	3661	135	92	91	18
799	21		90.491	3264	3607	128	102	94	19
157	22		90.488	3320	3669	128	111	92	18
044	23		90.465	3349	3702	186	56	99	12
1321	24		90.439	3320	3671	123	111	91	26
083	25		90.432	293	324	17	7	5	2
597	26		89.875	3320	3694	152	106	93	23
994	27		89.415	3269	3656	133	132	105	17
033	28		89.367	1992	2229	95	55	70	17
2786	29		89.304	3298	3693	158	118	95	24
33	30		89.241	3235	3625	187	76	109	18
F1	31		89.142	3292	3693	177	93	115	16
02	32		89.132	2567	2880	128	74	97	14
213	33		89.064	3274	3676	156	129	98	19
A	34		88.731	2992	3372	310		48	22
022	35		88.529	1860	2101	96	79	53	13
04	36		88.344	1296	1467	97	26	37	11
029	37		87.990	740	841	47	20	30	4
019	38		87.693	3242	3697	242	55	135	23
𝔓66C*	39		87.692	342	390	14	10	20	4
01Cca	40		87.640	390	445	18	9	23	5
070	41		86.435	771	892	50	33	31	7
03	42		86.078	3178	3692	289	25	176	24
892S	43		85.940	599	697	29	25	37	7
𝔓75	44		85.731	2986	3483	244	73	151	29
032S	45		85.551	450	526	34	31	5	6
032	46		85.095	2649	3113	184	108	142	30
213S		-	83.333	10	12			2	
05S		-	83.333	10	12			2	
𝔓66	47		82.397	2874	3488	218	196	162	38
01	48		81.102	2974	3667	236	220	202	35
𝔓45	49		78.894	471	597	37	33	49	7
04C1	50		75.000	33	44	5	3	2	1
𝔓60		-	75.000	9	12			3	
118S		-	50.000	6	12			6	
054		-	50.000	6	12			6	

Table 150: Genealogical coherence with W1=249

Genealogical coherence with W1=333

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
249	1		98.620	3644	3695	16	10	21	4
07C		-	98.198	109	111	1	1		
892SC		-	97.590	81	83	1	1		
037C	2		96.296	78	81	1		2	
869	3		96.161	2054	2136	33	24	22	3
1797C	4		95.876	93	97	2	1	1	
017C		-	95.522	64	67	1	1	1	
2192C	5		94.886	167	176	4	1	3	1
18	6		93.176	5789	6213	159	123	116	26
35	7		93.133	5805	6233	159	124	119	26
MT	8		93.127	5786	6213	154	130	119	24
1320			93.127	5786	6213	154	130	119	24
226	10		93.089	5792	6222	158	126	117	29
01Ccb2	11		93.056	67	72	1		1	3
397	12		92.569	5668	6123	192	130	107	26
028	13		92.366	5747	6222	163	145	141	26
045	14		92.213	5743	6228	174	160	129	22
07	15		92.130	5736	6226	172	158	134	26
1561	16		92.038	5722	6217	169	158	138	30
892	17		91.484	3427	3746	138	86	74	21
FΠ	18		91.435	5669	6200	223	126	154	28
799	19		91.272	5595	6130	195	156	152	32
1321	20		91.271	5657	6198	196	162	139	44
597	21		91.171	5669	6218	215	154	142	38
821	22		90.993	5597	6151	220	151	145	38
0141	23		90.970	5601	6157	218	158	144	36
0290	24		90.827	505	556	15	11	21	4
157	25		90.805	5629	6199	201	178	158	33
044	26		90.581	5645	6232	299	102	159	27
1582C1	27		90.286	158	175	4	3	4	6
054	28		90.054	1005	1116	39	38	30	4
02	29		89.880	4858	5405	223	142	153	29
33	30		89.490	5475	6118	298	136	171	38
213	31		89.412	5075	5676	230	183	157	31
1071	32		89.377	5553	6213	235	233	155	37
033	33		89.375	3945	4414	191	113	132	33
F1	34		89.161	5544	6218	282	178	183	31
086	35		88.817	413	465	24	18	10	
A	36		88.747	5087	5732	530		73	42
2786	37		88.689	5512	6215	247	244	164	48
022	38		88.625	3397	3833	169	151	92	24
04	39		87.853	2177	2478	165	47	74	15
019	40		87.620	5457	6228	411	95	219	46
𝔓66C*	41		87.480	545	623	22	18	29	9
01Cca	42		87.145	705	809	28	20	43	13
083	43		86.398	578	669	44	27	15	5
029	44		86.368	811	939	60	23	42	3
03	45		86.164	5362	6223	495	49	274	43
05S	46		85.660	448	523	24	23	24	4
𝔓75	47		85.573	3731	4360	310	101	185	33
𝔓60	48		85.288	400	469	20	17	29	3
070	49		85.288	771	904	54	34	37	8
032S	50		84.644	915	1081	68	64	23	11
032	51		84.643	3996	4721	275	181	219	50
𝔓66	52		82.614	4709	5700	350	319	263	59
01	53		81.310	5016	6169	395	390	310	58
04C1	54		78.409	69	88	8	5	4	2
𝔓45	55		77.181	460	596	37	33	57	9

Table 151: Genealogical coherence with W1=333

Genealogical coherence with W1=430

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
892SC		-	100.000	1	1				
892S*		-	100.000	1	1				
017C	1		95.833	23	24	1			
019C		-	95.238	20	21			1	
𝔓45	2		94.828	55	58	1		2	
249	3		94.667	1580	1669	51	21	16	1
869	4		93.792	695	741	24	11	11	
2192C	5		93.750	45	48	2			1
333	6		93.739	2096	2236	71	36	28	5
1797C	7		92.308	36	39	2		1	
MT	8		91.084	2033	2232	96	54	41	8
1320			91.084	2033	2232	96	54	41	8
028	10		91.002	2043	2245	96	56	43	7
226	11		90.970	2045	2248	99	53	42	9
35	12		90.925	2044	2248	104	52	41	7
045	13		90.865	2039	2244	98	60	40	7
18	14		90.799	2023	2228	104	53	40	8
07	15		90.302	2030	2248	104	65	42	7

W2	NRD	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
1571	16	90.208	1170	1297	63	47	13	4
086	17	89.936	420	467	27	16	4	
FΠ	18	89.928	2009	2234	122	51	43	9
037	19	89.750	2014	2244	109	62	47	12
157	20	89.730	1992	2220	106	66	46	10
01Ccb2	-	89.474	17	19				2
892	21	89.413	2010	2248	135	58	34	11
799	22	89.340	1936	2167	114	60	50	7
1561	23	89.267	1996	2236	108	70	51	11
397	24	89.242	1991	2231	125	55	47	13
1321	25	89.149	1980	2221	102	76	47	16
011	26	89.028	1996	2242	104	81	47	14
2790	27	88.945	1593	1791	82	69	35	12
2615	28	88.929	1984	2231	111	79	46	11
0211	29	88.924	1983	2230	100	79	54	14
1014	30	88.804	1618	1822	82	70	42	10
063	-	88.798	975	1098	48	48	24	3
044	31	88.785	1995	2247	154	34	55	9
992	32	88.770	1992	2244	110	86	45	11
02	33	88.756	1555	1752	112	45	35	5
1463	34	88.724	1975	2226	107	87	47	10
1242	35	88.675	1981	2234	116	78	49	10
1319	36	88.651	1984	2238	108	87	48	11
2766	37	88.568	1991	2248	99	92	55	11
1230	38	88.462	1978	2236	109	91	43	15
2680	39	88.414	1984	2244	99	97	51	13
821	40	88.174	1961	2224	138	59	53	13
0141	41	88.056	1961	2227	138	62	53	13
357	42	87.993	1942	2207	118	87	52	8
2768	43	87.968	1974	2244	101	99	54	16
2561	44	87.952	1898	2158	125	74	54	7
865	45	87.904	1279	1455	69	63	36	8
597	46	87.873	1971	2243	136	75	46	15
33	47	87.783	1940	2210	148	58	50	14
2223	48	87.684	1965	2241	105	104	52	15
994	49	87.631	1757	2005	108	83	50	7
2411	50	87.562	1929	2203	114	98	53	9
1293	51	87.517	1956	2235	112	100	46	21
037C	-	87.500	21	24	1	1	1	
544	52	87.455	1938	2216	105	104	61	8
138	53	87.447	1658	1896	98	84	48	8
1788	54	87.408	1902	2176	114	98	52	10
0233	55	87.385	568	650	39	34	6	3
1654	56	87.271	1954	2239	106	104	55	20
2575	57	87.267	1871	2144	110	102	54	7
1128	58	87.253	1807	2071	97	90	61	16
213	59	87.222	1959	2246	141	86	50	10
2786	60	87.210	1957	2244	130	91	52	14
F1	61	87.199	1955	2242	144	79	52	12
022	62	87.083	1591	1827	122	61	44	9
038	63	87.011	1956	2248	127	92	61	12
1071	64	86.943	1951	2244	124	106	51	12
A	65	86.670	1814	2093	236		27	16
1010	66	86.598	1932	2231	123	104	51	21
04	67	86.331	840	973	81	22	22	8
033	68	86.234	664	770	49	22	31	4
083	69	86.212	569	660	51	26	9	5
𝔓66C*	70	86.154	168	195	12	3	8	4
019	71	85.975	1931	2246	187	44	64	20
1241	72	85.881	1916	2231	120	117	58	20
03	73	85.076	1904	2238	215	23	80	16
029	74	84.411	666	789	62	19	37	5
032S	75	84.235	903	1072	77	62	18	12
01Cca	76	84.227	267	317	19	6	17	8
𝔓75	77	84.146	1863	2214	197	60	79	15
032	78	82.967	906	1092	82	42	50	12
070	79	82.840	140	169	16	6	5	2
𝔓66	80	82.059	1674	2040	166	102	73	25
0290	-	81.818	9	11			2	
04C1	-	75.000	18	24	2	2	1	1
213S	-	75.000	9	12			3	
05S	-	75.000	9	12			3	
109C	81	71.053	27	38	5	4	1	1
𝔓60	-	66.667	8	12			4	
317	-	46.154	6	13			7	
892S	-	41.667	5	12			7	
118S	-	41.667	5	12			7	
054	-	41.667	5	12			7	

Table 152: Genealogical coherence with W1=430

Genealogical coherence with W1=869

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
892SC		-	100.000	12	12				
086		-	100.000	12	12				
017C		-	100.000	11	11				
1797C	1		97.143	34	35	1			
019C		-	94.737	18	19			1	
01Ccb2	2		94.595	35	37	1			1
397	3		92.636	1975	2132	65	38	42	12
1571		-	92.308	12	13			1	
083		-	92.308	12	13			1	
032S		-	92.308	12	13			1	
2192C	4		91.429	64	70	4	1		1
0290		-	90.909	10	11			1	
892	5		90.817	1345	1481	52	44	29	11
226		-	90.454	1933	2137	68	68	54	14
35	6		90.360	1931	2137	70	68	53	15
18	7		90.316	1912	2117	70	68	52	15
044	8		89.752	1918	2137	103	42	64	10
063	9		89.744	70	78	3	1	4	
FΠ	10		89.734	1888	2104	87	63	56	10
168C		-	89.583	43	48	1	1	3	
821	11		89.357	1889	2114	82	63	67	13
0141	12		89.252	1885	2112	82	67	65	13
799	13		89.213	1894	2123	88	66	62	13
1321	14		88.873	1885	2121	78	74	66	18
597	15		88.821	1891	2129	86	73	62	17
033	16		88.800	1546	1741	84	50	46	15
33	17		88.793	1854	2088	108	39	76	11
A	18		88.543	1708	1929	169		36	16
F1	19		88.498	1885	2130	101	54	78	12
04	20		88.386	723	818	54	10	25	6
029	21		88.190	687	779	39	15	33	5
022	22		88.164	1162	1318	52	49	47	8
2786	23		87.981	1874	2130	97	81	64	14
865	24		87.853	1526	1737	83	63	51	14
213	25		87.849	1858	2115	94	88	63	12
019	26		87.576	1868	2133	135	30	85	15
01Cca	27		87.549	225	257	11	7	12	2
037C		-	87.500	14	16			2	
𝔓66C*	28		86.977	187	215	9	5	13	1
02	29		86.950	1146	1318	56	49	57	10
2561	30		86.760	1802	2077	91	84	80	20
03	31		85.915	1836	2137	163	6	116	16
𝔓75	32		85.673	1764	2059	140	38	100	17
032	33		85.430	1800	2107	128	65	96	18
070	34		84.965	599	705	45	27	28	6
038	35		84.792	1812	2137	107	106	91	21
892S		-	84.314	473	561	25	25	30	8
213S		-	83.333	10	12			2	
05S		-	83.333	10	12			2	
1242C		-	82.796	77	93	4	4	5	3
𝔓66	36		81.850	1637	2000	116	113	110	24
01	37		80.746	1711	2119	135	102	148	23
𝔓45	38		77.401	411	531	36	31	46	7
𝔓60		-	75.000	9	12			3	
04C1		-	66.667	14	21	2	2	2	1
118S		-	50.000	6	12			6	
054		-	50.000	6	12			6	

Table 153: Genealogical coherence with W1=869

Appendix 5: Chapter 7 John 18 Experiment

Some tables in Chapter 7's John 18 experiment were abridged. This appendix contains the unabridged data.

Pre-Genealogical Coherence for W1=032

IGNTP Data

W2	NR	PERC1	EQ	PASS
𝔓108	1	100.000	10	10
𝔓52		100.000	9	9
2418	3	94.828	55	58
1803		94.828	55	58
1143	5	93.220	110	118
𝔓90	6	92.857	13	14
1590	7	91.346	95	104
A	8	90.763	226	249
1804	9	90.698	78	86
2649	10	89.423	93	104
2634	11	89.091	49	55
2307	12	89.062	114	128
1349	13	88.889	120	135
475s		88.889	24	27
546	15	88.679	141	159
748	16	88.281	113	128
96	17	88.000	154	175
2717	18	87.975	139	158
1571	19	87.919	131	149
405	20	87.861	152	173
2722	21	87.817	173	197
2372		87.817	173	197
179	23	87.778	158	180
NA27	24	87.640	234	267
940	25	87.293	158	181
1071	26	87.266	233	267
03		87.266	233	267
947	28	87.097	27	31
731s		87.097	27	31
1400	30	87.050	121	139
274	31	86.957	140	161
711	32	86.875	139	160
657	33	86.765	177	204
2632	34	86.700	176	203
33	35	86.692	228	263
2398	36	86.667	143	165
649	37	86.567	58	67
233	38	86.466	230	266
1421	39	86.463	198	229
1816	40	86.364	228	264
930	41	86.344	196	227
1558	42	86.207	200	232
019	43	86.142	230	267
2524	44	85.932	226	263
904	45	85.827	109	127
926	46	85.784	175	204
1273	47	85.772	211	246
669	48	85.768	229	267
235		85.768	229	267
1566		85.768	229	267
139		85.768	229	267
1006		85.768	229	267
044		85.768	229	267
01		85.768	229	267
685	55	85.714	228	266
2613		85.714	228	266
1451		85.714	228	266
1192		85.714	228	266
114		85.714	228	266
2726		85.714	12	14
776	61	85.660	227	265
974	62	85.606	226	264
2907	63	85.567	166	194
419	64	85.551	225	263
04	65	85.526	195	228
1638	66	85.437	176	206
1564	67	85.433	217	254
1207	68	85.425	211	247
951	69	85.393	228	267
887		85.393	228	267
871		85.393	228	267
774		85.393	228	267
758		85.393	228	267
2902		85.393	228	267
2721		85.393	228	267
2472		85.393	228	267
1819		85.393	228	267
1699		85.393	228	267
1563		85.393	228	267
1544		85.393	228	267
1365		85.393	228	267
1278		85.393	228	267
1057		85.393	228	267
041		85.393	228	267
944	85	85.338	227	266
941		85.338	227	266
697		85.338	227	266
66		85.338	227	266
584		85.338	227	266
396		85.338	227	266
2728		85.338	227	266
215		85.338	227	266
195		85.338	227	266
1676		85.338	227	266
1604		85.338	227	266
1581		85.338	227	266
1392		85.338	227	266
112		85.338	227	266
1080		85.338	227	266
342	100	85.317	215	252
87	101	85.283	226	265
479		85.283	226	265
352		85.283	226	265
1338		85.283	226	265
1312		85.283	226	265
1229		85.283	226	265
0109	107	85.246	52	61
650	108	85.227	225	264

Pre-Genealogical Coherence for W1=032

W2	NR	PERC1	EQ	PASS
278		85.227	225	264
1142		85.227	225	264
2467	111	85.204	167	196
333	112	85.200	213	250
376	113	85.185	207	243
969	114	85.171	224	263
760		85.171	224	263
511	116	85.115	223	262
1306	117	85.081	211	248
2562	118	85.057	222	261
1388		85.057	222	261
2567	120	85.034	125	147
MTJ18	121	85.019	227	267
991		85.019	227	267
975		85.019	227	267
961		85.019	227	267
943		85.019	227	267
932		85.019	227	267
924		85.019	227	267
764		85.019	227	267
728		85.019	227	267
699		85.019	227	267
489		85.019	227	267
414		85.019	227	267
406		85.019	227	267
317		85.019	227	267
298		85.019	227	267
2686		85.019	227	267
260		85.019	227	267
2394		85.019	227	267
2346		85.019	227	267
226		85.019	227	267
22		85.019	227	267
2182		85.019	227	267
2122		85.019	227	267
210		85.019	227	267
190		85.019	227	267
1628		85.019	227	267
1620		85.019	227	267
160		85.019	227	267
1596		85.019	227	267
1586		85.019	227	267
1584		85.019	227	267
1535		85.019	227	267
150		85.019	227	267
1466		85.019	227	267
1448		85.019	227	267
1423		85.019	227	267
1346		85.019	227	267
1230		85.019	227	267
1220		85.019	227	267
1219		85.019	227	267
1141		85.019	227	267
1042		85.019	227	267
100		85.019	227	267
017		85.019	227	267
345	165	85.000	221	260
78	166	84.962	226	266
696		84.962	226	266
583		84.962	226	266
410		84.962	226	266
402		84.962	226	266
276		84.962	226	266
269		84.962	226	266
2549		84.962	226	266
2463		84.962	226	266
2400		84.962	226	266
2370		84.962	226	266
200		84.962	226	266
1693		84.962	226	266
1688		84.962	226	266
1627		84.962	226	266
1582		84.962	226	266
1444		84.962	226	266
1428		84.962	226	266
1367		84.962	226	266
1272		84.962	226	266
1179		84.962	226	266
1061		84.962	226	266
757	188	84.906	225	265
710		84.906	225	265
506		84.906	225	265
375		84.906	225	265
2502		84.906	225	265
2100		84.906	225	265
1779		84.906	225	265
1578		84.906	225	265
1136		84.906	225	265
1008		84.906	225	265
1182		84.906	135	159
1092		84.906	135	159
2774	200	84.848	224	264
158		84.848	224	264
1172s		84.848	224	264
1158		84.848	224	264
2442		84.848	196	231
2636	205	84.825	218	257
552	206	84.791	223	263
515		84.791	223	263
2405		84.791	223	263
2290s	209	84.766	217	256
1560	210	84.762	178	210
1556	211	84.674	221	261
1145	212	84.646	215	254
958	213	84.644	226	267
955		84.644	226	267
942		84.644	226	267
896		84.644	226	267
890		84.644	226	267
875		84.644	226	267
864		84.644	226	267
852		84.644	226	267
844		84.644	226	267
843		84.644	226	267
831		84.644	226	267
824		84.644	226	267
823		84.644	226	267
759		84.644	226	267
746		84.644	226	267
698		84.644	226	267
689		84.644	226	267
597		84.644	226	267
592		84.644	226	267
586		84.644	226	267
549		84.644	226	267
547		84.644	226	267
536		84.644	226	267
521		84.644	226	267
510		84.644	226	267
480		84.644	226	267
43		84.644	226	267
399		84.644	226	267
37		84.644	226	267
363		84.644	226	267
35		84.644	226	267
324		84.644	226	267
291		84.644	226	267
2863		84.644	226	267
285		84.644	226	267
2806		84.644	226	267
2756		84.644	226	267
2730		84.644	226	267
2665		84.644	226	267
2604		84.644	226	267
2554		84.644	226	267
2522		84.644	226	267
2520		84.644	226	267
2508		84.644	226	267
2503		84.644	226	267
2466		84.644	226	267
2439		84.644	226	267
2382		84.644	226	267
2362		84.644	226	267
2323		84.644	226	267
2321		84.644	226	267
232		84.644	226	267
2296		84.644	226	267
2273		84.644	226	267
2260		84.644	226	267
2249		84.644	226	267
2229		84.644	226	267
2221		84.644	226	267

W2	NR	PERC1	EQ	PASS
2204		84.644	226	267
2176		84.644	226	267
2175		84.644	226	267
214		84.644	226	267
2131		84.644	226	267
211		84.644	226	267
2109		84.644	226	267
2097		84.644	226	267
183		84.644	226	267
18		84.644	226	267
1680		84.644	226	267
1667		84.644	226	267
166		84.644	226	267
1647		84.644	226	267
1637		84.644	226	267
1636		84.644	226	267
1618		84.644	226	267
1617		84.644	226	267
1600		84.644	226	267
1576		84.644	226	267
1572		84.644	226	267
1562		84.644	226	267
1561		84.644	226	267
1557		84.644	226	267
1550		84.644	226	267
155		84.644	226	267
1548		84.644	226	267
1510		84.644	226	267
1503		84.644	226	267
1497		84.644	226	267
1496		84.644	226	267
1494		84.644	226	267
1492		84.644	226	267
1491		84.644	226	267
1490		84.644	226	267
1487		84.644	226	267
1482		84.644	226	267
1477		84.644	226	267
1472		84.644	226	267
1469		84.644	226	267
1456		84.644	226	267
1427		84.644	226	267
141		84.644	226	267
1389		84.644	226	267
1341		84.644	226	267
134		84.644	226	267
1339		84.644	226	267
1322		84.644	226	267
1313		84.644	226	267
130		84.644	226	267
1292		84.644	226	267
127		84.644	226	267
1251		84.644	226	267
1235		84.644	226	267
1205		84.644	226	267
1197		84.644	226	267
1181		84.644	226	267
1157		84.644	226	267
1147		84.644	226	267
1146		84.644	226	267
1110		84.644	226	267
111		84.644	226	267
1085		84.644	226	267
1075		84.644	226	267
1073		84.644	226	267
1072		84.644	226	267
1056		84.644	226	267
105		84.644	226	267
1040		84.644	226	267
1032		84.644	226	267
1013		84.644	226	267
1		84.644	226	267
033		84.644	226	267
830	344	84.615	220	260
2437		84.615	209	247
960	346	84.586	225	266
953		84.586	225	266
945		84.586	225	266
865		84.586	225	266
773		84.586	225	266
750		84.586	225	266
715		84.586	225	266
700		84.586	225	266
582		84.586	225	266
5		84.586	225	266
482		84.586	225	266
478		84.586	225	266
448		84.586	225	266
431		84.586	225	266
394		84.586	225	266
284		84.586	225	266
2788s		84.586	225	266
263		84.586	225	266
234		84.586	225	266
2132		84.586	225	266
21		84.586	225	266
20		84.586	225	266
1791		84.586	225	266
1789		84.586	225	266
17		84.586	225	266
1652		84.586	225	266
1554		84.586	225	266
1551		84.586	225	266
1493		84.586	225	266
1470		84.586	225	266
1426		84.586	225	266
1370		84.586	225	266
1357		84.586	225	266
1324		84.586	225	266
1285		84.586	225	266
1225		84.586	225	266
514	382	84.581	192	227
784	383	84.556	219	259
782	384	84.528	224	265
591		84.528	224	265
286		84.528	224	265
2812		84.528	224	265
2767		84.528	224	265
2715		84.528	224	265
2687		84.528	224	265
2621		84.528	224	265
2388		84.528	224	265
2112		84.528	224	265
1530		84.528	224	265
1484		84.528	224	265
1422		84.528	224	265
1331		84.528	224	265
1248		84.528	224	265
1565	399	84.519	202	239
836	400	84.507	180	213
047	401	84.500	169	200
505	402	84.470	223	264
496		84.470	223	264
2747		84.470	223	264
2653		84.470	223	264
2245		84.470	223	264
1186		84.470	223	264
2179	408	84.434	179	212
968	409	84.411	222	263
719		84.411	222	263
2894		84.411	222	263
1431		84.411	222	263
120		84.411	222	263
2369	414	84.375	216	256
350	415	84.362	205	243
374	416	84.351	221	262
245		84.351	221	262
403	418	84.348	194	230
2414	419	84.293	161	191
53	420	84.291	220	261
986	421	84.270	225	267
98		84.270	225	267
938		84.270	225	267
923		84.270	225	267
845		84.270	225	267
83		84.270	225	267
806		84.270	225	267
797		84.270	225	267
769		84.270	225	267
765		84.270	225	267
763		84.270	225	267
754		84.270	225	267

Pre-Genealogical Coherence for W1=032

W2	NR	PERC1	EQ	PASS
726		84.270	225	267
72		84.270	225	267
717		84.270	225	267
708		84.270	225	267
703		84.270	225	267
7		84.270	225	267
662		84.270	225	267
653		84.270	225	267
575		84.270	225	267
564		84.270	225	267
556		84.270	225	267
530		84.270	225	267
512		84.270	225	267
507		84.270	225	267
504		84.270	225	267
49		84.270	225	267
461		84.270	225	267
447		84.270	225	267
439		84.270	225	267
387		84.270	225	267
386		84.270	225	267
329		84.270	225	267
3		84.270	225	267
29		84.270	225	267
2856		84.270	225	267
2773		84.270	225	267
277		84.270	225	267
2754		84.270	225	267
275		84.270	225	267
2714		84.270	225	267
2706		84.270	225	267
2701		84.270	225	267
2695		84.270	225	267
2689		84.270	225	267
2635		84.270	225	267
2624		84.270	225	267
2623		84.270	225	267
26		84.270	225	267
2571		84.270	225	267
2525		84.270	225	267
2515		84.270	225	267
2510		84.270	225	267
2507		84.270	225	267
2415		84.270	225	267
2407		84.270	225	267
240		84.270	225	267
2367		84.270	225	267
2356		84.270	225	267
2354		84.270	225	267
2352		84.270	225	267
2297		84.270	225	267
2261		84.270	225	267
2224		84.270	225	267
2213		84.270	225	267
2201		84.270	225	267
2193		84.270	225	267
219		84.270	225	267
2181		84.270	225	267
204		84.270	225	267
201		84.270	225	267
1705		84.270	225	267
1698		84.270	225	267
1695		84.270	225	267
1687		84.270	225	267
1686		84.270	225	267
167		84.270	225	267
1650		84.270	225	267
1634		84.270	225	267
1632		84.270	225	267
1619		84.270	225	267
1514		84.270	225	267
151		84.270	225	267
1501		84.270	225	267
15		84.270	225	267
1489		84.270	225	267
1476		84.270	225	267
147		84.270	225	267
1462		84.270	225	267
1452		84.270	225	267
1445		84.270	225	267
1441		84.270	225	267
1398		84.270	225	267
1396		84.270	225	267
1356		84.270	225	267
1355		84.270	225	267
1347		84.270	225	267
1328		84.270	225	267
1323		84.270	225	267
1320		84.270	225	267
1300		84.270	225	267
1296		84.270	225	267
1242		84.270	225	267
1234		84.270	225	267
1222		84.270	225	267
1210		84.270	225	267
121		84.270	225	267
12		84.270	225	267
1189		84.270	225	267
1165		84.270	225	267
1120		84.270	225	267
1117		84.270	225	267
1076		84.270	225	267
1062		84.270	225	267
1049		84.270	225	267
1034		84.270	225	267
1020		84.270	225	267
1018		84.270	225	267
1015		84.270	225	267
1003		84.270	225	267
030		84.270	225	267
996	543	84.211	224	266
987s		84.211	224	266
972		84.211	224	266
880		84.211	224	266
80		84.211	224	266
76		84.211	224	266
718		84.211	224	266
694		84.211	224	266
664		84.211	224	266
587		84.211	224	266
563		84.211	224	266
473		84.211	224	266
446		84.211	224	266
395		84.211	224	266
346		84.211	224	266
288		84.211	224	266
2758s		84.211	224	266
2732		84.211	224	266
271		84.211	224	266
2703		84.211	224	266
268		84.211	224	266
2661		84.211	224	266
251		84.211	224	266
247		84.211	224	266
2317		84.211	224	266
230		84.211	224	266
2278		84.211	224	266
2252		84.211	224	266
2141		84.211	224	266
175		84.211	224	266
1649		84.211	224	266
1570		84.211	224	266
1552		84.211	224	266
149		84.211	224	266
1480		84.211	224	266
1468		84.211	224	266
1455		84.211	224	266
1442		84.211	224	266
137		84.211	224	266
1359		84.211	224	266
1358		84.211	224	266
1289		84.211	224	266
1195		84.211	224	266
1191		84.211	224	266
1178		84.211	224	266
116		84.211	224	266
1123		84.211	224	266
1089		84.211	224	266
1088		84.211	224	266
036		84.211	224	266
021		84.211	224	266
2813	594	84.190	213	253

W2	NR	PERC1	EQ	PASS
2794	595	84.188	197	234
1541	596	84.170	218	259
973	597	84.151	223	265
877		84.151	223	265
581		84.151	223	265
30		84.151	223	265
2545		84.151	223	265
24		84.151	223	265
2371		84.151	223	265
2101		84.151	223	265
1901		84.151	223	265
188		84.151	223	265
1781		84.151	223	265
1700		84.151	223	265
170		84.151	223	265
1678		84.151	223	265
1642		84.151	223	265
1519		84.151	223	265
1406		84.151	223	265
1096		84.151	223	265
1011		84.151	223	265
780	616	84.091	222	264
55		84.091	222	264
2159		84.091	222	264
2147		84.091	222	264
2282		84.091	74	88
534	621	84.047	216	257
2517	622	84.043	79	94
595	623	84.030	221	263
2389		84.030	221	263
303	625	83.969	220	262
280	626	83.922	214	255
745	627	83.908	219	261
2492		83.908	219	261
992	629	83.895	224	267
957		83.895	224	267
906		83.895	224	267
899		83.895	224	267
826		83.895	224	267
77		83.895	224	267
75		83.895	224	267
714		83.895	224	267
672		83.895	224	267
660		83.895	224	267
645		83.895	224	267
550		83.895	224	267
533		83.895	224	267
509		83.895	224	267
498		83.895	224	267
490		83.895	224	267
476		83.895	224	267
470		83.895	224	267
423		83.895	224	267
415		83.895	224	267
380		83.895	224	267
379		83.895	224	267
367		83.895	224	267
296		83.895	224	267
2860		83.895	224	267
281		83.895	224	267
2765		83.895	224	267
2550		83.895	224	267
2530		83.895	224	267
2518		83.895	224	267
2516		83.895	224	267
2511		83.895	224	267
2495		83.895	224	267
2458		83.895	224	267
2444		83.895	224	267
2397		83.895	224	267
2381		83.895	224	267
2355		83.895	224	267
2315		83.895	224	267
2292		83.895	224	267
2277		83.895	224	267
227		83.895	224	267
2178		83.895	224	267
2135		83.895	224	267
2099		83.895	224	267
2098		83.895	224	267
199		83.895	224	267
194		83.895	224	267
1787		83.895	224	267
1783		83.895	224	267
1703		83.895	224	267
1672		83.895	224	267
1668		83.895	224	267
1648		83.895	224	267
1594		83.895	224	267
1545		83.895	224	267
1520		83.895	224	267
1513		83.895	224	267
1508		83.895	224	267
1504		83.895	224	267
1475		83.895	224	267
1465		83.895	224	267
1459		83.895	224	267
143		83.895	224	267
1373		83.895	224	267
1316		83.895	224	267
1309		83.895	224	267
1305		83.895	224	267
129		83.895	224	267
1238		83.895	224	267
1237		83.895	224	267
1227		83.895	224	267
1193		83.895	224	267
1163		83.895	224	267
1155		83.895	224	267
1095		83.895	224	267
1078		83.895	224	267
1035		83.895	224	267
1033		83.895	224	267
1019		83.895	224	267
07		83.895	224	267
039		83.895	224	267
037		83.895	224	267
034		83.895	224	267
1820	713	83.871	208	248
995	714	83.835	223	266
990		83.835	223	266
959		83.835	223	266
952		83.835	223	266
929		83.835	223	266
905		83.835	223	266
883		83.835	223	266
861		83.835	223	266
825		83.835	223	266
793		83.835	223	266
655		83.835	223	266
527		83.835	223	266
443		83.835	223	266
364		83.835	223	266
360		83.835	223	266
351		83.835	223	266
270		83.835	223	266
2592		83.835	223	266
2514		83.835	223	266
2509		83.835	223	266
2396		83.835	223	266
2322		83.835	223	266
2295		83.835	223	266
2284		83.835	223	266
2281		83.835	223	266
2236		83.835	223	266
2215		83.835	223	266
2142		83.835	223	266
2120		83.835	223	266
171		83.835	223	266
1440		83.835	223	266
1342		83.835	223	266
131		83.835	223	266
1303		83.835	223	266
1202		83.835	223	266
1201		83.835	223	266
1180		83.835	223	266
1039		83.835	223	266
1038		83.835	223	266
1435	753	83.830	197	235
1119	754	83.799	150	179
1808	755	83.784	217	259
988	756	83.774	222	265

Pre-Genealogical Coherence for W1=032

W2	NR	PERC1	EQ	PASS
931		83.774	222	265
73		83.774	222	265
682		83.774	222	265
486		83.774	222	265
48		83.774	222	265
349		83.774	222	265
2760		83.774	222	265
267		83.774	222	265
2546		83.774	222	265
2474		83.774	222	265
2462		83.774	222	265
229		83.774	222	265
1615		83.774	222	265
045		83.774	222	265
503	771	83.712	221	264
294		83.712	221	264
2487		83.712	221	264
1478s		83.712	221	264
1457		83.712	221	264
1196		83.712	221	264
054		83.712	221	264
2693s	778	83.658	215	257
45	779	83.650	220	263
287		83.650	220	263
2446		83.650	220	263
1567		83.650	220	263
1250		83.650	220	263
1050		83.650	220	263
600	785	83.588	219	262
736	786	83.525	218	261
TR	787	83.521	223	267
999		83.521	223	267
956		83.521	223	267
927		83.521	223	267
925		83.521	223	267
821		83.521	223	267
808		83.521	223	267
799		83.521	223	267
787		83.521	223	267
777		83.521	223	267
707		83.521	223	267
651		83.521	223	267
65		83.521	223	267
6		83.521	223	267
57		83.521	223	267
568		83.521	223	267
560		83.521	223	267
528		83.521	223	267
522		83.521	223	267
495		83.521	223	267
481		83.521	223	267
44		83.521	223	267
393		83.521	223	267
39		83.521	223	267
358		83.521	223	267
347		83.521	223	267
331		83.521	223	267
297		83.521	223	267
2724		83.521	223	267
265		83.521	223	267
261		83.521	223	267
2496		83.521	223	267
2451		83.521	223	267
2324		83.521	223	267
2265		83.521	223	267
218		83.521	223	267
2173		83.521	223	267
2172		83.521	223	267
2137		83.521	223	267
2136		83.521	223	267
213		83.521	223	267
2107		83.521	223	267
185		83.521	223	267
1823		83.521	223	267
1792		83.521	223	267
1701		83.521	223	267
1684		83.521	223	267
1660		83.521	223	267
1623		83.521	223	267
1602		83.521	223	267
1580		83.521	223	267
157		83.521	223	267
1539		83.521	223	267
1505		83.521	223	267
1474		83.521	223	267
1453		83.521	223	267
1450		83.521	223	267
1438		83.521	223	267
142		83.521	223	267
1329		83.521	223	267
1297		83.521	223	267
1279		83.521	223	267
1268		83.521	223	267
1241		83.521	223	267
1211		83.521	223	267
106		83.521	223	267
1058		83.521	223	267
1037		83.521	223	267
1031		83.521	223	267
1029		83.521	223	267
1026		83.521	223	267
1010		83.521	223	267
028		83.521	223	267
0141		83.521	223	267
95	861	83.459	222	266
933		83.459	222	266
809		83.459	222	266
801		83.459	222	266
783		83.459	222	266
762		83.459	222	266
690		83.459	222	266
668		83.459	222	266
585		83.459	222	266
538		83.459	222	266
409		83.459	222	266
359		83.459	222	266
2673		83.459	222	266
2633		83.459	222	266
2598		83.459	222	266
2523		83.459	222	266
2483		83.459	222	266
2314		83.459	222	266
2304		83.459	222	266
2174		83.459	222	266
2145		83.459	222	266
2117		83.459	222	266
205		83.459	222	266
196		83.459	222	266
189		83.459	222	266
1685		83.459	222	266
1673		83.459	222	266
1629		83.459	222	266
1479		83.459	222	266
1415		83.459	222	266
1298		83.459	222	266
1226		83.459	222	266
1223		83.459	222	266
1208		83.459	222	266
1185		83.459	222	266
1149		83.459	222	266
108		83.459	222	266
1068		83.459	222	266
1064		83.459	222	266
1043		83.459	222	266
1017		83.459	222	266
2779	902	83.429	146	175
475	903	83.402	201	241
851	904	83.396	221	265
796		83.396	221	265
70		83.396	221	265
588		83.396	221	265
545		83.396	221	265
520		83.396	221	265
2707		83.396	221	265
2454		83.396	221	265
2426		83.396	221	265
23		83.396	221	265
1802		83.396	221	265
1646		83.396	221	265
159		83.396	221	265
1543		83.396	221	265
1446		83.396	221	265

W2	NR	PERC1	EQ	PASS
1425		83.396	221	265
117		83.396	221	265
113		83.396	221	265
109		83.396	221	265
1084		83.396	221	265
1077		83.396	221	265
1024		83.396	221	265
64	926	83.333	220	264
491		83.333	220	264
2606		83.333	220	264
2406		83.333	220	264
2283		83.333	220	264
2121		83.333	220	264
1542		83.333	220	264
1375		83.333	220	264
1059		83.333	220	264
2679		83.333	170	204
888		83.333	135	162
1498	937	83.270	219	263
976	938	83.260	189	227
1540		83.260	189	227
305	940	83.206	218	262
2238		83.206	218	262
231	942	83.203	213	256
948	943	83.146	222	267
900		83.146	222	267
828		83.146	222	267
788		83.146	222	267
785		83.146	222	267
775		83.146	222	267
705		83.146	222	267
688		83.146	222	267
561		83.146	222	267
497		83.146	222	267
484		83.146	222	267
449		83.146	222	267
413		83.146	222	267
397		83.146	222	267
390		83.146	222	267
388		83.146	222	267
353		83.146	222	267
2757		83.146	222	267
2685		83.146	222	267
2684		83.146	222	267
266		83.146	222	267
2645		83.146	222	267
2637		83.146	222	267
259		83.146	222	267
2563		83.146	222	267
2500		83.146	222	267
2494		83.146	222	267
2476		83.146	222	267
2465		83.146	222	267
2404		83.146	222	267
237		83.146	222	267
2301		83.146	222	267
2263		83.146	222	267
2146		83.146	222	267
202		83.146	222	267
1813		83.146	222	267
1670		83.146	222	267
1664		83.146	222	267
1654		83.146	222	267
1635		83.146	222	267
1597		83.146	222	267
1592		83.146	222	267
1538		83.146	222	267
1512		83.146	222	267
1511		83.146	222	267
1485		83.146	222	267
148		83.146	222	267
145		83.146	222	267
144s		83.146	222	267
1443		83.146	222	267
1394		83.146	222	267
1352		83.146	222	267
1345		83.146	222	267
133		83.146	222	267
1318		83.146	222	267
1315		83.146	222	267
1310		83.146	222	267
1266		83.146	222	267
1256		83.146	222	267
118s		83.146	222	267
1187		83.146	222	267
1173		83.146	222	267
1144		83.146	222	267
1121		83.146	222	267
1074		83.146	222	267
1001		83.146	222	267
1000		83.146	222	267
1327	1010	83.137	212	255
1041	1011	83.099	118	142
937	1012	83.083	221	266
874		83.083	221	266
854		83.083	221	266
74		83.083	221	266
677		83.083	221	266
656		83.083	221	266
525		83.083	221	266
508		83.083	221	266
493		83.083	221	266
477		83.083	221	266
438		83.083	221	266
408		83.083	221	266
348		83.083	221	266
2886		83.083	221	266
279		83.083	221	266
2709		83.083	221	266
2705		83.083	221	266
2656		83.083	221	266
2591		83.083	221	266
2127		83.083	221	266
193		83.083	221	266
19		83.083	221	266
1644		83.083	221	266
1626		83.083	221	266
1613		83.083	221	266
1583		83.083	221	266
1577		83.083	221	266
153		83.083	221	266
1509		83.083	221	266
125		83.083	221	266
1194		83.083	221	266
1065		83.083	221	266
1054s		83.083	221	266
1048		83.083	221	266
901	1046	83.019	220	265
820		83.019	220	265
742		83.019	220	265
565		83.019	220	265
518		83.019	220	265
2620		83.019	220	265
2616		83.019	220	265
225		83.019	220	265
1692		83.019	220	265
1643		83.019	220	265
164		83.019	220	265
1432		83.019	220	265
1212		83.019	220	265
895	1059	82.955	219	264
734		82.955	219	264
706		82.955	219	264
422		82.955	219	264
2897		82.955	219	264
2374		82.955	219	264
162		82.955	219	264
156		82.955	219	264
1082		82.955	219	264
031s		82.955	219	264
2900	1069	82.946	214	258
2584	1070	82.927	204	246
2612	1071	82.890	218	263
2375		82.890	218	263
1966		82.890	218	263
1200s		82.890	218	263
1139		82.890	218	263
1053		82.890	218	263
1569	1077	82.879	213	257
1288	1078	82.869	208	251
798	1079	82.857	29	35
1152	1080	82.824	217	262

Pre-Genealogical Coherence for W1=032

W2	NR	PERC1	EQ	PASS
997	1081	82.772	221	267
994		82.772	221	267
90		82.772	221	267
89		82.772	221	267
811		82.772	221	267
68		82.772	221	267
676		82.772	221	267
59		82.772	221	267
580		82.772	221	267
579s		82.772	221	267
578		82.772	221	267
513		82.772	221	267
51		82.772	221	267
483		82.772	221	267
412		82.772	221	267
36		82.772	221	267
344s		82.772	221	267
34		82.772	221	267
2718		82.772	221	267
2710		82.772	221	267
2573		82.772	221	267
2266		82.772	221	267
217		82.772	221	267
209		82.772	221	267
178		82.772	221	267
1639		82.772	221	267
1624		82.772	221	267
1418		82.772	221	267
1364		82.772	221	267
1240		82.772	221	267
1214		82.772	221	267
1086		82.772	221	267
107		82.772	221	267
1036		82.772	221	267
10		82.772	221	267
165	1116	82.759	216	261
725	1117	82.707	220	266
716		82.707	220	266
666s		82.707	220	266
46		82.707	220	266
373		82.707	220	266
2608		82.707	220	266
2586		82.707	220	266
2108		82.707	220	266
192		82.707	220	266
1454		82.707	220	266
1410		82.707	220	266
1397		82.707	220	266
1348		82.707	220	266
1236		82.707	220	266
1203		82.707	220	266
1190		82.707	220	266
1127		82.707	220	266
1091		82.707	220	266
28	1135	82.703	153	185
1269	1136	82.692	215	260
827	1137	82.642	219	265
819		82.642	219	265
720		82.642	219	265
71		82.642	219	265
435s		82.642	219	265
40		82.642	219	265
295		82.642	219	265
2735		82.642	219	265
2220		82.642	219	265
2133		82.642	219	265
1393		82.642	219	265
1333		82.642	219	265
1267		82.642	219	265
1188		82.642	219	265
1047s		82.642	219	265
248	1152	82.625	214	259
2727	1153	82.609	190	230
891	1154	82.576	218	264
1788		82.576	218	264
1413		82.576	218	264
1263		82.576	218	264
1166		82.576	218	264
886	1159	82.510	217	263
1403		82.510	217	263
1009		82.510	217	263
1625	1162	82.490	212	257
011s	1163	82.474	80	97
1677	1164	82.443	216	262
858	1165	82.397	220	267
661		82.397	220	267
574		82.397	220	267
555		82.397	220	267
544		82.397	220	267
502		82.397	220	267
444		82.397	220	267
289		82.397	220	267
2713		82.397	220	267
2702		82.397	220	267
2247		82.397	220	267
186		82.397	220	267
1707		82.397	220	267
1641		82.397	220	267
1595		82.397	220	267
1588		82.397	220	267
152		82.397	220	267
1350		82.397	220	267
124		82.397	220	267
1239		82.397	220	267
1217		82.397	220	267
1087		82.397	220	267
1063		82.397	220	267
889	1188	82.331	219	266
551		82.331	219	266
54		82.331	219	266
2611		82.331	219	266
228		82.331	219	266
2223		82.331	219	266
191		82.331	219	266
1651		82.331	219	266
1645		82.331	219	266
1495		82.331	219	266
1458		82.331	219	266
138		82.331	219	266
1314		82.331	219	266
1204		82.331	219	266
1171		82.331	219	266
1090		82.331	219	266
1012		82.331	219	266
949	1205	82.308	214	260
727	1206	82.264	218	265
684		82.264	218	265
494		82.264	218	265
2775s		82.264	218	265
2658		82.264	218	265
2561		82.264	218	265
2482		82.264	218	265
184		82.264	218	265
1704		82.264	218	265
163		82.264	218	265
1531		82.264	218	265
1299		82.264	218	265
1215		82.264	218	265
993	1219	82.197	217	264
440		82.197	217	264
1780		82.197	217	264
1060		82.197	217	264
977	1223	82.022	219	267
652		82.022	219	267
524		82.022	219	267
501		82.022	219	267
355		82.022	219	267
293		82.022	219	267
2868		82.022	219	267
2809		82.022	219	267
1528		82.022	219	267
1515		82.022	219	267
1395		82.022	219	267
1387		82.022	219	267
1319		82.022	219	267
119		82.022	219	267
749	1237	81.955	218	266
306		81.955	218	266
301		81.955	218	266
1573		81.955	218	266
1555		81.955	218	266
1402		81.955	218	266

W2	NR	PERC1	EQ	PASS
1317		81.955	218	266
1302		81.955	218	266
500	1245	81.928	136	166
86	1246	81.887	217	265
829		81.887	217	265
523		81.887	217	265
1536		81.887	217	265
0211		81.887	217	265
569	1251	81.818	216	264
4		81.818	216	264
1170		81.818	216	264
1343s		81.818	171	209
2575	1255	81.783	211	258
169	1256	81.749	215	263
1630		81.749	215	263
1553		81.749	215	263
835	1259	81.648	218	267
69		81.648	218	267
686		81.648	218	267
554		81.648	218	267
1436		81.648	218	267
1353		81.648	218	267
13		81.648	218	267
1291		81.648	218	267
1265		81.648	218	267
1004		81.648	218	267
445	1269	81.641	209	256
2452	1270	81.624	191	234
834	1271	81.579	217	266
391		81.579	217	266
2711		81.579	217	266
1533		81.579	217	266
1262		81.579	217	266
1160		81.579	217	266
1113		81.579	217	266
2188	1278	81.553	168	206
679	1279	81.509	216	265
2605		81.509	216	265
1301		81.509	216	265
1243		81.509	216	265
752	1283	81.439	215	264
741		81.439	215	264
428		81.439	215	264
1081		81.439	215	264
2185	1287	81.423	206	253
011	1288	81.395	140	172
855	1289	81.273	217	267
817		81.273	217	267
357		81.273	217	267
2660		81.273	217	267
1044		81.273	217	267
𝔓60	1294	81.250	78	96
747	1295	81.203	216	266
884	1296	81.132	215	265
47		81.132	215	265
273s		81.132	215	265
2291		81.132	215	265
1506		81.132	215	265
58	1301	81.061	214	264
472	1302	81.057	184	227
878	1303	80.899	216	267
392		80.899	216	267
2214		80.899	216	267
2106		80.899	216	267
1294		80.899	216	267
022		80.899	216	267
1663	1309	80.859	207	256
61	1310	80.827	215	266
2206		80.827	215	266
1122		80.827	215	266
56	1313	80.608	212	263
2478		80.608	212	263
729	1315	80.556	203	252
732	1316	80.524	215	267
1784s		80.524	215	267
1326		80.524	215	267
05s	1319	80.460	140	174
683	1320	80.377	213	265
16	1321	80.303	212	264
744	1322	80.150	214	267
723		80.150	214	267
1689		80.150	214	267
182	1325	80.075	213	266
1135		80.075	213	266
903	1327	80.000	208	260
733		80.000	204	255
1293		80.000	16	20
𝔓59		80.000	8	10
2290		80.000	8	10
2526	1332	79.894	151	189
857	1333	79.775	213	267
2680		79.775	213	267
1574		79.775	213	267
2399	1336	79.747	126	158
370	1337	79.730	59	74
013	1338	79.605	121	152
05	1339	79.570	74	93
2535	1340	79.508	97	122
246	1341	79.365	100	126
2679s		79.365	50	63
1534	1343	79.245	210	265
1633	1344	79.231	103	130
1335	1345	79.087	208	263
𝔓66	1346	79.032	98	124
1424	1347	78.868	209	265
792	1348	78.838	190	241
841	1349	78.788	208	264
2192		78.788	208	264
2311	1351	78.652	210	267
1797		78.652	210	267
2529	1353	78.626	103	131
2148	1354	78.491	208	265
2810	1355	78.113	207	265
2643	1356	77.903	208	267
2676	1357	77.863	102	131
2650		77.863	102	131
779	1359	77.778	42	54
892	1360	77.143	54	70
27	1361	76.923	10	13
2766	1362	76.779	205	267
731	1363	76.471	143	187
2786	1364	75.564	201	266
1571s	1365	73.504	86	117
087	1366	73.333	22	30
2992	1367	73.214	41	56

Table 154: Pre-genealogical coherence with W1=032 using ***IGNTP data***

ECM Data

W2	NR	PERC1	EQ	PASS
𝔓108	1	100.000	21	21
𝔓90	2	93.750	60	64
𝔓52		93.750	30	32
A	4	90.789	276	304
1071	5	88.924	281	316
03		88.924	281	316
1571S	7	88.889	136	153
33	8	88.608	280	316
019	9	87.975	278	316
1321	10	87.937	277	315
0109	11	87.368	83	95
FΠ	12	87.342	276	316
397		87.342	276	316
044		87.342	276	316
1293	15	87.097	27	31
865	16	87.025	275	316
033		87.025	275	316
02		87.025	275	316
01		87.025	275	316
04	20	86.905	219	252
333	21	86.901	272	313
0290	22	86.873	225	259
F1	23	86.709	274	316
226		86.709	274	316
1561		86.709	274	316
MT	26	86.392	273	316
597		86.392	273	316
35		86.392	273	316
18		86.392	273	316
1320		86.392	273	316
1242		86.392	273	316
07		86.392	273	316
037		86.392	273	316
992	34	86.076	272	316
821		86.076	272	316
213		86.076	272	316
0141		86.076	272	316
317	38	85.942	269	313
157	39	85.759	271	316
1463		85.759	271	316
1014		85.759	271	316
1010		85.759	271	316
028		85.759	271	316
579	44	85.714	270	315
1230	45	85.484	265	310
799	46	85.443	270	316
2768		85.443	270	316
1654		85.443	270	316
038		85.443	270	316
2561	50	85.397	269	315
045		85.397	269	315
F13	52	85.127	269	316
994		85.127	269	316
1241		85.127	269	316
118S	55	85.050	256	301
𝔓66	56	85.034	250	294
295	57	84.984	266	313
011	58	84.862	185	218
138	59	84.810	268	316
1009		84.810	268	316
807	61	84.494	267	316
1546		84.494	267	316
1093		84.494	267	316
2223	64	84.444	266	315
𝔓60	65	84.390	173	205
892S*	66	84.314	86	102
544	67	84.177	266	316
2718		84.177	266	316
109		84.177	266	316
357	70	84.127	265	315
1319		84.127	265	315
054	72	84.076	264	314
2575	73	83.974	262	312
2615	74	83.861	265	316
05S	75	83.810	176	210
2411	76	83.729	247	295
377	77	83.706	262	313
1788	78	83.548	259	310
1029	79	83.544	264	316
022		83.544	264	316
892S	81	83.178	178	214
0211	82	83.175	262	315
732	83	82.911	262	316
2106	84	82.595	261	316
1797		82.595	261	316
1344	86	82.000	123	150
1424	87	81.962	259	316
1128	88	81.699	250	306
2192	89	81.646	258	316
2680	90	81.329	257	316
841	91	81.029	252	311
087	92	80.952	34	42
05	93	80.000	76	95
792	94	79.866	238	298
168	95	79.801	241	302
2766	96	79.618	250	314
2786	97	79.114	250	316
𝔓59	98	68.750	11	16

*Table 155: Pre-genealogical coherence with W1=032 using **ECM data***

Genealogical coherence for W1=032

IGNTP Data

The equivalent table for the ECM data is printed in full in Chapter 7.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
𝔓108		-	100.000	10	10				
𝔓52		-	100.000	9	9				
1803	1		94.828	55	58	2		1	
1143		-	93.220	110	118	4	4		
A	2		90.763	226	249	20		3	
2307	3		89.062	114	128	7	3	4	
1349	4		88.889	120	135	6	5	3	1
475s		-	88.889	24	27	1	1	1	
96	5		88.000	154	175	8	6	7	
2717	6		87.975	139	158	7	6	5	1
1571		-	87.919	131	149	7	7	3	1
2372	7		87.817	173	197	9	8	4	3
NA27	8		87.640	234	267	23	1	9	
1071	9		87.266	233	267	15	10	8	1
03			87.266	233	267	18	2	11	3
1400	11		87.050	121	139	6	3	8	1
274	12		86.957	140	161	8	6	7	
2632	13		86.700	176	203	10	9	7	1
33	14		86.692	228	263	14	11	8	2
2398	15		86.667	143	165	7	6	8	1
233	16		86.466	230	266	13	9	11	3
1816	17		86.364	228	264	13	10	11	2
930		-	86.344	196	227	11	11	8	1
1558		-	86.207	200	232	11	11	8	2
019	18		86.142	230	267	18	7	10	2
2524	19		85.932	226	263	13	12	10	2
904	20		85.827	109	127	7	6	5	
669	21		85.768	229	267	13	11	12	2
235		-	85.768	229	267	12	12	12	2
1566			85.768	229	267	13	11	13	1
139			85.768	229	267	13	11	12	2
1006			85.768	229	267	16	11	9	2
044			85.768	229	267	15	10	11	2
685	26		85.714	228	266	13	11	12	2
2613			85.714	228	266	14	11	11	2
1192		-	85.714	228	266	13	13	11	1
114			85.714	228	266	14	11	12	1
2726			85.714	12	14	1		1	
776	30		85.660	227	265	13	11	12	2
974	31		85.606	226	264	16	11	9	2
2907	32		85.567	166	194	9	6	11	2
419		-	85.551	225	263	13	13	10	2
04	33		85.526	195	228	15	12	6	
1564	34		85.433	217	254	12	11	13	1
1207	35		85.425	211	247	13	11	11	1
951		-	85.393	228	267	13	13	12	1
887	36		85.393	228	267	13	12	12	2
871			85.393	228	267	13	12	13	1
774			85.393	228	267	13	12	13	1
758			85.393	228	267	13	12	12	2
2902			85.393	228	267	14	10	13	2
2721			85.393	228	267	13	12	12	2
2472			85.393	228	267	13	12	13	1
1819			85.393	228	267	14	13	9	3
1699			85.393	228	267	13	12	12	2
1563		-	85.393	228	267	13	13	12	1
1544			85.393	228	267	13	12	12	2
1365			85.393	228	267	13	11	12	3
1278			85.393	228	267	14	12	12	1
1057			85.393	228	267	13	12	12	2
041			85.393	228	267	14	11	12	2
944		-	85.338	227	266	13	13	12	1
941		-	85.338	227	266	13	13	11	2
697	50		85.338	227	266	14	11	12	2
66			85.338	227	266	13	10	13	3
584			85.338	227	266	13	11	13	2
396			85.338	227	266	13	12	13	1
2728			85.338	227	266	14	12	11	2
215			85.338	227	266	13	12	13	1
195			85.338	227	266	13	12	13	1
1676			85.338	227	266	16	12	9	2
1604		-	85.338	227	266	13	13	12	1
1392			85.338	227	266	12	10	13	4
112			85.338	227	266	12	11	13	3
1080		-	85.338	227	266	12	12	13	2
479		-	85.283	226	265	12	12	13	2
352	60		85.283	226	265	16	13	8	2
1338		-	85.283	226	265	13	13	12	1
1229			85.283	226	265	13	11	14	1
0109	62		85.246	52	61	2	1	5	1
650		-	85.227	225	264	13	13	12	1
278	63		85.227	225	264	14	11	12	2

Genealogical coherence for W1=032

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
1142			85.227	225	264	13	12	13	1
376		-	85.185	207	243	12	12	11	1
1306	65		85.081	211	248	13	11	11	2
2562		-	85.057	222	261	13	13	12	1
1388	66		85.057	222	261	13	10	15	1
MTJ18		-	85.019	227	267	13	13	13	1
991	67		85.019	227	267	13	12	13	2
975		-	85.019	227	267	13	13	12	2
961		-	85.019	227	267	13	13	12	2
943		-	85.019	227	267	13	13	13	1
932		-	85.019	227	267	13	13	12	2
924			85.019	227	267	14	12	12	2
728			85.019	227	267	13	12	14	1
699			85.019	227	267	14	11	13	2
489			85.019	227	267	14	12	12	2
414			85.019	227	267	13	12	14	1
406		-	85.019	227	267	13	13	13	1
317		-	85.019	227	267	13	13	11	3
298			85.019	227	267	13	12	13	2
2686		-	85.019	227	267	13	13	12	2
260			85.019	227	267	13	11	13	3
2394			85.019	227	267	13	12	13	2
2346		-	85.019	227	267	12	12	14	2
226		-	85.019	227	267	13	13	13	1
2182		-	85.019	227	267	13	13	13	1
2122		-	85.019	227	267	13	13	12	2
210			85.019	227	267	13	12	13	2
190		-	85.019	227	267	13	13	13	1
1628		-	85.019	227	267	13	13	12	2
1620		-	85.019	227	267	13	13	12	2
1596		-	85.019	227	267	13	13	12	2
1586		-	85.019	227	267	13	13	13	1
1584		-	85.019	227	267	12	12	14	2
1535		-	85.019	227	267	13	13	13	1
1466			85.019	227	267	13	12	14	1
1448			85.019	227	267	14	12	12	2
1423			85.019	227	267	13	12	13	2
1346			85.019	227	267	14	12	12	2
1230		-	85.019	227	267	12	12	13	3
1219			85.019	227	267	14	11	13	2
1141		-	85.019	227	267	13	13	13	1
1042		-	85.019	227	267	13	13	13	1
100			85.019	227	267	13	12	13	2
017			85.019	227	267	14	12	11	3
78	84		84.962	226	266	13	12	13	2
696			84.962	226	266	13	12	13	2
583		-	84.962	226	266	12	12	13	3
410			84.962	226	266	13	12	14	1
402		-	84.962	226	266	13	13	12	2
276			84.962	226	266	13	12	13	2
269		-	84.962	226	266	13	13	13	1
2549			84.962	226	266	13	12	14	1

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
2463		-	84.962	226	266	13	13	12	2
2400		-	84.962	226	266	14	14	10	2
1688			84.962	226	266	13	12	13	2
1627			84.962	226	266	13	12	11	4
1367		-	84.962	226	266	13	13	13	1
1272		-	84.962	226	266	13	13	11	3
1179		-	84.962	226	266	13	13	13	1
757		-	84.906	225	265	13	13	12	2
506		-	84.906	225	265	13	13	12	2
375	91		84.906	225	265	16	12	9	3
2100		-	84.906	225	265	13	13	13	1
1779			84.906	225	265	13	12	12	3
1136			84.906	225	265	14	13	12	1
1008		-	84.906	225	265	13	13	14	
1092			84.906	135	159	8	6	9	1
158		-	84.848	224	264	14	14	10	2
1158	95		84.848	224	264	13	12	13	2
2405	96		84.791	223	263	14	12	11	3
1556		-	84.674	221	261	13	13	11	3
958		-	84.644	226	267	13	13	13	2
955		-	84.644	226	267	13	13	13	2
864		-	84.644	226	267	13	13	13	2
844	97		84.644	226	267	13	12	14	2
824		-	84.644	226	267	13	13	13	2
597		-	84.644	226	267	13	13	13	2
592		-	84.644	226	267	13	13	13	2
586		-	84.644	226	267	13	13	13	2
547		-	84.644	226	267	13	13	13	2
536		-	84.644	226	267	13	13	13	2
521		-	84.644	226	267	13	13	13	2
510		-	84.644	226	267	13	13	13	2
399		-	84.644	226	267	13	13	14	1
363		-	84.644	226	267	13	13	12	3
35		-	84.644	226	267	13	13	13	2
324		-	84.644	226	267	12	12	14	3
291		-	84.644	226	267	13	13	13	2
285		-	84.644	226	267	13	13	13	2
2806		-	84.644	226	267	13	13	13	2
2730			84.644	226	267	13	12	15	1
2665		-	84.644	226	267	13	13	14	1
2554		-	84.644	226	267	13	13	13	2
2520		-	84.644	226	267	13	13	13	2
2508		-	84.644	226	267	13	13	13	2
2503		-	84.644	226	267	13	13	13	2
2439		-	84.644	226	267	13	13	14	1
2382		-	84.644	226	267	13	13	13	2
2362			84.644	226	267	14	13	13	1
2323		-	84.644	226	267	13	13	13	2
2321		-	84.644	226	267	13	13	13	2
2296		-	84.644	226	267	13	13	13	2
2273		-	84.644	226	267	13	13	13	2
2260		-	84.644	226	267	13	13	13	2

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
2249		-	84.644	226	267	13	13	13	2
2221		-	84.644	226	267	13	13	13	2
2204		-	84.644	226	267	13	13	13	2
214		-	84.644	226	267	13	13	13	2
2131		-	84.644	226	267	13	13	13	2
211		-	84.644	226	267	13	13	14	1
18		-	84.644	226	267	13	13	13	2
1680		-	84.644	226	267	13	13	13	2
1667		-	84.644	226	267	13	13	13	2
166		-	84.644	226	267	13	13	14	1
1636		-	84.644	226	267	13	13	13	2
1618			84.644	226	267	13	12	14	2
1617		-	84.644	226	267	13	13	12	3
1576		-	84.644	226	267	13	13	13	2
1572		-	84.644	226	267	13	13	13	2
1561		-	84.644	226	267	14	14	11	2
1550		-	84.644	226	267	13	13	13	2
155		-	84.644	226	267	13	13	13	2
1548		-	84.644	226	267	13	13	13	2
1510		-	84.644	226	267	13	13	13	2
1503		-	84.644	226	267	13	13	13	2
1497			84.644	226	267	14	13	12	2
1496		-	84.644	226	267	13	13	13	2
1492		-	84.644	226	267	13	13	13	2
1491			84.644	226	267	13	12	14	2
1490		-	84.644	226	267	13	13	13	2
1487		-	84.644	226	267	13	13	13	2
1482		-	84.644	226	267	13	13	13	2
1477			84.644	226	267	14	13	12	2
1472		-	84.644	226	267	13	13	14	1
1469		-	84.644	226	267	13	13	13	2
1456		-	84.644	226	267	13	13	13	2
1427		-	84.644	226	267	13	13	13	2
141		-	84.644	226	267	13	13	13	2
1389			84.644	226	267	14	13	12	2
1339		-	84.644	226	267	13	13	13	2
1313		-	84.644	226	267	14	14	10	3
1292			84.644	226	267	13	12	14	2
127			84.644	226	267	13	12	12	4
1251		-	84.644	226	267	13	13	13	2
1235		-	84.644	226	267	13	13	13	2
1205		-	84.644	226	267	13	13	14	1
1197		-	84.644	226	267	13	13	13	2
1181		-	84.644	226	267	13	13	13	2
1157		-	84.644	226	267	13	13	13	2
1147		-	84.644	226	267	13	13	13	2
1146		-	84.644	226	267	13	13	13	2
1075		-	84.644	226	267	13	13	13	2
1072		-	84.644	226	267	13	13	13	2
105		-	84.644	226	267	13	13	14	1
1040		-	84.644	226	267	13	13	13	2
033		-	84.644	226	267	16	16	9	

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
830		-	84.615	220	260	13	13	13	1
960		-	84.586	225	266	13	13	13	2
953	107		84.586	225	266	14	13	12	2
865			84.586	225	266	16	15	10	
773		-	84.586	225	266	13	13	14	1
700			84.586	225	266	14	11	13	3
5		-	84.586	225	266	13	13	14	1
21		-	84.586	225	266	12	12	14	3
20		-	84.586	225	266	13	13	14	1
1789		-	84.586	225	266	13	13	13	2
1554			84.586	225	266	13	11	15	2
1551		-	84.586	225	266	13	13	13	2
1493		-	84.586	225	266	13	13	13	2
1426		-	84.586	225	266	13	13	14	1
1370			84.586	225	266	15	14	11	1
1285		-	84.586	225	266	13	13	14	1
1225		-	84.586	225	266	13	13	14	1
782	112		84.528	224	265	16	14	10	1
591			84.528	224	265	14	13	12	2
2767		-	84.528	224	265	13	13	13	2
2621		-	84.528	224	265	13	13	13	2
1530			84.528	224	265	14	13	13	1
1484		-	84.528	224	265	13	13	13	2
1422		-	84.528	224	265	14	14	12	1
1248		-	84.528	224	265	13	13	13	2
836	115		84.507	180	213	11	10	11	1
719		-	84.411	222	263	13	13	14	1
1431	116		84.411	222	263	14	11	14	2
806		-	84.270	225	267	14	14	12	2
754		-	84.270	225	267	13	13	14	2
72		-	84.270	225	267	13	13	14	2
530		-	84.270	225	267	13	13	14	2
387		-	84.270	225	267	14	14	12	2
2624		-	84.270	225	267	14	14	12	2
2623	117		84.270	225	267	14	13	13	2
2415		-	84.270	225	267	14	14	13	1
2193			84.270	225	267	14	13	13	2
1695			84.270	225	267	13	12	14	3
1634		-	84.270	225	267	13	13	14	2
1462		-	84.270	225	267	13	13	14	2
1398			84.270	225	267	14	13	12	3
1356			84.270	225	267	14	13	14	1
1355		-	84.270	225	267	14	14	13	1
1347		-	84.270	225	267	13	13	14	2
121		-	84.270	225	267	13	13	13	3
1018		-	84.270	225	267	13	13	14	2
030			84.270	225	267	15	13	12	2
395		-	84.211	224	266	13	13	14	2
271		-	84.211	224	266	14	14	13	1
2252		-	84.211	224	266	15	15	10	2
1552		-	84.211	224	266	14	14	12	2
1468		-	84.211	224	266	13	13	15	1

Genealogical coherence for W1=032

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
2794	123		84.188	197	234	14	12	9	2
188		-	84.151	223	265	13	13	15	1
2282	124		84.091	74	88	4	3	7	
2517	125		84.043	79	94	7	6	1	1
490		-	83.895	224	267	13	13	15	2
470		-	83.895	224	267	13	13	15	2
2516		-	83.895	224	267	14	14	11	4
2495		-	83.895	224	267	13	13	14	3
2397		-	83.895	224	267	15	15	10	3
1520	126		83.895	224	267	14	13	13	3
1513		-	83.895	224	267	14	14	13	2
861	127		83.835	223	266	17	13	12	1
655		-	83.835	223	266	14	14	13	2
443		-	83.835	223	266	13	13	15	2
045	128		83.774	222	265	15	14	13	1
821		-	83.521	223	267	15	15	11	3
799	129		83.521	223	267	16	15	12	1
331			83.521	223	267	15	14	13	2
1268			83.521	223	267	16	14	12	2
0141		-	83.521	223	267	15	15	11	3
1208		-	83.459	222	266	15	15	12	2
888		-	83.333	135	162	10	10	7	
1041	132		83.099	118	142	10	7	7	
798		-	82.857	29	35	3	3		
500		-	81.928	136	166	9	9	12	
1293	133		80.000	16	20	2	1	1	
2290			80.000	8	10	1		1	
370		-	79.730	59	74	5	5	5	
2535		-	79.508	97	122	7	7	11	
2679s	135		79.365	50	63	3	2	8	
2676	136		77.863	102	131	9	8	12	
779		-	77.778	42	54	2	2	8	
27		-	76.923	10	13	1	1	1	
087	137		73.333	22	30	3	1	4	
2992		-	73.214	41	56	5	5	5	

Table 156: Potential ancestors for W1=032 using ***IGNTP data***

Genealogical Coherence for W1=397

IGNTP Data

The equivalent table for the ECM data is printed in full in Chapter 7.

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
2726		-	100.000	14	14				
𝔓59		-	100.000	10	10				
2290		-	100.000	10	10				
𝔓52		-	100.000	9	9				
1143	1		96.610	114	118	1		2	1
1803		-	96.552	56	58			1	1
2418	2		94.828	55	58	2			1
2398		-	94.545	156	165	1	1	6	1
2307		-	94.531	121	128	2	2	3	
376	3		94.239	229	243	4	3	4	3
1590	4		94.231	98	104	3	2		1
1638	5		94.175	194	206	5	2	1	4
2774		-	93.939	248	264	4	4	5	3
2562	6		93.870	245	261	4	2	7	3
2856	7		93.633	250	267	5	3	6	3
2665			93.633	250	267	4	2	7	4
210			93.633	250	267	5	3	6	3
1586			93.633	250	267	4	3	6	4
1292			93.633	250	267	4	2	7	4
1141			93.633	250	267	4	3	6	4
926	13		93.627	191	204	5	3	1	4
657		-	93.627	191	204	4	4	3	2
2141	14		93.609	249	266	4	3	7	3
137			93.609	249	266	5	3	6	3
2632	16		93.596	190	203	4	3	2	4
2388	17		93.585	248	265	4	3	7	3
1312			93.585	248	265	5	3	6	3
1136			93.585	248	265	5	3	7	2
947		-	93.548	29	31			2	
534	20		93.385	240	257	5	3	6	3
1560	21		93.333	196	210	6	3	2	3
MTJ18	22		93.258	249	267	5	3	6	4
975			93.258	249	267	4	3	7	4
774			93.258	249	267	6	2	6	4
769		-	93.258	249	267	4	4	6	4
746		-	93.258	249	267	4	4	6	4
414			93.258	249	267	5	2	7	4
2472			93.258	249	267	5	3	6	4
226			93.258	249	267	5	3	6	4
1792			93.258	249	267	5	4	6	3
1472			93.258	249	267	4	3	7	4

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
1466			93.258	249	267	5	2	7	4
1452			93.258	249	267	4	3	7	4
1445		-	93.258	249	267	4	4	6	4
1300			93.258	249	267	4	3	7	4
105			93.258	249	267	4	3	7	4
1042			93.258	249	267	5	3	6	4
972		-	93.233	248	266	4	4	6	4
952	35		93.233	248	266	5	4	6	3
696			93.233	248	266	5	3	6	4
431		-	93.233	248	266	4	4	6	4
402			93.233	248	266	5	4	5	4
2732			93.233	248	266	5	4	6	3
269			93.233	248	266	5	3	6	4
2117			93.233	248	266	5	4	6	3
195			93.233	248	266	5	3	6	4
1688			93.233	248	266	5	3	6	4
1652		-	93.233	248	266	4	4	6	4
1493			93.233	248	266	5	4	6	3
1367			93.233	248	266	5	3	6	4
1064		-	93.233	248	266	5	5	5	3
2502	45		93.208	247	265	5	4	6	3
55		-	93.182	246	264	4	4	6	4
2245	46		93.182	246	264	6	4	5	3
1186		-	93.182	246	264	5	5	5	3
719	47		93.156	245	263	4	3	8	3
2894			93.156	245	263	5	3	7	3
96	49		93.143	163	175	5		6	1
600		-	93.130	244	262	5	5	5	3
53		-	93.103	243	261	4	4	7	3
830	50		93.077	242	260	6	3	6	3
930	51		92.952	211	227	5	3	4	4
1145	52		92.913	236	254	5	4	6	3
991	53		92.884	248	267	5	3	7	4
958			92.884	248	267	5	4	6	4
955			92.884	248	267	5	4	6	4
943			92.884	248	267	5	4	6	4
890			92.884	248	267	5	4	6	4
845			92.884	248	267	5	4	6	4
83			92.884	248	267	6	4	6	3
824			92.884	248	267	5	4	6	4
806		-	92.884	248	267	4	4	7	4

Genealogical Coherence for W1=397

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
759		-	92.884	248	267	5	5	5	4
754		-	92.884	248	267	4	4	7	4
586			92.884	248	267	5	4	6	4
549		-	92.884	248	267	4	4	6	5
547			92.884	248	267	5	4	6	4
536			92.884	248	267	5	4	6	4
510			92.884	248	267	5	4	6	4
476		-	92.884	248	267	4	4	7	4
43			92.884	248	267	5	4	6	4
380		-	92.884	248	267	4	4	7	4
367		-	92.884	248	267	4	4	7	4
35			92.884	248	267	5	4	6	4
285			92.884	248	267	5	4	6	4
2806			92.884	248	267	5	4	6	4
2765		-	92.884	248	267	4	4	7	4
2701			92.884	248	267	5	4	6	4
2554			92.884	248	267	5	4	6	4
2525			92.884	248	267	5	4	6	4
2520			92.884	248	267	5	4	6	4
2503			92.884	248	267	5	4	6	4
2439			92.884	248	267	5	3	7	4
2382			92.884	248	267	5	4	6	4
2323			92.884	248	267	5	4	6	4
232			92.884	248	267	5	4	6	4
2296			92.884	248	267	5	4	6	4
2273			92.884	248	267	5	4	6	4
2260			92.884	248	267	5	4	6	4
2221			92.884	248	267	5	4	6	4
2201		-	92.884	248	267	4	4	7	4
2175			92.884	248	267	5	4	6	4
214			92.884	248	267	5	4	6	4
2131			92.884	248	267	5	4	6	4
2109		-	92.884	248	267	4	4	7	4
18			92.884	248	267	5	4	6	4
1783			92.884	248	267	5	4	6	4
1705			92.884	248	267	6	4	6	3
1698			92.884	248	267	6	4	6	3
1667			92.884	248	267	5	4	6	4
1637			92.884	248	267	5	4	6	4
1636			92.884	248	267	5	4	6	4
1576			92.884	248	267	5	4	6	4
1572			92.884	248	267	5	4	6	4
1562			92.884	248	267	5	4	6	4
1550			92.884	248	267	5	4	6	4
155			92.884	248	267	5	4	6	4
1548			92.884	248	267	5	4	6	4
1503			92.884	248	267	5	4	6	4
1496			92.884	248	267	5	4	6	4
1492			92.884	248	267	5	4	6	4
1490			92.884	248	267	5	4	6	4
1487			92.884	248	267	5	4	6	4
1482			92.884	248	267	5	4	6	4
1469			92.884	248	267	5	4	6	4
1427			92.884	248	267	5	4	6	4
141			92.884	248	267	5	4	6	4
139			92.884	248	267	6	3	7	3
1339			92.884	248	267	5	4	6	4
1328			92.884	248	267	5	4	6	4
1320			92.884	248	267	5	4	6	4
1251			92.884	248	267	5	4	6	4
1238		-	92.884	248	267	4	4	7	4
1163		-	92.884	248	267	4	4	7	4
1147			92.884	248	267	5	4	6	4
1146			92.884	248	267	5	4	6	4
1075			92.884	248	267	5	4	6	4
1072			92.884	248	267	5	4	6	4
960	116		92.857	247	266	5	4	6	4
750		-	92.857	247	266	5	5	5	4
715		-	92.857	247	266	6	6	4	3
664			92.857	247	266	5	4	6	4
66			92.857	247	266	5	4	6	4
410			92.857	247	266	5	3	7	4
2314		-	92.857	247	266	5	5	6	3
215			92.857	247	266	5	2	8	4
1791		-	92.857	247	266	5	5	5	4
1554			92.857	247	266	6	3	7	3
1551			92.857	247	266	5	4	6	4
1442			92.857	247	266	6	5	5	3
1426			92.857	247	266	5	3	7	4
1179			92.857	247	266	5	3	6	5
1088			92.857	247	266	6	4	6	3
757	127		92.830	246	265	5	4	6	4
479			92.830	246	265	5	4	6	4
2767			92.830	246	265	5	4	6	4
2715			92.830	246	265	6	5	5	3
2621			92.830	246	265	5	4	6	4
2462			92.830	246	265	5	4	7	3
1781		-	92.830	246	265	5	5	6	3
1229			92.830	246	265	6	2	7	4
1008			92.830	246	265	5	3	7	4
505	135		92.803	245	264	6	4	7	2
2747			92.803	245	264	6	3	6	4
1457			92.803	245	264	6	4	6	3
1158			92.803	245	264	5	4	6	4
1142			92.803	245	264	6	4	6	3
245		-	92.748	243	262	5	5	7	2
2636	140		92.607	238	257	5	4	6	4
1349	141		92.593	125	135	4	3	1	2
274	142		92.547	149	161	3	2	6	1
2567	143		92.517	136	147	4	3	2	2
986		-	92.509	247	267	4	4	8	4
98		-	92.509	247	267	5	5	6	4
961	144		92.509	247	267	5	4	7	4
951			92.509	247	267	5	4	7	4

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
938		-	92.509	247	267	5	5	6	4
932			92.509	247	267	5	4	7	4
923		-	92.509	247	267	5	5	6	4
887			92.509	247	267	6	4	7	3
852			92.509	247	267	5	4	7	4
797		-	92.509	247	267	5	5	6	4
764			92.509	247	267	5	4	7	4
763		-	92.509	247	267	5	5	6	4
72			92.509	247	267	6	5	6	3
717		-	92.509	247	267	5	5	6	4
669			92.509	247	267	6	4	6	4
653			92.509	247	267	5	4	7	4
645			92.509	247	267	6	4	7	3
575		-	92.509	247	267	5	5	6	4
556		-	92.509	247	267	5	5	7	3
521			92.509	247	267	5	4	7	4
512		-	92.509	247	267	5	5	6	4
480		-	92.509	247	267	5	5	6	4
386		-	92.509	247	267	5	5	6	4
37			92.509	247	267	6	4	6	4
363			92.509	247	267	5	4	7	4
329		-	92.509	247	267	5	5	6	4
298			92.509	247	267	6	4	6	4
2773		-	92.509	247	267	5	5	6	4
275		-	92.509	247	267	5	5	7	3
2721			92.509	247	267	7	4	6	3
2706		-	92.509	247	267	5	5	6	4
2635		-	92.509	247	267	5	5	6	4
2518		-	92.509	247	267	5	5	6	4
2510		-	92.509	247	267	5	5	6	4
2466			92.509	247	267	6	4	7	3
240		-	92.509	247	267	5	5	6	4
2367		-	92.509	247	267	5	5	6	4
2356			92.509	247	267	5	4	7	4
2352		-	92.509	247	267	5	5	6	4
2261		-	92.509	247	267	5	5	6	4
2249			92.509	247	267	5	4	7	4
2229			92.509	247	267	5	4	6	5
2204			92.509	247	267	5	4	7	4
2182			92.509	247	267	6	4	6	4
2181		-	92.509	247	267	4	4	8	4
2122			92.509	247	267	5	4	7	4
204		-	92.509	247	267	5	5	6	4
201		-	92.509	247	267	5	5	6	4
190			92.509	247	267	7	4	6	3
1703		-	92.509	247	267	5	5	6	4
1686		-	92.509	247	267	5	5	6	4
1680		-	92.509	247	267	5	5	6	4
167		-	92.509	247	267	5	5	6	4
166			92.509	247	267	5	4	7	4
1650		-	92.509	247	267	5	5	6	4
1628			92.509	247	267	5	4	7	4
1620			92.509	247	267	5	4	7	4
1619		-	92.509	247	267	5	5	6	4
1618			92.509	247	267	5	4	7	4
1617		-	92.509	247	267	5	5	6	4
1596			92.509	247	267	5	4	7	4
1563			92.509	247	267	6	4	5	5
1544			92.509	247	267	6	3	7	4
1514			92.509	247	267	5	4	7	4
1501		-	92.509	247	267	5	5	6	4
15		-	92.509	247	267	4	4	8	4
1497			92.509	247	267	5	4	7	4
1491			92.509	247	267	6	4	7	3
1489		-	92.509	247	267	5	5	6	4
1477			92.509	247	267	5	4	7	4
1465			92.509	247	267	6	3	8	3
1441		-	92.509	247	267	5	5	6	4
1389			92.509	247	267	5	4	7	4
1356			92.509	247	267	7	3	7	3
1347			92.509	247	267	5	4	7	4
1323		-	92.509	247	267	5	5	6	4
1322		-	92.509	247	267	5	5	6	4
1189		-	92.509	247	267	5	5	6	4
1181			92.509	247	267	5	4	7	4
1165		-	92.509	247	267	5	5	6	4
1120			92.509	247	267	5	4	8	3
1117		-	92.509	247	267	5	5	6	4
1095		-	92.509	247	267	5	5	6	4
1076			92.509	247	267	5	4	7	4
1062		-	92.509	247	267	5	5	6	4
1040			92.509	247	267	6	4	6	4
1034			92.509	247	267	5	4	7	4
1032		-	92.509	247	267	5	5	6	4
1020		-	92.509	247	267	5	5	6	4
1018			92.509	247	267	5	4	7	4
1015		-	92.509	247	267	5	5	6	4
1010		-	92.509	247	267	5	5	7	3
1003		-	92.509	247	267	5	5	6	4
836		-	92.488	197	213	3	3	7	3
959	188		92.481	246	266	6	5	6	3
953			92.481	246	266	5	4	7	4
95		-	92.481	246	266	5	5	7	3
685			92.481	246	266	6	4	6	4
563		-	92.481	246	266	5	5	6	4
448			92.481	246	266	6	5	6	3
284		-	92.481	246	266	5	5	6	4
2788s		-	92.481	246	266	5	5	6	4
2673			92.481	246	266	6	5	6	3
263		-	92.481	246	266	5	5	6	4
2549			92.481	246	266	6	3	7	4
2322		-	92.481	246	266	5	5	6	4
1789		-	92.481	246	266	5	5	6	4
17			92.481	246	266	5	3	8	4

Genealogical Coherence for W1=397

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
1693			92.481	246	266	5	4	7	4
1470		-	92.481	246	266	5	5	6	4
1225			92.481	246	266	7	3	7	3
1061			92.481	246	266	6	2	7	5
036			92.481	246	266	6	5	6	3
591	199		92.453	245	265	5	4	7	4
2812		-	92.453	245	265	5	5	6	4
2545			92.453	245	265	6	4	7	3
2112		-	92.453	245	265	6	6	5	3
2101		-	92.453	245	265	5	5	6	4
2100			92.453	245	265	5	4	8	3
1578			92.453	245	265	7	4	6	3
1338			92.453	245	265	7	4	6	3
2179			92.453	196	212	5	4	3	4
1092			92.453	147	159	3	2	5	2
2717	206		92.405	146	158	3	2	5	2
333	207		92.400	231	250	7	6	4	2
2722	208		92.386	182	197	5	3	4	3
1435		-	92.340	217	235	5	5	4	4
1388	209		92.337	241	261	5	4	7	4
1207	210		92.308	228	247	5	4	7	3
784		-	92.278	239	259	6	6	6	2
924	211		92.135	246	267	7	4	7	3
906		-	92.135	246	267	6	6	6	3
899		-	92.135	246	267	5	5	7	4
864			92.135	246	267	5	4	8	4
843		-	92.135	246	267	5	5	7	4
823			92.135	246	267	6	5	6	4
758			92.135	246	267	6	4	7	4
728			92.135	246	267	6	4	7	4
698			92.135	246	267	6	5	6	4
672		-	92.135	246	267	5	5	7	4
597			92.135	246	267	5	3	9	4
550			92.135	246	267	6	5	7	3
461		-	92.135	246	267	5	5	7	4
399			92.135	246	267	6	4	7	4
387		-	92.135	246	267	5	5	7	4
3			92.135	246	267	7	5	6	3
291			92.135	246	267	5	4	8	4
2863			92.135	246	267	6	5	6	4
2860		-	92.135	246	267	5	5	7	4
2724			92.135	246	267	5	4	8	4
261		-	92.135	246	267	5	5	7	4
2604		-	92.135	246	267	5	5	7	4
260			92.135	246	267	8	4	6	3
2571			92.135	246	267	6	4	7	4
2508			92.135	246	267	6	5	6	4
2507		-	92.135	246	267	5	5	7	4
2496			92.135	246	267	6	5	7	3
2407			92.135	246	267	6	5	7	3
2362			92.135	246	267	7	4	6	4
2354			92.135	246	267	7	4	7	3
2292		-	92.135	246	267	5	5	7	4
2224		-	92.135	246	267	5	5	7	4
2098		-	92.135	246	267	5	5	7	4
2097			92.135	246	267	6	5	6	4
199		-	92.135	246	267	5	5	7	4
183			92.135	246	267	6	5	6	4
1701			92.135	246	267	7	6	6	2
1687		-	92.135	246	267	5	5	7	4
1684		-	92.135	246	267	5	5	7	4
1647			92.135	246	267	6	5	7	3
1634		-	92.135	246	267	5	5	7	4
1600		-	92.135	246	267	5	5	7	4
1584		-	92.135	246	267	5	5	7	4
1566			92.135	246	267	7	4	6	4
1557			92.135	246	267	6	4	7	4
1520			92.135	246	267	7	5	6	3
1504		-	92.135	246	267	5	5	7	4
150			92.135	246	267	5	4	7	5
1494			92.135	246	267	6	5	6	4
148			92.135	246	267	6	5	7	3
1476		-	92.135	246	267	5	5	7	4
1462		-	92.135	246	267	5	5	7	4
1459		-	92.135	246	267	5	5	7	4
1456			92.135	246	267	6	5	6	4
144s			92.135	246	267	6	5	7	3
1423			92.135	246	267	5	3	9	4
130		-	92.135	246	267	6	6	5	4
1296		-	92.135	246	267	6	6	6	3
1278			92.135	246	267	7	3	7	4
1205			92.135	246	267	6	4	7	4
12			92.135	246	267	5	4	8	4
111			92.135	246	267	6	5	6	4
1056		-	92.135	246	267	4	4	9	4
100			92.135	246	267	5	3	9	4
475		-	92.116	222	241	5	5	5	4
880	249		92.105	245	266	5	4	8	4
783			92.105	245	266	6	5	7	3
78			92.105	245	266	8	5	6	2
773			92.105	245	266	5	3	9	4
587		-	92.105	245	266	5	5	7	4
538		-	92.105	245	266	5	5	8	3
5		-	92.105	245	266	5	5	8	3
364		-	92.105	245	266	5	5	7	4
2633			92.105	245	266	6	5	7	3
2598			92.105	245	266	6	5	7	3
2370		-	92.105	245	266	5	5	7	4
234		-	92.105	245	266	5	5	7	4
230		-	92.105	245	266	5	5	7	4
2281		-	92.105	245	266	5	5	7	4
2278			92.105	245	266	5	4	9	3
21			92.105	245	266	7	4	7	3
20			92.105	245	266	5	3	9	4

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
1604			92.105	245	266	6	4	7	4
1552		-	92.105	245	266	5	5	7	4
1370			92.105	245	266	8	5	5	3
1359		-	92.105	245	266	6	6	6	3
1358			92.105	245	266	7	4	7	3
1298			92.105	245	266	5	4	8	4
1285			92.105	245	266	6	4	7	4
1208		-	92.105	245	266	5	5	7	4
1192			92.105	245	266	6	4	7	4
1191		-	92.105	245	266	5	5	7	4
1149			92.105	245	266	6	5	7	3
2813	265		92.095	233	253	7	6	4	3
682		-	92.075	244	265	5	5	7	4
506	266		92.075	244	265	6	4	8	3
286			92.075	244	265	7	5	6	3
2371			92.075	244	265	4	3	8	6
188			92.075	244	265	5	3	8	5
1779		-	92.075	244	265	6	6	6	3
170		-	92.075	244	265	5	5	7	4
1543		-	92.075	244	265	6	6	6	3
1530			92.075	244	265	7	4	6	4
1484		-	92.075	244	265	5	5	7	4
1248			92.075	244	265	5	4	8	4
1096		-	92.075	244	265	5	5	7	4
1565	272		92.050	220	239	5	4	5	5
650	273		92.045	243	264	6	4	6	5
496		-	92.045	243	264	5	5	7	4
2159		-	92.045	243	264	5	5	7	4
120	274		92.015	242	263	5	4	8	4
047	275		92.000	184	200	6	5	2	3
1571	276		91.946	137	149	5	4	1	2
2372	277		91.878	181	197	8	2	3	3
546	278		91.824	146	159	5	3	2	3
942	279		91.760	245	267	5	4	8	5
875			91.760	245	267	7	6	6	3
871			91.760	245	267	7	5	6	4
765			91.760	245	267	7	5	7	3
714		-	91.760	245	267	6	6	6	4
7			91.760	245	267	6	5	7	4
509		-	91.760	245	267	6	6	7	3
507			91.760	245	267	7	5	6	4
504			91.760	245	267	6	5	7	4
490		-	91.760	245	267	5	5	8	4
484		-	91.760	245	267	5	5	8	4
470		-	91.760	245	267	5	5	8	4
449		-	91.760	245	267	6	6	7	3
423		-	91.760	245	267	8	8	5	1
379		-	91.760	245	267	6	6	7	3
317		-	91.760	245	267	7	7	5	3
2754		-	91.760	245	267	5	5	8	4
2645		-	91.760	245	267	5	5	8	4
2522		-	91.760	245	267	6	6	6	4
235			91.760	245	267	6	5	7	4
2346			91.760	245	267	5	3	10	4
2321		-	91.760	245	267	5	5	8	4
2265		-	91.760	245	267	5	5	8	4
194		-	91.760	245	267	5	5	7	5
1787		-	91.760	245	267	5	5	8	4
1632		-	91.760	245	267	6	6	6	4
160			91.760	245	267	6	5	8	3
1594		-	91.760	245	267	5	5	8	4
1513			91.760	245	267	8	6	6	2
147		-	91.760	245	267	6	6	6	4
1341			91.760	245	267	6	5	7	4
127			91.760	245	267	7	6	7	2
1235			91.760	245	267	5	4	9	4
1197		-	91.760	245	267	5	5	8	4
1110		-	91.760	245	267	6	6	6	4
1057			91.760	245	267	6	4	8	4
1049		-	91.760	245	267	6	6	6	4
2907	294		91.753	178	194	4	3	6	3
941	295		91.729	244	266	6	5	7	4
700			91.729	244	266	6	5	8	3
697			91.729	244	266	8	3	8	3
666s			91.729	244	266	6	5	8	3
478			91.729	244	266	6	5	7	4
473		-	91.729	244	266	6	6	6	4
446			91.729	244	266	8	5	6	3
443			91.729	244	266	6	4	8	4
395		-	91.729	244	266	5	5	8	4
2132			91.729	244	266	7	6	6	3
1581		-	91.729	244	266	5	5	8	4
1570		-	91.729	244	266	5	5	8	4
149		-	91.729	244	266	6	6	6	4
1468		-	91.729	244	266	5	5	8	4
131		-	91.729	244	266	5	5	8	4
87	303		91.698	243	265	8	3	8	3
710			91.698	243	265	7	5	6	4
1700		-	91.698	243	265	6	6	6	4
1422			91.698	243	265	7	5	6	4
503		-	91.667	242	264	6	6	7	3
422		-	91.667	242	264	7	7	5	3
278	306		91.667	242	264	7	4	9	2
162		-	91.667	242	264	6	6	6	4
968	307		91.635	241	263	7	5	7	3
775		-	91.386	244	267	6	6	8	3
660	308		91.386	244	267	6	5	8	4
592			91.386	244	267	7	4	8	4
530			91.386	244	267	8	6	7	2
447		-	91.386	244	267	6	6	7	4
406			91.386	244	267	7	6	6	4
2730			91.386	244	267	6	5	8	4
2563		-	91.386	244	267	6	6	7	4
2495		-	91.386	244	267	6	6	7	4

Genealogical Coherence for W1=397

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
2415			91.386	244	267	6	5	8	4
237		-	91.386	244	267	6	6	7	4
2324		-	91.386	244	267	5	5	9	4
22			91.386	244	267	6	5	8	4
219		-	91.386	244	267	5	5	8	5
1602		-	91.386	244	267	6	6	8	3
1535			91.386	244	267	7	5	7	4
151		-	91.386	244	267	7	7	6	3
1505			91.386	244	267	7	6	7	3
145		-	91.386	244	267	5	5	9	4
1448			91.386	244	267	9	5	7	2
1396		-	91.386	244	267	6	6	7	4
1297		-	91.386	244	267	7	7	6	3
1157			91.386	244	267	6	4	8	5
1073			91.386	244	267	6	5	8	4
1029		-	91.386	244	267	6	6	7	4
039			91.386	244	267	6	5	7	5
030			91.386	244	267	6	2	11	4
945	322		91.353	243	266	6	5	8	4
944			91.353	243	266	7	5	7	4
905		-	91.353	243	266	6	6	7	4
584			91.353	243	266	7	5	8	3
396			91.353	243	266	7	4	8	4
360		-	91.353	243	266	5	5	8	5
276			91.353	243	266	6	4	9	4
271			91.353	243	266	6	5	8	4
175		-	91.353	243	266	7	7	7	2
1444			91.353	243	266	8	5	7	3
1440		-	91.353	243	266	6	6	7	4
1392			91.353	243	266	7	5	8	3
1185		-	91.353	243	266	6	6	8	3
112			91.353	243	266	6	4	8	5
1080			91.353	243	266	7	6	6	4
2442	332		91.342	211	231	6	5	6	3
1564	333		91.339	232	254	5	4	8	5
30	334		91.321	242	265	7	5	8	3
2687			91.321	242	265	7	6	7	3
1646		-	91.321	242	265	5	5	9	4
159			91.321	242	265	6	5	8	4
403	337		91.304	210	230	8	6	3	3
246		-	91.270	115	126	2	2	5	2
1421	338		91.266	209	229	6	4	6	4
552		-	91.255	240	263	6	6	7	4
419	339		91.255	240	263	6	5	8	4
287			91.255	240	263	6	5	8	4
2524			91.255	240	263	9	5	7	2
2405			91.255	240	263	8	6	7	2
1567		-	91.255	240	263	6	6	7	4
2437	343		91.093	225	247	8	5	6	3
2369	344		91.016	233	256	7	6	7	3
2290s			91.016	233	256	6	5	8	4
821	346		91.011	243	267	9	8	4	3
388		-	91.011	243	267	6	6	9	3
297		-	91.011	243	267	6	6	9	3
29			91.011	243	267	6	5	9	4
2530			91.011	243	267	8	7	7	2
2465		-	91.011	243	267	7	7	7	3
2213		-	91.011	243	267	6	6	8	4
2176		-	91.011	243	267	7	7	6	4
1561		-	91.011	243	267	5	5	10	4
1510			91.011	243	267	6	5	9	4
1365			91.011	243	267	8	4	9	3
134			91.011	243	267	7	6	7	4
1210			91.011	243	267	7	6	8	3
987s		-	90.977	242	266	7	7	6	4
80		-	90.977	242	266	5	5	8	6
718		-	90.977	242	266	7	7	7	3
694		-	90.977	242	266	6	6	8	4
655	353		90.977	242	266	7	6	7	4
583			90.977	242	266	6	5	8	5
351		-	90.977	242	266	7	7	7	3
2592		-	90.977	242	266	5	5	9	5
200		-	90.977	242	266	6	6	7	5
1451			90.977	242	266	8	5	8	3
1428			90.977	242	266	8	5	9	2
1357		-	90.977	242	266	6	6	9	3
1324		-	90.977	242	266	6	6	7	5
375	357		90.943	241	265	7	5	10	2
352			90.943	241	265	8	5	9	2
974	359		90.909	240	264	9	3	10	2
1166			90.909	240	264	8	6	7	3
969	361		90.875	239	263	7	6	7	4
515			90.875	239	263	8	5	8	3
1431			90.875	239	263	9	6	7	2
342		-	90.873	229	252	6	6	7	4
1288		-	90.837	228	251	6	6	8	3
1306	364		90.726	225	248	7	5	8	3
277	365		90.637	242	267	7	5	9	4
2624		-	90.637	242	267	5	5	11	4
1699			90.637	242	267	7	5	9	4
1668		-	90.637	242	267	6	6	9	4
1355			90.637	242	267	7	6	8	4
1309			90.637	242	267	8	6	8	3
1230			90.637	242	267	6	5	10	4
1187		-	90.637	242	267	7	7	7	4
1006			90.637	242	267	9	3	10	3
0141			90.637	242	267	9	8	5	3
690	372		90.602	241	266	9	6	8	2
279		-	90.602	241	266	8	8	7	2
2728			90.602	241	266	7	4	12	2
2613			90.602	241	266	8	4	9	4
1289		-	90.602	241	266	6	6	9	4
045	375		90.566	240	265	7	6	8	4
500		-	90.361	150	166	4	4	6	2

W2	NR	D	PERC1	EQ	PASS	W1<W2	W1>W2	UNCL	NOREL
992		-	90.262	241	267	6	6	9	5
2397		-	90.262	241	267	7	7	11	1
2394	376		90.262	241	267	9	5	8	4
2193		-	90.262	241	267	5	5	12	4
1695			90.262	241	267	7	6	9	4
1220			90.262	241	267	10	5	8	3
106			90.262	241	267	9	7	8	2
017			90.262	241	267	6	5	11	4
233	381		90.226	240	266	8	4	10	4
782	382		90.189	239	265	8	5	11	2
776			90.189	239	265	10	5	8	3
2794	384		90.171	211	234	7	3	11	2
1816	385		90.152	238	264	8	4	10	4
158		-	90.152	238	264	6	6	11	3
760	386		90.114	237	263	9	6	8	3
2529		-	90.076	118	131	2	2	7	2
726		-	89.888	240	267	6	6	11	4
2623	387		89.888	240	267	8	6	10	3
211		-	89.888	240	267	7	7	9	4
157		-	89.888	240	267	7	7	8	5
1346			89.888	240	267	7	5	11	4
1071			89.888	240	267	13	6	7	1
041			89.888	240	267	7	5	11	4
861	391		89.850	239	266	9	5	10	3
2463			89.850	239	266	7	6	10	4
2400			89.850	239	266	8	7	8	4
1676			89.850	239	266	9	5	11	2
1627		-	89.850	239	266	7	7	10	3
48		-	89.811	238	265	7	7	10	3
1425		-	89.811	238	265	7	7	9	4
345	395		89.615	233	260	8	7	8	4
489	396		89.513	239	267	9	5	11	3
1268			89.513	239	267	10	6	11	1
1626		-	89.474	238	266	8	8	9	3
1455	398		89.474	238	266	9	7	8	4
1273	399		89.431	220	246	11	5	6	4
2535	400		89.344	109	122	2	1	8	2
2676	401		89.313	117	131	4	2	7	1
331	402		89.139	238	267	9	7	10	3
2902			89.139	238	267	8	5	12	4
1219			89.139	238	267	8	5	12	4
037		-	89.139	238	267	9	9	7	4
482	405		89.098	237	266	9	8	8	4
1272			89.098	237	266	9	8	8	4
114			89.098	237	266	8	6	11	4
581		-	89.057	236	265	8	8	10	3
33	408		88.973	234	263	12	5	9	3
799	409		88.764	237	267	11	8	9	2
699			88.764	237	267	8	6	12	4
1398			88.764	237	267	10	8	9	3
044			88.764	237	267	13	6	9	2
A	413		88.755	221	249	24		2	2
0109		-	88.525	54	61	1	1	5	
1819	414		87.640	234	267	13	8	10	2
2992		-	87.500	49	56	2	2	2	1
04	415		87.281	199	228	15	6	6	2
033	416		87.266	233	267	15	11	6	2
2684		-	86.891	232	267	11	11	10	3
865	417		86.466	230	266	15	12	7	2
1820	418		86.290	214	248	13	10	9	2
2517	419		85.106	80	94	5	3	4	2
NA27	420		85.019	227	267	24	1	13	2
019			85.019	227	267	20	7	12	1
1582		-	84.586	225	266	13	13	12	3
03	422		84.270	225	267	22	3	13	4
032	423		83.146	222	267	15	13	14	3

*Table 157: Potential ancestors for W1=397 using **IGNTP data***

Appendix 6: Optimal Substemmata

Optimal Substemmata

In Chapter 10 I defined the Optimal Substemmata for all witnesses. The following notes show the working behind my decisions. The comment “simple” means that my script `analyse_combinations_of_ancestors.py` in my CBGM software found a single combination that had minimal ‘Offen’ and lowest summed rank – with no combinations with fewer witnesses able to challenge it.[1]

```
# Thu 14 Dec 13:44:53 GMT 2017

optimal_substemmata = {
    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 011.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 4
    # > 1 rows have Offen=4
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 18, 2786, 03
    '011': [{'18', '2786', '03'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 0141.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 14 rows have Offen=1
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 821, FΠ, F1, 03
    '0141': [{'821', 'FΠ', 'F1', '03'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 017C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 8191 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 7930 rows have Offen=0
    # > Restricted to the best of each combination size: 13 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 038
    '017C': [{'038'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 019C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
```

---

[1] See https://doi.org/10.5281/zenodo.1296287 or https://github.com/edmondac/CBGM

```
    # > 3401 rows have Offen=0
    # > Restricted to the best of each combination size: 7 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'1241', '01'} explains the most by agreement
    # > Combination {'1241'} explains the following by posterity (compared to
{'1241', '01'}): {'B04K13V37/6-8'}
    # Just an article
    '019C': [{'1241'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 019.csv
    # > Loaded 7 rows
    # > Min 'Offen' is 4
    # > 2 rows have Offen=4
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'A', '03', 'P75'} explains the most by agreement
    # > Too large a difference by posterity (21), skipping combination {'A', '03'}
    '019': [{'A', '03', 'P75'}],  # simple, in the end

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 01Cca.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 9
    # > 117 rows have Offen=9
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'019', '044', '2786', '032'} explains the most by agreement
    # > Combination {'019', 'FΠ', '032'} explains the following by posterity
(compared to {'019', '044', '2786', '032'}): {'B04K03V03/4-10', 'B04K05V15/2-8',
'B04K11V52/11', 'B04K06V11/22', 'B04K20V16/27', 'B04K13V32/2-12'}
    # B04K20V16/27 is an added phrase "και προσεδραμεν αψασθαι αυτου" - too much to
be coincidence...
    '01Cca': [{'019', '044', '2786', '032'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py
01Ccb2.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 391 rows have Offen=0
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: F13, 01
    '01Ccb2': [{'F13', '01'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 01.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 4
    # > 8 rows have Offen=4
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: A, 226, P66
    '01': [{'A', '226', 'P66'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 0211.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 10
    # > 4 rows have Offen=10
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 045, A
    '0211': [{'045', 'A'}],  # simple
```

```
    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 022.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 4 rows have Offen=2
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 044, FΠ, 397, 03
    '022': [{'044', 'FΠ', '397', '03'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 0233.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 13 rows have Offen=1
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 226, 038, 397
    '0233': [{'226', '038', '397'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 028.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 3
    # > 29 rows have Offen=3
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 35, FΠ, A
    '028': [{'35', 'FΠ', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 0290.csv
    # > Found Hinweis '<<' entries
    # > Loaded 8191 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 113 rows have Offen=0
    # > Restricted to the best of each combination size: 12 comb. left
    # > Excluded rows with more combinations but no better results: 6 comb. left
    # > Human thought required (probably)...
    # > Combination {'2561C3', '0141', '33', 'A', '1071', '044', '04'} explains the
most by agreement
    # > Combination {'2561C3', '0141', '33', 'A', '1071', '044'} explains the
following by posterity (compared to {'2561C3', '0141', '33', 'A', '1071', '044',
'04'}): {'B04K18V34/2-4'}
    # added article
    # > Combination {'2561C3', '0141', '33', 'A', '044'} explains the following by
posterity (compared to {'2561C3', '0141', '33', 'A', '1071', '044'}):
{'B04K19V34/22-24'}
    # two words reversed
    # > Combination {'0141', '33', 'A', '044'} explains the following by posterity
(compared to {'2561C3', '0141', '33', 'A', '044'}): {'B04K19V31/8-34',
'B04K19V27/10'}
    # επει παρασκευη ην moved to end of phrase instead of start (I.e. word order),
ιδου vs. ιδε
    # > Combination {'0141', 'A', '044'} explains the following by posterity
(compared to {'0141', '33', 'A', '044'}): {'B04K19V38/2-6', 'B04K19V39/14-22',
'B04K19V38/13'}
    # removed δε, τον ιησουν vs. αυτον, added article
    # I think τον ιησουν is sufficiently connective here
    # > Too large a difference by posterity (20), skipping combination {'A', '044'}
    '0290': [{'0141', '33', 'A', '044'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 029.csv
    # > Found Hinweis '<<' entries
    # > Loaded 1023 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
```

```
    # > 35 rows have Offen=0
    # > Restricted to the best of each combination size: 9 comb. left
    # > Excluded rows with more combinations but no better results: 4 comb. left
    # > Human thought required (probably)...
    # > Combination {'03', 'P75', '019', '04', '032S'} explains the most by
agreement
    # > Combination {'03', 'P75', '019', '032S'} explains the following by
posterity (compared to {'03', 'P75', '019', '04', '032S'}): {'B04K04V53/9'}
    # Added αυτου
    # > Combination {'A', '019', '032S'} explains the following by posterity
(compared to {'03', 'P75', '019', '032S'}): {'B04K08V31/24', 'B04K03V12/30',
'B04K06V30/20-24', 'B04K06V32/24', 'B04K06V52/20-30'}
    # > Too large a difference by posterity (12), skipping combination {'A',
'032S'}
    '029': [{'03', 'P75', '019', '032S'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 02.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 122 rows have Offen=1
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 3 comb. left
    # > Human thought required (probably)...
    # > Combination {'FΠ', '892', 'A', 'F1', '33'} explains the most by agreement
    # > Combination {'FΠ', '892', 'A', 'F1'} explains the following by posterity
(compared to {'FΠ', '892', 'A', 'F1', '33'}): {'B04K15V08/24', 'B04K06V09/6',
'B04K18V25/40-44', 'B04K19V27/26', 'B04K03V03/4-10', 'B04K12V09/50-56',
'B04K20V04/12-18', 'B04K19V38/13'}
    # 'B04K19V27/26': ημερας vs. ωρας, 'B04K20V04/12-18': ο δε vs. και ο,
'B04K18V25/40-44': λεγει vs. ειπεν... that's enough to say "no".
    # > Too large a difference by posterity (18), skipping combination {'FΠ', 'A',
'F1'}
    '02': [{'FΠ', '892', 'A', 'F1', '33'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 032.csv
    # > Loaded 4095 rows
    # > Min 'Offen' is 6
    # > 1280 rows have Offen=6
    # > Restricted to the best of each combination size: 11 comb. left
    # > Excluded rows with more combinations but no better results: 8 comb. left
    # > Human thought required (probably)...
    # > Combination {'A', '029', 'P75', '03', '019', '04', '33', '044', '0290'}
explains the most by agreement
    # > Combination {'A', '029', '03', '019', '04', '33', '044', '0290'} explains
the following by posterity (compared to {'A', '029', 'P75', '03', '019', '04',
'33', '044', '0290'}): {'B04K10V09/14'}
    # εαν vs. αν
    # > Combination {'A', '029', '03', '019', '04', '33', '044'} explains the
following by posterity (compared to {'A', '029', '03', '019', '04', '33', '044',
'0290'}): {'B04K19V13/28-30', 'B04K18V20/12-16'}
    # added article, ελαλησα vs λελαληκα
    # > Combination {'A', '029', '019', '04', '33', '044'} explains the following
by posterity (compared to {'A', '029', '03', '019', '04', '33', '044'}):
{'B04K12V04/4', 'B04K20V31/36-38'}
    # δε vs. ουν, ζωην αιωνιον εχητε vs. ζωην εχητε - eye skip to the ν?
    # > Combination {'A', '019', '04', '33', '044'} explains the following by
posterity (compared to {'A', '029', '019', '04', '33', '044'}): {'B04K08V31/24',
'B04K06V57/34'}
    # μεινητε vs. μενητε, ζησεται vs. ζησει
    # > Too large a difference by posterity (14), skipping combination {'A', '019',
'33', '044'}
    # > Too large a difference by posterity (21), skipping combination {'A', '33',
'044'}
    # > Too large a difference by posterity (52), skipping combination {'A', '044'}
    '032': [{'A', '019', '04', '33', '044'}],
```

```
    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 032S.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 4 rows have Offen=0
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'A', '892', '397'} explains the most by agreement
    # > Combination {'A', '213'} explains the following by posterity (compared to
{'A', '892', '397'}): {'B04K05V01/18-22', 'B04K03V21/36', 'B04K03V20/36-40',
'B04K04V01/34', 'B04K02V15/4-10', 'B04K04V30/3', 'B04K01V39/10', 'B04K03V20/41',
'B04K05V01/6', 'B04K02V17/15', 'B04K01V44/4-6', 'B04K04V05/24', 'B04K04V53/31',
'B04K01V40/2'}
    # > ... and the other way round: {'B04K04V54/4-6', 'B04K01V15/16-22',
'B04K01V37/6-14', 'B04K04V31/10-14'}
    # That's too many - so just take {A, 892, 397}
    '032S': [{'A', '892', '397'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 033.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 4
    # > 66 rows have Offen=4
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'A', '044', '33', '019', '032'} explains the most by agreement
    # > Combination {'A', '044', '33', '019'} explains the following by posterity
(compared to {'A', '044', '33', '019', '032'}): {'B04K14V15/16', 'B04K10V38/26-28',
'B04K20V31/36-38', 'B04K14V02/26-30', 'B04K10V32/34-36', 'B04K16V29/4-8',
'B04K09V30/20-24'}
    # τηρησετε vs. τηρησατε, και γινωσκητε vs. και γινωσκετε, ζωην αιωνιον εχητε
vs. ζωην εχητε - eye skip to the ν?, missing αν, εμε λιθαζετε vs. λιθαζετε με, odd
- 033 and 33 agree at 'B04K16V29/4-8', article - seems a bit much to me.
    '033': [{'A', '044', '33', '019', '032'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 037C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 511 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 189 rows have Offen=0
    # > Restricted to the best of each combination size: 9 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'044', '019'} explains the most by agreement
    # > Combination {'044'} explains the following by posterity (compared to
{'044', '019'}): {'B04K05V35/24-28'}
    # αγαλλιασθηναι vs. αγαλλιαθηναι
    '037C': [{'044'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 037.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 3
    # > 14 rows have Offen=3
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 045, A, 1071
    '037': [{'045', 'A', '1071'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 038.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 5
```

```
    # > 2 rows have Offen=5
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: FΠ, A, 022
    '038': [{'FΠ', 'A', '022'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 03.csv
    # > Loaded 1 rows
    # > Min 'Offen' is 1
    # > 1 rows have Offen=1
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: A
    '03': [{'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 044.csv
    # > Loaded 63 rows
    # > Min 'Offen' is 21
    # > 4 rows have Offen=21
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 3 comb. left
    # > Human thought required (probably)...
    # > Combination {'A', '019', '04', '03', 'P75', '029'} explains the most by
agreement
    # > Combination {'A', '019', '04', '03', 'P75'} explains the following by
posterity (compared to {'A', '019', '04', '03', 'P75', '029'}): {'B04K06V45/23',
'B04K01V26/21'}
    # added ουν, added δε
    # > Too large a difference by posterity (14), skipping combination {'A', '019',
'04', '03'}
    '044': [{'A', '019', '04', '03', 'P75'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 045.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 3
    # > 3 rows have Offen=3
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 028, 2786, A
    '045': [{'028', '2786', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 04C1.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 12 rows have Offen=0
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'02', '1242', '01', '04'} explains the most by agreement
    # > Combination {'02', '01', '04'} explains the following by posterity
(compared to {'02', '1242', '01', '04'}): {'B04K17V01/43', 'B04K11V12/6-10'}
    # added και, αυτου vs. αυτω
    '04C1': [{'02', '01', '04'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 04C2.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 16 rows have Offen=0
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
```

```
    # > Combination {'1278C1', '07', '03'} explains the most by agreement
    # > Combination {'07', '03'} explains the following by posterity (compared to
{'1278C1', '07', '03'}): {'B04K21V01/6-18'}
    # 04C2 and 1271C1 had unique readings anyway, albeit very similar.
Regularisation?
    '04C2': [{'07', '03'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 04.csv
    # > Loaded 15 rows
    # > Min 'Offen' is 3
    # > 2 rows have Offen=3
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'A', '03', '019', 'P75'} explains the most by agreement
    # > Combination {'A', '03', '019'} explains the following by posterity
(compared to {'A', '03', '019', 'P75'}): {'B04K08V41/18', 'B04K06V52/6-12',
'B04K11V22/12-14', 'B04K01V33/65', 'B04K04V12/20'}
    # added ουν, word order, οσα εαν vs. οσα αν, addition of και πυρι, δεδωκεν vs.
εδωκεν
    # P75 is also a close ancestor - so keep it
    '04': [{'A', '03', '019', 'P75'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 054.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 1 rows have Offen=0
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 35, 019, F1
    '054': [{'35', '019', 'F1'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 05.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 20
    # > 2 rows have Offen=20
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: A, F13
    '05': [{'A', 'F13'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 05S.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 9 rows have Offen=0
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 3 comb. left
    # > Human thought required (probably)...
    # > Combination {'597', '038', '044', 'P66'} explains the most by agreement
    # > Combination {'597', 'A', '044'} explains the following by posterity
(compared to {'597', '038', '044', 'P66'}): {'B04K18V40/24-26', 'B04K18V34/10',
'B04K18V36/38-44'}
    # article, om. vs. συ, η εμη βασιλεια vs. η βασιλεια η εμη
    # > Combination {'597', 'A'} explains the following by posterity (compared to
{'597', 'A', '044'}): {'B04K19V31/46-50', 'B04K19V14/18'}
    # word order, τριτη vs. εκτη - that's ok
    '05S': [{'597', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 063.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
```

```
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 3 rows have Offen=0
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 226, 579
    '063': [{'226', '579'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 070.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 4 rows have Offen=0
    # > Restricted to the best of each combination size: 4 comb. left
    # > Excluded rows with more combinations but no better results: 4 comb. left
    # > Human thought required (probably)...
    # > Combination {'019', '33', '03', 'P75', '044'} explains the most by
agreement
    # > Combination {'04C1', '086', '019', '03', 'P75', '044'} explains the
following by posterity (compared to {'019', '33', '03', 'P75', '044'}):
{'B04K11V51/30-34'}
    # Useful to check for the next line's benefit: just an article
    # > Combination {'019', '03', 'P75', '044'} explains the following by posterity
(compared to {'04C1', '086', '019', '03', 'P75', '044'}): {'B04K09V09/27'}
    # added δε
    # > Combination {'03', 'P75', '044'} explains the following by posterity
(compared to {'019', '03', 'P75', '044'}): {'B04K09V09/30', 'B04K09V08/27',
'B04K09V04/18', 'B04K11V56/18-24'}
    # removed οτι, added οτι, ως vs. εως, word order
    '070': [{'03', 'P75', '044'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 07C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 2204 rows have Offen=0
    # > Restricted to the best of each combination size: 4 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'18', '157'} explains the most by agreement
    # > Combination {'18'} explains the following by posterity (compared to {'18',
'157'}): {'B04K03V04/4-10'}
    # article
    '07C': [{'18'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 07.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 31 rows have Offen=2
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 028, 157, 03
    '07': [{'028', '157', '03'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 083.csv
    # > Found Hinweis '<<' entries
    # > Loaded 2047 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 57 rows have Offen=0
    # > Restricted to the best of each combination size: 10 comb. left
    # > Excluded rows with more combinations but no better results: 5 comb. left
    # > Human thought required (probably)...
```

```
    # > Combination {'029', 'A', 'P75', '044', '019', '033'} explains the most by
agreement
    # > Combination {'A', 'P75', '044', '019', '033'} explains the following by
posterity (compared to {'029', 'A', 'P75', '044', '019', '033'}): {'B04K01V28/4-8'}
    # βηθαβαρα vs. βηθανια
    # > Combination {'A', 'P75', '044', '019'} explains the following by posterity
(compared to {'A', 'P75', '044', '019', '033'}): {'B04K01V35/6-12', 'B04K01V33/62-
64'}
    # παλιν ιστηκει ο ιωαννης vs. ειστηκει ιωαννης (P75) παλιν ειστηκει ιωαννης
(019) and ιστηκει ο ιωαννης (044), article
    # > Combination {'A', 'P75', '044'} explains the following by posterity
(compared to {'A', 'P75', '044', '019'}): {'B04K04V46/28-30', 'B04K04V08/10'}
    # ην δε vs. και ην, απηλθον (unique to 083 and 019) vs. απεληλυθεισαν but 019
is 6th ranked - so accept multiple emergence? No. There's no other reading like it.
    # > Combination {'A', '044'} explains the following by posterity (compared to
{'A', 'P75', '044'}): {'B04K04V50/26-38', 'B04K03V20/41', 'B04K04V03/12',
'B04K02V18/22', 'B04K03V12/30'}
    # Not continuing...
    '083': [{'A', 'P75', '044', '019'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 086.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 22 rows have Offen=0
    # > Restricted to the best of each combination size: 4 comb. left
    # > Excluded rows with more combinations but no better results: 4 comb. left
    # > Human thought required (probably)...
    # > Combination {'18', 'P75', '213', '022'} explains the most by agreement
    # > Combination {'18', 'P75', '022'} explains the following by posterity
(compared to {'18', 'P75', '213', '022'}): {'B04K03V27/16', 'B04K03V21/24-28'}
    # λαμβανειν αφ εαυτου vs. λαμβανειν - relatively common words in John so
accept, word order
    # > Combination {'18', 'P66'} explains the following by posterity (compared to
{'18', 'P75', '022'}): {'B04K03V12/30', 'B04K03V27/3', 'B04K04V13/3'}
    # πιστευετε vs. πιστευσετε, added article, added article
    # > Too large a difference by posterity (27), skipping combination {'A'}
    '086': [{'18', 'P66'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1009.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 6
    # > 2 rows have Offen=6
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 2223, A
    '1009': [{'2223', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1010.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 10
    # > 11 rows have Offen=10
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'118S', '07', 'A'} explains the most by agreement
    # > Combination {'07', 'A'} explains the following by posterity (compared to
{'118S', '07', 'A'}): {'B04K19V11/38', 'B04K21V04/12', 'B04K19V35/14-20',
'B04K20V11/8-16'}
    # παραδιδους vs. παραδους, article, word order, το vs. τω
    '1010': [{'07', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1014.csv
    # > Loaded 9999 rows
```

```
    # > Min 'Offen' is 6
    # > 28 rows have Offen=6
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1320, FΠ, A
    '1014': [{'1320', 'FΠ', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1029.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 17
    # > 12 rows have Offen=17
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 045, A
    '1029': [{'045', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1071.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 8
    # > 4 rows have Offen=8
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 33, 045, A
    '1071': [{'33', '045', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1093.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 18
    # > 2 rows have Offen=18
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1788, A
    '1093': [{'1788', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 109C.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 2107 rows have Offen=1
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1010, 0141
    '109C': [{'1010', '0141'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 109.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 5
    # > 3 rows have Offen=5
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'019C', '226', 'A'} explains the most by agreement
    # > Combination {'226', 'A'} explains the following by posterity (compared to
{'019C', '226', 'A'}): {'B04K15V24/28', 'B04K13V37/6-8'}
    # ειχον vs. ειχοσαν, article
    '109': [{'226', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1128.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 14
    # > 2 rows have Offen=14
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: FΠ, A
    '1128': [{'FΠ', 'A'}],  # simple
```

```
    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 118S.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 1097 rows have Offen=1
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 3 comb. left
    # > Human thought required (probably)...
    # > Combination {'18', '317', '1321', '044'} explains the most by agreement
    # > Combination {'18', '1321', '044'} explains the following by posterity
(compared to {'18', '317', '1321', '044'}): {'B04K19V40/24', 'B04K16V29/4-8'}
    # om. article, same as B04K16V29/4-8 above...
    # > Combination {'18', 'A'} explains the following by posterity (compared to
{'18', '1321', '044'}): {'B04K19V28/6-10', 'B04K20V11/8-16', 'B04K18V11/32',
'B04K20V13/6-8'}
    # ιδων vs. ειδως, μνημειον vs. μνημειω, δεδωκεν vs. εδωκεν, αυτοι vs. αυτη
    '118S': [{'18', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1230.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 5
    # > 6 rows have Offen=5
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 35, 1071, A
    '1230': [{'35', '1071', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1241.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 9
    # > 6 rows have Offen=9
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1320, A
    '1241': [{'1320', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1242C.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 3
    # > 14 rows have Offen=3
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'083', '0211', '05'} explains the most by agreement
    # > Combination {'0211', '03'} explains the following by posterity (compared to
{'083', '0211', '05'}): {'B04K06V07/24-28', 'B04K04V52/22-24'}
    # εκαστος αυτων βραχυ τι vs.εκαστος βραχυ, και ειπον vs. ειπον ουν
    # > ... and the other way round: {'B04K05V10/32'}
    # om. σου
    '1242C': [{'0211', '03'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1242.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 9
    # > 10 rows have Offen=9
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'1571', '35', 'A'} explains the most by agreement
    # > Too large a difference by posterity (31), skipping combination {'35', 'A'}
    '1242': [{'1571', '35', 'A'}],  # simple, in the end

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1253.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
```

```
    # > Min 'Offen' is 0
    # > 1 rows have Offen=0
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1230, 2790, A
    '1253': [{'1230', '2790', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py
1278C1.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 166 rows have Offen=0
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 35, 841
    '1278C1': [{'35', '841'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1293.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 7 rows have Offen=2
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1010, 011, A
    '1293': [{'1010', '011', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1319.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 7
    # > 6 rows have Offen=7
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1463, A
    '1319': [{'1463', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1320.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 2 rows have Offen=1
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 35, F1, 03
    '1320': [{'35', 'F1', '03'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1321.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 7
    # > 1 rows have Offen=7
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: FΠ, 397, A
    '1321': [{'FΠ', '397', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1344C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 8061 rows have Offen=0
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1320
    '1344C': [{'1320'}],  # simple
```

```
    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1344.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 9
    # > 4 rows have Offen=9
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'019C', '1320', 'A'} explains the most by agreement
    # > Combination {'1320', 'A'} explains the following by posterity (compared to
{'019C', '1320', 'A'}): {'B04K13V37/6-8'}
    # article
    '1344': [{'1320', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 138.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 5 rows have Offen=2
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 994, 157, A
    '138': [{'994', '157', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1424.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 14
    # > 2 rows have Offen=14
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 045, A
    '1424': [{'045', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1463.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 6
    # > 3 rows have Offen=6
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 045, A
    '1463': [{'045', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1546C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 95 rows have Offen=0
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'1561', '029'} explains the most by agreement
    # > Combination {'1561'} explains the following by posterity (compared to
{'1561', '029'}): {'B04K01V25/2-12'}
    # και ηρωτησαν αυτον vs. και ηρωτησαν αυτον και ειπον αυτω - I.e. om. και ειπον
αυτω
    '1546C': [{'1561'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1546.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 7
    # > 3 rows have Offen=7
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'07C', 'FΠ', 'A'} explains the most by agreement
```

```
    # > Combination {'FΠ', 'A'} explains the following by posterity (compared to
{'07C', 'FΠ', 'A'}): {'B04K08V09/2-40'}
    # FΠ is lac in that verse... so without further evidence we'll include 07C.
Initial text is to omit the verse.
    '1546': [{'07C', 'FΠ', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1561.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 3
    # > 4 rows have Offen=3
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: FΠ, 35, 03
    '1561': [{'FΠ', '35', '03'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1571.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 2 rows have Offen=0
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 2561, 994, A
    '1571': [{'2561', '994', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1571S.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 1 rows have Offen=0
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: FΠ, F1, A
    '1571S': [{'FΠ', 'F1', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 157.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 6
    # > 1 rows have Offen=6
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 028, F1, A
    '157': [{'028', 'F1', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py
1582C1.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 5 rows have Offen=0
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 226, 799, 138
    '1582C1': [{'226', '799', '138'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1654.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 19
    # > 10 rows have Offen=19
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
```

```
    # > Combination {'2192C', '226', 'A'} explains the most by agreement
    # > Combination {'226', 'A'} explains the following by posterity (compared to
{'2192C', '226', 'A'}): {'B04K08V42/3'}
    # added ουν
    '1654': [{'226', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 168C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 33 rows have Offen=0
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'1093', '045'} explains the most by agreement
    # > Combination {'045'} explains the following by posterity (compared to
{'1093', '045'}): {'B04K01V29/6', 'B04K03V19/28-38', 'B04K14V14/2-4'}
    # βλεπει ο ιωαννης vs. βλεπει, το σκοτος μαλλον vs. μαλλον το σκοτος, added και
    '168C': [{'045'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 168.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 14
    # > 8 rows have Offen=14
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'07C', '07', 'A'} explains the most by agreement
    # > Combination {'07', 'A'} explains the following by posterity (compared to
{'07C', '07', 'A'}): {'B04K05V02/26', 'B04K16V13 /44', 'B04K03V04/4-10'}
    # βηθεσδα vs. βηθσαιδα, ακουση vs. ακουσει, article
    '168': [{'07', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 173.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 14
    # > 3 rows have Offen=14
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1029, A
    '173': [{'1029', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1788C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 67 rows have Offen=0
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'377', '226'} explains the most by agreement
    # > Combination {'1321'} explains the following by posterity (compared to
{'377', '226'}): {'B04K08V10/2-24', 'B04K11V34/8', 'B04K11V40/20', 'B04K12V26/4-8',
'B04K08V09/2-40', 'B04K08V06/2-41'}
    # 'B04K08V10/2-24', 'B04K08V09/2-40' and 'B04K08V06/2-41' are om. verses in
1321 - that's not good enough.
    '1788C': [{'377', '226'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1788.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 9
    # > 3 rows have Offen=9
    # > Restricted to the best of each combination size: 2 comb. left
```

```
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 045, A
    '1788': [{'045', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1797C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 803 rows have Offen=0
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'1546', '028'} explains the most by agreement
    # > Combination {'028'} explains the following by posterity (compared to
{'1546', '028'}): {'B04K19V30/4'}
    # δε vs. ουν
    '1797C': [{'028'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 1797.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 7
    # > 6 rows have Offen=7
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1320, A
    '1797': [{'1320', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 18.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 3
    # > 14 rows have Offen=3
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: FΠ, 597, 799, A
    '18': [{'FΠ', '597', '799', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2106.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 22
    # > 1 rows have Offen=22
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1561, A
    '2106': [{'1561', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 213.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 4 rows have Offen=2
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 033, 33, A, FΠ
    '213': [{'033', '33', 'A', 'FΠ'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 213S.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 394 rows have Offen=1
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'992', 'F1', 'A'} explains the most by agreement
```

```
    # > Combination {'317', 'FΠ'} explains the following by posterity (compared to
{'992', 'F1', 'A'}): {'B04K19V24/60', 'B04K19V41/24', 'B04K19V27/30-34',
'B04K20V23/20', 'B04K20V26/14-20', 'B04K19V13/28-30', 'B04K21V01/6-18'}
    # εβαλλον vs. εβαλον, τοπω vs. κηπω, word order, κρατειτε vs. κρατητε, removed
αυτου, added article, added αυτου.
    # > ... and the other way round: {'B04K20V30/22-26'}
    # removes αυτου
    # Also, {'317', 'FΠ'} sum_rank is 35, whereas {992, F1, A} is 104
    '213S': [{'317', 'FΠ'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2192C.csv
    # > Loaded 4095 rows
    # > Min 'Offen' is 1
    # > 4004 rows have Offen=1
    # > Restricted to the best of each combination size: 12 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'FΠ', '226'} explains the most by agreement
    # > Combination {'A'} explains the following by posterity (compared to {'FΠ',
'226'}): {'B04K06V17/8', 'B04K20V21/6-12', 'B04K08V54/48', 'B04K08V59/31'}
    # article, added ο ιησους, ημων vs. υμων, addition of "διελθων δια μεσου αυτων
και παρηγεν ουτως - too much to accept"
    '2192C': [{'FΠ', '226'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2192.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 9
    # > 2 rows have Offen=9
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 045, A
    '2192': [{'045', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2223.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 6
    # > 3 rows have Offen=6
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'07C', '2680', 'A'} explains the most by agreement
    # > Combination {'2680', 'A'} explains the following by posterity (compared to
{'07C', '2680', 'A'}): {'B04K03V04/4-10', 'B04K16V13/44'}
    # article, ακουση vs. ακουσει
    '2223': [{'2680', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 226C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 2199 rows have Offen=0
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'35', 'FΠ'} explains the most by agreement
    # > Combination {'35'} explains the following by posterity (compared to {'35',
'FΠ'}): {'B04K07V29/2'}
    # added δε
    '226C': [{'35'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 226.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 2 rows have Offen=2
```

```
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 18, FΠ, 03, P66
    '226': [{'18', 'FΠ', '03', 'P66'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2411.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 1 rows have Offen=1
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: FΠ, A, 01
    '2411': [{'FΠ', 'A', '01'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 249.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 22 rows have Offen=2
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 397, 226, 157
    '249': [{'397', '226', '157'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py
2561C1.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 9 rows have Offen=0
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'35', '1788', '05'} explains the most by agreement
    # > Combination {'35', '05'} explains the following by posterity (compared to
{'35', '1788', '05'}): {'B04K19V02 /30-32'}
    # αυτω vs. αυτον
    '2561C1': [{'35', '05'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py
2561C3.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 5 rows have Offen=0
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'18', '1561', '1321'} explains the most by agreement
    # > Combination {'1561', '1321'} explains the following by posterity (compared
to {'18', '1561', '1321'}): {'B04K16V33/26', 'B04K04V37/8-14'}
    # εξετε vs. εχετε
    '2561C3': [{'1561', '1321'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2561.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 7
    # > 5 rows have Offen=7
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 045, A
    '2561': [{'045', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2575.csv
```

```
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 19 rows have Offen=2
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 994, 18, A
    '2575': [{'994', '18', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2585.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 10
    # > 1 rows have Offen=10
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1788, A
    '2585': [{'1788', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2615.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 4
    # > 1 rows have Offen=4
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: FΠ, 1561, A
    '2615': [{'FΠ', '1561', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2680.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 7
    # > 9 rows have Offen=7
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'037C', '1320', 'A'} explains the most by agreement
    # > Combination {'1320', 'A'} explains the following by posterity (compared to
{'037C', '1320', 'A'}): {'B04K05V35/24-28'}
    # αγαλλιασθηναι vs. αγαλλιαθηναι
    '2680': [{'1320', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2718.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 9
    # > 11 rows have Offen=9
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'019C', '226', 'A'} explains the most by agreement
    # > Combination {'226', 'A'} explains the following by posterity (compared to
{'019C', '226', 'A'}): {'B04K13V37/6-8'}
    # article
    '2718': [{'226', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2766C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 143 rows have Offen=0
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 3 comb. left
    # > Human thought required (probably)...
    # > Combination {'35', '2786', '333'} explains the most by agreement
    # > Combination {'1546', '2786'} explains the following by posterity (compared
to {'35', '2786', '333'}): {'B04K16V13/22-28'}
    # προς vs. εις
```

```
    # > Combination {'18'} explains the following by posterity (compared to
{'1546', '2786'}): {'B04K19V31/46-50'}
    # word order
    '2766C': [{'18'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2766.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 8
    # > 6 rows have Offen=8
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'07C', '35', 'A'} explains the most by agreement
    # > Combination {'35', 'A'} explains the following by posterity (compared to
{'07C', '35', 'A'}): {'B04K03V04/4-10'}
    # article
    '2766': [{'35', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2768.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 7
    # > 15 rows have Offen=7
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 226, A
    '2768': [{'226', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2786.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 11
    # > 9 rows have Offen=11
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1320, A, 213
    '2786': [{'1320', 'A', '213'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 2790.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 29 rows have Offen=2
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'35', '038', 'A'} explains the most by agreement
    # > Too large a difference by posterity (28), skipping combination {'045', 'A'}
    '2790': [{'35', '038', 'A'}],  # simple, in the end

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 295.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 9
    # > 6 rows have Offen=9
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1320, A
    '295': [{'1320', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 317.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 2 rows have Offen=1
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 333, F1, A
    '317': [{'333', 'F1', 'A'}],  # simple
```

```
    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 333.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 4
    # > 2 rows have Offen=4
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 249, 1561, A
    '333': [{'249', '1561', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 33.csv
    # > Loaded 4095 rows
    # > Min 'Offen' is 5
    # > 256 rows have Offen=5
    # > Restricted to the best of each combination size: 9 comb. left
    # > Excluded rows with more combinations but no better results: 7 comb. left
    # > Human thought required (probably)...
    # > Combination {'A', '044', '019', '04', 'P66C*', '03', 'P75', '083', '029',
'04C1'} explains the most by agreement
    # > Combination {'A', '044', '019', '04', 'P66C*', '03', 'P75', '083', '04C1'}
explains the following by posterity (compared to {'A', '044', '019', '04', 'P66C*',
'03', 'P75', '083', '029', '04C1'}): {'B04K01V24/2'}
    # added οι
    # > Combination {'A', '044', '019', '04', 'P66C*', '03', 'P75', '04C1'}
explains the following by posterity (compared to {'A', '044', '019', '04', 'P66C*',
'03', 'P75', '083', '04C1'}): {'B04K04V14/43', 'B04K02V20/36'}
    # added εγω, εγειρεις vs. εγερεις
    # > Combination {'A', '044', '019', '04', '03', 'P75', '04C1'} explains the
following by posterity (compared to {'A', '044', '019', '04', 'P66C*', '03', 'P75',
'04C1'}): {'B04K12V16/24', 'B04K17V16/2-24', 'B04K11V41/18-22'}
    # article, om. "εκ του κοσμου ουκ εισιν καθως εγω ουκ ειμι εκ του κοσμου" - 33
also omits v15 - but v14 also ends "εκ του κοσμου" so that's just a big eye skip,
added αυτου
    # > Combination {'A', '044', '019', '04', '03', '04C1'} explains the following
by posterity (compared to {'A', '044', '019', '04', '03', 'P75', '04C1'}):
{'B04K12V36/30', 'B04K09V17/12-16', 'B04K01V42/12', 'B04K04V31/2'}
    # added article, word order, added δε, added δε
    # > Too large a difference by posterity (14), skipping combination {'A', '044',
'019', '04', '04C1'}
    # > Too large a difference by posterity (21), skipping combination {'A', '044',
'019', '04'}
    '33': [{'A', '044', '019', '04', '03', '04C1'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 357.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 9 rows have Offen=2
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 138, 045, A
    '357': [{'138', '045', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 35.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 30 rows have Offen=1
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'18', 'FΠ', '03', 'P66'} explains the most by agreement
    # > Combination {'18', '03', 'P66'} explains the following by posterity
(compared to {'18', 'FΠ', '03', 'P66'}): {'B04K07V33/12-14', 'B04K10V42/2-12',
'B04K13V37/6-8', 'B04K19V26/32', 'B04K12V14/14-16', 'B04K05V26/20-34'}
    # word order, word order, added article, ιδου vs. ιδε, αυτω vs. αυτο, word
order
    '35': [{'18', '03', 'P66'}],
```

```
    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 377.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 3
    # > 7 rows have Offen=3
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 807, 226, 03
    '377': [{'807', '226', '03'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 382.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 10
    # > 8 rows have Offen=10
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'1344', 'A'} explains the most by agreement
    # > Too large a difference by posterity (21), skipping combination {'07C',
'045', 'A'}
    '382': [{'1344', 'A'}],  # simple, in the end

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 397.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 5
    # > 9 rows have Offen=5
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'037C', '044', '597', 'A'} explains the most by agreement
    # > Combination {'044', '597', 'A'} explains the following by posterity
(compared to {'037C', '044', '597', 'A'}): {'B04K05V35/24-28'}
    # αγαλλιασθηναι vs. αγαλλιαθηναι
    '397': [{'044', '597', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 430.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 2 rows have Offen=1
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 249, 333, 03
    '430': [{'249', '333', '03'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 544.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 8
    # > 2 rows have Offen=8
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1561, A
    '544': [{'1561', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 579.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 21
    # > 4 rows have Offen=21
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'07C', 'A', '0211'} explains the most by agreement
    # > Combination {'A', '0211'} explains the following by posterity (compared to
{'07C', 'A', '0211'}): {'B04K05V02/26', 'B04K08V09/2-40'}
    # βηθεσδα vs. βηθζαθα or βηθσαιδα, om. verse - no good reason
    '579': [{'07C', 'A', '0211'}],
```

```
    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 597.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 4
    # > 1 rows have Offen=4
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: FΠ, F1, A, P75
    '597': [{'FΠ', 'F1', 'A', 'P75'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 732.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 15
    # > 8 rows have Offen=15
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1320, A
    '732': [{'1320', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 792.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 24
    # > 1 rows have Offen=24
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1029, A
    '792': [{'1029', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 799C.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 72 rows have Offen=0
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'1320', 'P66C*', 'FΠ'} explains the most by agreement
    # > Combination {'1320', 'FΠ'} explains the following by posterity (compared to
{'1320', 'P66C*', 'FΠ'}): {'B04K11V45/26-28'}
    # οσα εποιησεν vs. α εποιησεν
    '799C': [{'1320', 'FΠ'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 799.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 5
    # > 2 rows have Offen=5
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'017C', '213', 'FΠ', 'A'} explains the most by agreement
    # > Combination {'213', 'FΠ', 'A'} explains the following by posterity
(compared to {'017C', '213', 'FΠ', 'A'}): {'B04K14V28/33'}
    # added ειπον
    '799': [{'213', 'FΠ', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 807.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 11
    # > 5 rows have Offen=11
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'226', '35', 'A'} explains the most by agreement
```

```
    # > Too large a difference by posterity (34), skipping combination {'1463',
'A'}
    '807': [{'226', '35', 'A'}],  # simple, in the end

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 821.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 10
    # > 7 rows have Offen=10
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'2561C3', '597', 'A', 'F1'} explains the most by agreement
    # > Too large a difference by posterity (17), skipping combination {'FΠ', 'A',
'F1'}
    '821': [{'2561C3', '597', 'A', 'F1'}],  # simple, in the end

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 841.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 7
    # > 4 rows have Offen=7
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1320, A
    '841': [{'1320', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 865.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 8 rows have Offen=2
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 033, 226, 397
    '865': [{'033', '226', '397'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 869.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 15 rows have Offen=2
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 397, 35, A
    '869': [{'397', '35', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 892.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 6
    # > 144 rows have Offen=6
    # > Restricted to the best of each combination size: 4 comb. left
    # > Excluded rows with more combinations but no better results: 3 comb. left
    # > Human thought required (probably)...
    # > Combination {'044', 'A', '33', '033', '019', 'P75'} explains the most by
agreement
    # > Combination {'044', 'A', '33', '033', 'P75'} explains the following by
posterity (compared to {'044', 'A', '33', '033', '019', 'P75'}): {'B04K09V37/2-4',
'B04K05V04/8-20', 'B04K05V36/6-12', 'B04K13V08/6', 'B04K03V36/31', 'B04K06V07/6',
'B04K08V14/58-62', 'B04K04V50/20', 'B04K05V44/6-8'}
    # 019 is needed for the large addition 'B04K05V04/8-20': "κατα καιρον
κατεβαινεν εν τη κολυμβηθρα και"
    # > Too large a difference by posterity (25), skipping combination {'044', 'A',
'33', 'P75'}
    '892': [{'044', 'A', '33', '033', '019', 'P75'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 892SC.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
```

```
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 9925 rows have Offen=0
    # > Restricted to the best of each combination size: 4 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1071
    '892SC': [{'1071'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 892S.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 3
    # > 4 rows have Offen=3
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'1242C', '045', 'A'} explains the most by agreement
    # > Combination {'045', 'A'} explains the following by posterity (compared to
{'1242C', '045', 'A'}): {'B04K17V11/50-54'}
    # ους vs. ω
    '892S': [{'045', 'A'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py
'892S*.csv'
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 238 rows have Offen=1
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'35', '1546', 'F13'} explains the most by agreement
    # > Combination {'028', 'F13'} explains the following by posterity (compared to
{'35', '1546', 'F13'}): {'B04K15V06/42', 'B04K15V16/48-52', 'B04K18V10/45'}
    # βαλουσιν vs, βαλλουσιν, ο τι εαν vs. οταν or ο τι αν, article
    # NOTE: This hand shouldn't be here, as it should have been merged with 892S.
But the
    # S and the * seem to have ganged up and created a bug. We will live with it.
    # > ... and the other way round: {'B04K20V02/56'}
    # οιδα vs. οιδαμεν
    '892S*': [{'028', 'F13'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 992.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 5
    # > 10 rows have Offen=5
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: FΠ, 18, A
    '992': [{'FΠ', '18', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py 994.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 2 rows have Offen=0
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: F1, 1320, 019
    '994': [{'F1', '1320', '019'}],  # simple

    'A': [{}],  # obviously no ancestors

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py F13.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 23
```

```
    # > 4 rows have Offen=23
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 1561, A
    'F13': [{'1561', 'A'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py F1.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 6
    # > 2 rows have Offen=6
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'2192C', '892', '044', 'A', 'P75', '032'} explains the most by
agreement
    # > Combination {'044', 'A', '33', 'P75', '032'} explains the following by
posterity (compared to {'2192C', '892', '044', 'A', 'P75', '032'}): {'B04K08V01/2-
17', 'B04K05V04/8-20', 'B04K05V36/6-12', 'B04K09V18/16-24', 'B04K08V59/31',
'B04K04V09/32-36', 'B04K04V06/42', 'B04K04V09/4', 'B04K06V35/3', 'B04K12V25/12',
'B04K08V03/2-32', 'B04K07V41/3', 'B04K08V07/2-34', 'B04K09V01/12-14', 'B04K07V43/6-
12', 'B04K08V04/2-20', 'B04K07V32/2', 'B04K05V02/26', 'B04K07V53/2-14',
'B04K04V36/1', 'B04K06V71/18-22', 'B04K06V46/10-12', 'B04K12V36/2', 'B04K05V27/2-
12', 'B04K13V30/12-14', 'B04K13V20/36', 'B04K08V02/2-36', 'B04K01V20/18-22',
'B04K09V37/2-4', 'B04K08V06/2-41', 'B04K04V30/3', 'B04K12V40/40', 'B04K06V17/30-
32', 'B04K08V05/2-29', 'B04K07V33/6-8', 'B04K08V58/2'}
    # > ... and the other way round: {'B04K09V10/12-14', 'B04K04V47/24',
'B04K07V29/2', 'B04K12V26/4-8', 'B04K10V29/8-12', 'B04K18V15/15', 'B04K21V04/14',
'B04K21V20/4-12', 'B04K08V28/52-54', 'B04K11V40/22', 'B04K11V20/22', 'B04K02V19/2-
10', 'B04K21V04/12', 'B04K16V19/4', 'B04K01V32/18-26', 'B04K04V51/20',
'B04K06V10/42-44', 'B04K15V16/54', 'B04K10V23/22', 'B04K04V05/24', 'B04K05V02/20-
22', 'B04K21V14/10-14', 'B04K02V16/19', 'B04K01V28/4-8', 'B04K19V07/23',
'B04K06V09/8-10', 'B04K05V44/6-8'}
    # Too much change (36 vs. 27). We'll take the larger set.
    'F1': [{'2192C', '892', '044', 'A', 'P75', '032'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py FΠ.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 2
    # > 2 rows have Offen=2
    # > Restricted to the best of each combination size: 2 comb. left
    # > Excluded rows with more combinations but no better results: 2 comb. left
    # > Human thought required (probably)...
    # > Combination {'892', '044', 'A', '33', '033', 'P75', '032'} explains the
most by agreement
    # > Too large a difference by posterity (11), skipping combination {'892',
'044', 'A', '33', 'P75', '032'}
    'FΠ': [{'892', '044', 'A', '33', '033', 'P75', '032'}],  # simple, in the end

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py MT.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 1
    # > 2 rows have Offen=1
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 35, F1, 03
    'MT': [{'35', 'F1', '03'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py P45.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 1 rows have Offen=0
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
```

```
    # > Only one combination left: P66C*, 03, 138
    'P45': [{'P66C*', '03', '138'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py P60.csv
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 6 rows have Offen=0
    # > Restricted to the best of each combination size: 3 comb. left
    # > Excluded rows with more combinations but no better results: 3 comb. left
    # > Human thought required (probably)...
    # > Combination {'A', '397', '03', '05S'} explains the most by agreement
    # > Combination {'A', '397', '05S'} explains the following by posterity
(compared to {'A', '397', '03', '05S'}): {'B04K19V12/48-50', 'B04K18V39/18-22'}
    # αυτον vs. εαυτον, om. εν
    # > Combination {'A', '397'} explains the following by posterity (compared to
{'A', '397', '05S'}): {'B04K18V20/38', 'B04K18V34/10', 'B04K19V17/24-28'}
    # παντοτε vs. παντες, omit συ, ος vs. ο
    'P60': [{'A', '397'}],

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py
'P66C*.csv'
    # > Found Hinweis '<<' entries
    # > Loaded 9999 rows
    # > Restricting search to Hinweis '<<' rows
    # > Min 'Offen' is 0
    # > 1 rows have Offen=0
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: 070, 03, 1788, 01
    'P66C*': [{'070', '03', '1788', '01'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py P66.csv
    # > Loaded 9999 rows
    # > Min 'Offen' is 4
    # > 8 rows have Offen=4
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: A, F1, 397, P66C*
    'P66': [{'A', 'F1', '397', 'P66C*'}],  # simple

    # $ /home/ed/coding/CBGM/scripts/analyse_combinations_of_ancestors.py P75.csv
    # > Loaded 3 rows
    # > Min 'Offen' is 3
    # > 1 rows have Offen=3
    # > Restricted to the best of each combination size: 1 comb. left
    # > Excluded rows with more combinations but no better results: 1 comb. left
    # > Only one combination left: A, 03
    'P75': [{'A', '03'}],  # simple
}
```

Bibliography

Software

1. Andrew Edmondson
    a) My CBGM implementation: https://doi.org/10.5281/zenodo.1296287 or https://github.com/edmondac/CBGM
    b) Scripts for manipulating NEXUS files and related things: https://doi.org/10.5281/zenodo.1296329 or https://github.com/edmondac/nexus_scripts
    c) Stripey: https://doi.org/10.5281/zenodo.1296295 or https://github.com/edmondac/stripey
2. Archaeopteryx, an evolutionary tree explorer. Part of the forester collection of software: https://sites.google.com/site/cmzmasek/home/software/forester Accessed 14 March 2017.
    a) See Edmondson's fork of the code for automated tree creation here: https://github.com/edmondac/forester Accessed 14 March 2017.
3. CollateX
    a) Java-based service: http://collatex.net/ Accessed 30 December 2016. Versions used: Initially collatex-0.9 then collatex-tools-1.3, collatex-tools-1.5.1 and finally collatex-tools-1.7.1 (in January 2017).
    b) Python-based library: https://pypi.python.org/pypi/collatex Accessed 17 January 2017. Version 2.1.1 (in January 2017). The source code is here, although the releases are not tagged there: https://github.com/interedition/collatex Accessed 17 January 2017.
4. FigTree, http://tree.bio.ed.ac.uk/software/figtree/ Accessed 23 September 2014. Versions used 1.4.2 and 1.4.3.
5. Graphviz, http://www.graphviz.org/ Accessed 10 October 2017.[1]
    a) See https://graphviz.gitlab.io/_pages/pdf/dot.1.pdf for information about the different algorithms, including NEATO and SFDP. Accessed 26 August 2018.
6. Mesquite, http://mesquiteproject.org/ Accessed 17 January 2017. Maddison, W. P. and D.R. Maddison. (2017) *Mesquite: a modular system for evolutionary analysis.* Version used: 3.2 (build 801).

---

[1] See Gansner and North, 'An Open Graph Visualization System and Its Applications to Software Engineering'.

7. MrBayes, *Bayesian Inference of Phylogeny*, http://mrbayes.sourceforge.net/download.php Accessed 2 February 2017. Versions used 3.2.2 to 3.2.6.[2]
    a) Edmondson patch for MrBayes 3.2.6: https://github.com/edmondac/MrBayes (commit a0ad10f applied to upstream 3.2.6).
8. Network, http://www.fluxus-engineering.com/sharenet.htm Accessed 10 October 2017. Versions used 4.613, 5.0.0, 5.0.0.1.[3]
9. SplitsTree, http://www.splitstree.org/ Accessed 23 September 2014. Version used: 4.13.1[4]

General bibliography

10. Aland, Barbara, Kurt Aland, Johannes Karavidopoulos, Carlo Maria Martini, and Bruce Manning Metzger, eds. *Novum Testamentum Graece 27*. 9th corrected. Stuttgart: Deutsche Bibelgesellschaft, 2006.
11. Aland, Barbara, Kurt Aland, Johannes Karavidopoulos, Carlo Maria Martini, Bruce Manning Metzger, and Holger Strutwolf, eds. *Novum Testamentum Graece 28*. Stuttgart: Deutsche Bibelgesellschaft, 2012.
12. Aland, Barbara, Kurt Aland, Gerd Mink, Holger Strutwolf, and Klaus Wachtel, eds. *Novum Testamentum Graecum, Editio Critica Maior: IV Catholic Letters*. 2nd Revised Edition. Stuttgart: Deutsche Bibelgesellschaft, 2013.
13. ———, eds. *Novum Testamentum Graecum Editio Critica Maior: IV Catholic Letters: Part 1 Text: Installment 4: The The Second and Third Letter of John, the Letter of Jude*. Stuttgart: Deutsche Bibelgesellschaft, 2005.
14. Aland, Barbara, Kurt Aland, Gerd Mink, and Klaus Wachtel, eds. *Novum Testamentum Graecum Editio Critica Maior: IV Catholic Letters: Part 1 Text: Installment 1: James*. Stuttgart: Deutsche Bibelgesellschaft, 1997.
15. ———, eds. *Novum Testamentum Graecum Editio Critica Maior: IV Catholic Letters: Part 1 Text: Installment 2: The Letters of Peter*. Stuttgart: Deutsche Bibelgesellschaft, 2000.
16. ———, eds. *Novum Testamentum Graecum Editio Critica Maior: IV Catholic Letters: Part 1 Text: Installment 3: The First Letter of John*. Stuttgart: Deutsche Bibelgesellschaft, 2003.
17. Aland, Kurt. *Text Und Textwert Der Griechischen Handschriften Des Neuen Testaments*. Berlin: De Gruyter, 1987–2005.

---

[2] See Huelsenbeck and Ronquist, 'MRBAYES'; Ronquist and Deans, 'Bayesian Phylogenetics and Its Influence on Insect Systematics'; Ronquist and Huelsenbeck, 'MrBayes 3'; 'MrBayes Tutorial, v3.2'., Ronquist and Deans, 'Bayesian Phylogenetics and Its Influence on Insect Systematics'.

[3] See Bandelt, Forster, and Rohl, 'Median-Joining Networks for Inferring Intraspecific Phylogenies'.

[4] See Huson and Bryant, 'Application of Phylogenetic Networks in Evolutionary Studies'.

18. Aland, Kurt, and Barbara Aland. *The Text of the New Testament: An Introduction to the Critical Editions and to the Theory and Practice of Modern Textual Criticism*. Grand Rapids: Eerdmans and Leiden: Brill, 1987.
19. Aland, Kurt, Klaus Junack, Beate Köster, and Michael Welte. *Kurzgefaßte Liste der griechischen Handschriften des Neuen Testaments*, 2011. https://doi.org/10.1515/9783110877588.
20. Alexanderson, Bengt. 'Problems in the New Testament: Old Manuscripts and Papyri, the New Genealogical Method (CBGM) and the Editio Critica Maior (ECM)'. *Humaniora* 48 (2014).
21. Altekar, Gautam, Sandhya Dwarkadas, John P. Huelsenbeck, and Fredrik Ronquist. 'Parallel Metropolis Coupled Markov Chain Monte Carlo for Bayesian Phylogenetic Inference'. *Bioinformatics* 20, no. 3 (12 February 2004): 407–15. https://doi.org/10.1093/bioinformatics/btg427.
22. Anderson, Amy S. *The Textual Tradition of the Gospels: Family 1 in Matthew*. New Testament Tools and Studies, v. 32. Leiden ; Boston: Brill, 2004.
23. Andrews, Tara L. 'Analysis of Variation Significance in Artificial Traditions Using Stemmaweb'. *Digital Scholarship in the Humanities* 31, no. 3 (September 2016): 523–39. https://doi.org/10.1093/llc/fqu072.
24. Atallah, Mikhail J., and Susan Fox, eds. *Algorithms and Theory of Computation Handbook*. 1st ed. Boca Raton, FL, USA: CRC Press, Inc., 1998.
25. Bandelt, H. J., P. Forster, and A. Rohl. 'Median-Joining Networks for Inferring Intraspecific Phylogenies'. *Molecular Biology and Evolution* 16, no. 1 (1 January 1999): 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036.
26. Barbrook, Adrian C., Christopher J. Howe, Norman Blake, and Peter Robinson. 'The Phylogeny of The Canterbury Tales'. *Nature* 394, no. 6696 (August 1998): 839–839. https://doi.org/10.1038/29667.
27. Bentley, Richard. *Dr. Bentley's Proposals for Printing a New Edition of the Greek Testament, and St. Hierom's Latin Version: With a Full Answer to All the Remarks of a Late Pamphleteer*. J. Knapton, 1721.
28. ———. *Remarks upon a Late Discourse of Free-Thinking: In a Letter to N. N.* W. Thurlbourn, 1737.
29. Bergel, Giles, Christopher J. Howe, and Heather F. Windram. 'Lines of Succession in an English Ballad Tradition: The Publishing History and Textual Descent of *The Wandering Jew's Chronicle*'. *Digital Scholarship in the Humanities* 31, no. 3 (September 2016): 540–62. https://doi.org/10.1093/llc/fqv003.
30. Bhagwat, V. M., and B. V. Ramachandran. 'Malathion A and B Esterases of Mouse Liver-I'. *Biochemical Pharmacology* 24, no. 18 (15 September 1975): 1713–17.
31. Boc, A., A. B. Diallo, and V. Makarenkov. 'T-REX: A Web Server for Inferring, Validating and Visualizing Phylogenetic Trees and Networks'. *Nucleic Acids Research* 40, no. W1 (1 July 2012): W573–79. https://doi.org/10.1093/nar/gks485.

32. Bordalejo, Barbara. 'The Genealogy of Texts: Manuscript Traditions and Textual Traditions'. *Digital Scholarship in the Humanities* 31, no. 3 (September 2016): 563–77. https://doi.org/10.1093/llc/fqv038.

33. Bryant, David, and Vincent Moulton. 'NeighborNet: An Agglomerative Method for the Construction of Planar Phylogenetic Networks'. In *Algorithms in Bioinformatics*, edited by Roderic Guigó and Dan Gusfield, 2452:375–91. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002. https://doi.org/10.1007/3-540-45784-4_28.

34. Buneman, Peter. 'The Recovery of Trees from Measures of Dissimilarity'. *Mathematics in the Archaeological and Historical Sciences*, 1971. https://ci.nii.ac.jp/naid/10018379872/en/.

35. Buzzoni, Marina, Eugenio Burgio, Martina Modena, and Samuela Simion. 'Open versus Closed Recensions (Pasquali): Pros and Cons of Some Methods for Computer-Assisted Stemmatology'. *Digital Scholarship in the Humanities* 31, no. 3 (September 2016): 652–69. https://doi.org/10.1093/llc/fqw014.

36. Carlson, Stephen C. 'Coherence-Based Genealogical Method'. *Hypotyposeis* (blog), 1 December 2009. http://hypotyposeis.org/weblog/2009/12/coherence-based-genealogical-method.html.

37. ———. 'Comments on the Coherence-Based Genealogical Method'. *TC – A Journal of Biblical Textual Criticism* 20 (2015). http://rosetta.reltech.org/TC/v20/TC-2015-CBGM-Carlson.pdf.

38. ———. 'Fundamentals of New Testament Textual Criticism. By Stanley E. Porter and Andrew W. Pitts.' *The Journal of Theological Studies* 68, no. 1 (April 2017): 269–73. https://doi.org/10.1093/jts/flx065.

39. ———. 'The Text of Galatians and Its History'. PhD, Duke University, 2012. http://dukespace.lib.duke.edu/dspace/bitstream/handle/10161/5597/Carlson_duke_0066D_11426.pdf.

40. ———. *The Text of Galatians and Its History*. Wissenschaftliche Untersuchunger zum Neuen Testament. 2. Reihe 385. Tübingen, Germany: Mohr Siebeck, 2015.

41. Caruso, Carlo. 'Review: The Genesis of Lachmann's Method by Sebastiano Timpanaro; Glenn W. Most'. *The Modern Language Review* 102 (2007). https://doi.org/10.2307/20467293.

42. Clark, Michael A. 'The Catena of Nicetas of Heraclea and Its Johannine Text'. PhD, University of Birmingham, 2016. http://etheses.bham.ac.uk/6424/1/ClarkM16PhD.pdf.

43. 'Codex Sinaiticus: Experience the Oldest Bible. The Transcription.' Accessed 6 March 2018. http://codexsinaiticus.org/en/project/transcription_detailed.aspx.

44. Courtney, E. 'Review: The Genesis of Lachmann's Method by Sebastiano Timpanaro; Glenn W. Most'. *The Classical Journal* 102, no. 3 (2007): 312–315.

45. Darwin, Charles. *Journal of Researches into the Natural History and Geology of the Countries Visited during the Voyage of H.M.S. Beagle (Reprint of 3rd Edn., 1860)*. London, 2003.

46. ———. *On the Origin of Species by Means of Natural Selection, or, The Preservation of Favoured Races in the Struggle for Life*. London: John Murry, 1859.
47. De Vos, Ilse, Caroline Macé, and Koen Geuten. 'Comparing Stemmatological and Phylogenetic Methods to Understand the Transmission History of the "Florilegium Coislinianum"'. In *Ars Edendi Lecture Series Vol. II*, edited by Alessandra Bucossi and Erika Kihlman, 107–129. Acta Universitatis Stockholmiensis, Studia Latina Stockholmiensia LVIII. Stockholm University Publications, 2012.
48. Dress, Andreas, ed. *Basic Phylogenetic Combinatorics*. Cambridge, UK ; New York: Cambridge University Press, 2012.
49. Eagleton, Catherine, and Matthew Spencer. 'Copying and Conflation in Geoffrey Chaucer's Treatise on the Astrolabe: A Stemmatic Analysis Using Phylogenetic Software'. *Studies in History and Philosophy of Science Part A* 37, no. 2 (June 2006): 237–68. https://doi.org/10.1016/j.shpsa.2005.08.020.
50. Ehrman, Bart D. *The Orthodox Corruption of Scripture: The Effect of Early Christological Controversies on the Text of the New Testament*. 2nd edition. Oxford/New York:OUP, 2011.
51. Ehrman, Bart D., and Michael W. Holmes, eds. *The Text of the New Testament in Contemporary Research: Essays on the Status Quaestionis*. 2nd ed. New Testament Tools, Studies, and Documents, v. 42. Leiden ; Boston: Brill, 2013.
52. Elliott, James K. 'A New Edition of Nestle-Aland, Greek New Testament'. *The Journal of Theological Studies* 64, no. 1 (2013): 47–65.
53. ———. 'Review of "How We Got the New Testament: Text, Transmission, Translation." Porter, S.E.' *The Journal of Theological Studies* 65, no. 2 (2014): 672–674.
54. ———. 'Review of "Novum Testamentum Graecum: Editio Critica Maior", INTF (ed.)'. *The Journal of Theological Studies* 64, no. 1 (2013): 636–642.
55. Epp, Eldon Jay. 'In the Beginning Was the New Testament Text, but Which Text?' In *Texts and Traditions*, edited by Jeffrey Kloha and Peter Doble, 35–70. Brill, 2014. https://doi.org/10.1163/9789004273931_005.
56. ———. 'Textual Clusters: Their Past and Future in New Testament Textual Criticism'. In *The Text of the New Testament in Contemporary Research*, edited by Bart D. Ehrman and Michael W. Holmes, 519–77. Brill, 2012. https://doi.org/10.1163/9789004236554_021.
57. Fee, Gordon D. 'On the Types, Classification, and Presentation of Textual Variation'. In *Studies in the Theory and Method of New Testament Textual Criticism*, edited by Gordon D Fee and Eldon Jay Epp. Grand Rapids, Mich.: W.B. Eerdmans, 1993.
58. Felsenstein, Joseph. 'The Number of Evolutionary Trees'. *Systematic Zoology* 27, no. 1 (March 1978): 27. https://doi.org/10.2307/2412810.
59. Finney, Timothy. 'How to Discover Textual Groups'. *Digital Studies/Le Champ Numérique* 8, no. 1 (10 May 2018). https://doi.org/10.16995/dscn.291.

60. Foster, Paul. ‘Book Review: Recent Developments and Future Directions in New Testament Textual Criticism: Report on a Conference at the University of Edinburgh’. *Journal for the Study of the New Testament* 29, no. 2 (December 2006): 229–35. https://doi.org/10.1177/0142064X06072929.

61. Gäbel, Georg, Annette Hüffmeier, Gerd Mink, Holger Strutwolf, and Klaus Wachtel. ‘The CBGM Applied to Variants from Acts: Methodological Background’. *TC – A Journal of Biblical Textual Criticism* 20 (2015). http://rosetta.reltech.org/TC/v20/TC-2015-CBGM-background.pdf.

62. Gabriel, Edgar, Graham E. Fagg, George Bosilca, Thara Angskun, Jack J. Dongarra, Jeffrey M. Squyres, Vishal Sahay, et al. ‘Open MPI: Goals, Concept, and Design of a Next Generation MPI Implementation’. In *Proceedings, 11th European PVM/MPI Users’ Group Meeting*, 97–104. Budapest, Hungary, 2004.

63. Gansner, Emden R., and Stephen C. North. ‘An Open Graph Visualization System and Its Applications to Software Engineering’. *Software - Practice and Experience* 30, no. 11 (2000): 1203–1233.

64. Gelman, Andrew, and Donald B. Rubin. ‘Inference from Iterative Simulation Using Multiple Sequences’. *Statistical Science* 7, no. 4 (November 1992): 457–72. https://doi.org/10.1214/ss/1177011136.

65. Goodrich, Richard J., and Albert L. Lukaszewski, eds. *A Reader’s Greek New Testament*. 2nd ed. Grand Rapids, Mich: Zondervan, 2007.

66. Gregory, Caspar René. *Textkritik des Neuen Testaments*. Vol. 1. Hinrichs:Leipzig, 1900.

67. Gurry, Peter J. *A Critical Examination of the Coherence-Based Genealogical Method in New Testament Textual Criticism*. Leiden: Brill, 2017.

68. Halonen, Marko. ‘Computer-Assisted Stemmatology in Studying Paulus Juusten’s 16th-Century Chronicle *Catalogus et Ordinaria Successio Episcoporum Finlandensium*’. *Digital Scholarship in the Humanities* 31, no. 3 (September 2016): 578–93. https://doi.org/10.1093/llc/fqv004.

69. Haugen, Odd Einar. ‘The *Silva Portentosa* of Stemmatology: Bifurcation in the Recension of Old Norse Manuscripts’. *Digital Scholarship in the Humanities* 31, no. 3 (September 2016): 594–610. https://doi.org/10.1093/llc/fqv002.

70. Head, Peter. ‘Editio Critica Maior: An Introduction and Assessment’. *Tyndale Bulletin* 61, no. 1 (2010): 131–152.

71. Heikkilä, Tuomas, and Teemu Roos. ‘Thematic Section on Studia Stemmatologica’. *Digital Scholarship in the Humanities* 31, no. 3 (September 2016): 520–22. https://doi.org/10.1093/llc/fqw038.

72. Holmes, Michael W. *The Greek New Testament. SBL Edition*. SBL:Atlanta and Logos Bible Software:Bellingham, 2010.

73. Houghton, Hugh A. G. ‘Recent Developments in New Testament Textual Criticism’. *Early Christianity* 2, no. 2 (1 May 2011): 245–68. https://doi.org/10.1628/186870311795777463.

74. ———. *The Latin New Testament*. Oxford: OUP, 2016.

75. Houghton, Hugh A. G., and David C. Parker. *Textual Variation: Theological and Social Tendencies?* New Jersey: Gorgias Press, 2008.

76. Houghton, Hugh A. G., and Catherine J. Smith. ‘Digital Editing and the Greek New Testament’. In *Ancient Worlds in Digital Culture*, 110–127. Digital Biblical Studies. Brill, 2016.

77. Howe, Christopher J., Adrian C Barbrook, Linne R. Mooney, and Peter Robinson. ‘Parallels Between Stemmatology and Phylogenetics’. In *Studies in Stemmatology II*, edited by Pieter van Reenen, August den Hollander, and Margot van Mulken, 3–11. John Benjamins, 2004.

78. Howe, Christopher J., Adrian C. Barbrook, Matthew Spencer, Peter Robinson, Barbara Bordalejo, and Linne R. Mooney. ‘Manuscript Evolution’. *Trends in Genetics* 17, no. 3 (March 2001): 147–52. https://doi.org/10.1016/S0168-9525(00)02210-1.

79. Howe, Christopher J., Ruth Connolly, and Heather F. Windram. ‘Responding to Criticisms of Phylogenetic Methods in Stemmatology’. *SEL Studies in English Literature 1500-1900* 52, no. 1 (2012): 51–67. https://doi.org/10.1353/sel.2012.0008.

80. Howe, Christopher J., and Heather F. Windram. ‘Phylomemetics—Evolutionary Analysis beyond the Gene’. *PLoS Biology* 9, no. 5 (31 May 2011): e1001069. https://doi.org/10.1371/journal.pbio.1001069.

81. Huelsenbeck, John P., and Fredrik Ronquist. ‘MRBAYES: Bayesian Inference of Phylogenetic Trees’. *Bioinformatics* 17, no. 8 (1 August 2001): 754–55. https://doi.org/10.1093/bioinformatics/17.8.754.

82. Hüffmeier, Annette. ‘The CBGM Applied to Variants from Acts’. *TC – A Journal of Biblical Textual Criticism* 20 (2015). http://rosetta.reltech.org/TC/v20/TC-2015-CBGM-H%C3%BCffmeier.pdf.

83. Hurtado, Larry. ‘Muenster’. *Larry Hurtado’s Blog: Comments on the New Testament and Early Christianity (and Related Matters)* (blog). Accessed 6 March 2018. https://larryhurtado.wordpress.com/2010/07/05/muenster/.

84. Huson, Daniel H., and David Bryant. ‘Application of Phylogenetic Networks in Evolutionary Studies’. *Molecular Biology and Evolution* 23, no. 2 (1 February 2006): 254–67. https://doi.org/10.1093/molbev/msj030.

85. IGNTP. ‘IGNTP Guidelines for XML Transcriptions of New Testament Manuscripts Using the TEI P5’, 2012. http://epapers.bham.ac.uk/1656/10/IGNTP_XML_guidelines_1-3.pdf.

86. ———. ‘Online Electronic Edition of John’. Accessed 16 August 2018. http://www.iohannes.com/transcriptions/.

87. INTF. ‘“Genealogical Queries” v1 (for the 1st Edition of ECM Vol IV)’. Accessed 25 October 2013. http://intf.uni-muenster.de/cbgm/GenQ.html.

88. ———. ‘“Genealogical Queries” v2 (for the 2nd Edition of ECM Vol IV)’. Accessed 25 October 2013. http://intf.uni-muenster.de/cbgm2/GenQ.html.

89. INTF, and Cologne Center for eHumanities. '"Genealogical Queries" Phase 4 (for ECM Vol III)', 2017. http://ntg.cceh.uni-koeln.de/ph4/.

90. Jongkind, Dirk. 'On the Nature and Limitations of the Coherence Based Genealogical Method.' In *SBL Annual Meeting*, 2014. https://www.sbl-site.org/meetings/Congresses_Abstracts.aspx?MeetingId=25.

91. ———. 'On the Weighing and Counting of Variants: The Coherence Based Genealogical Method, Potential Ancestors and Statistical Significance'. Baltimore, 2013. https://www.sbl-site.org/meetings/Congresses_Abstracts.aspx?MeetingId=23.

92. ———. 'Review of "The Erotic Life of Manuscripts: New Testament Textual Criticism and the Biological Sciences", Yii-Jan Lin'. *Journal for the Study of the New Testament* 39, no. 5 (August 2017): 117–18.

93. ———. 'Review of "The Textual History of the Greek New Testament", Wachtel, Klaus and Michael Holmes Eds.' *Journal for the Study of the New Testament* 35, no. 5 (2013): 147.

94. Kloppenborg, John S., and Judith H. Newman, eds. *Editing the Bible: Assessing the Task Past and Present*. Atlanta: Society of Biblical Literature, 2012. http://www.jstor.org/stable/j.ctt32bzxf.

95. Lachmann, Karl. *Novum Testamentum Graece*. Berlin: G. Reimer, 1831. http://archive.org/details/novumtestamentu00lachgoog.

96. Lafleur, Didier. 'Which Criteria for Family 13 (f^{13}) Manuscripts?' *Novum Testamentum* 54, no. 2 (2012): 105–48.

97. Lakner, Clemens, Paul van der Mark, John P. Huelsenbeck, Bret Larget, and Fredrik Ronquist. 'Efficiency of Markov Chain Monte Carlo Tree Proposals in Bayesian Phylogenetics'. *Systematic Biology* 57, no. 1 (February 2008): 86–103. https://doi.org/10.1080/10635150801886156.

98. Lee, Arthur. 'Numerical Taxonomy Revisited: John Griffith, Cladistic Analysis and St. Augustine's Quaestiones in Heptateuchum'. *Studia Patristica* XX, Originally delivered at the 10th International Conference of Patristic Studies, Oxford, 1987. (1989): 24–32.

99. Levenshtein, V.I. 'Binary Codes Capable of Correcting Deletions, Insertions, and Reversals'. *(Russian). English Translation in Soviet Physics Doklady* 163, no. 4 (1965): 845–48.

100. Lewis, Paul O. 'A Likelihood Approach to Estimating Phylogeny from Discrete Morphological Character Data'. *Systematic Biology* 50, no. 6 (1 November 2001): 913–25. https://doi.org/10.1080/106351501753462876.

101. Lin, Yii-Jan. *The Erotic Life of Manuscripts: New Testament Textual Criticism and the Biological Sciences*. New York, New York: Oxford University Press, 2016.

102. Maas, Paul. *Textual Criticism (Translation of Textkritic, 3rd Edn, Leipzig, 1957)*. Oxford, 1958.

103. Macé, C, C Peersman, Riccardo Mazza, J Noret, E Van Mulken, Evert Wattel, Paolo Canettieri, et al. 'Testing Methods on an Artificially Created Textual Tradition'. *Macé, Caroline: The Evolution of Texts : Confronting Stemmatological and Genetical Methods ; Proceedings of the International Workshop Held in Louvain-La-Neuve on September 1 - 2, 2004, Ist. Ed. e Poligrafici Internazionli, 255-283 (2006)*, 2006.

104. Macé, Caroline, ed. *The Evolution of Texts: Confronting Stemmatological and Genetical Methods ; Proceedings of the International Workshop Held in Louvain-La Neuve on September 1-2, 2004*. Linguistica Computazionale, 24/25. Pisa: Istituti editoriali e poligrafici internazionali, 2006.

105. Maddison, David R., David L. Swofford, and Wayne P. Maddison. 'Nexus: An Extensible File Format for Systematic Information'. Edited by David Cannatella. *Systematic Biology* 46, no. 4 (1 December 1997): 590–621. https://doi.org/10.1093/sysbio/46.4.590.

106. Matzke, Nick. 'MrBayes Lab (Handout), PRINCIPLES OF PHYLOGENETICS, Integrative Biology 200A'. University of California, Berkeley, 2008. http://ib.berkeley.edu/courses/ib200a/labs/ib200a_lab11_mrbayes.pdf.

107. 'Mesquite FAQ'. Accessed 6 March 2018. http://mesquiteproject.wikispaces.com/FAQ.

108. Metzger, Bruce Manning. *A Textual Commentary on the Greek New Testament*. 2nd ed. Stuttgart: Deutsche Bibelgesellschaft, 1994.

109. ———. *The Text of the New Testament. Its Transmission, Corruption, and Restoration*. Oxford: OUP, 1964.

110. Mink, Gerd. 'Contamination, Coherence, and Coincidence in Textual Transmission: The Coherence-Based Genealogical Method (CBGM) as a Complement and Corrective to Existing Approaches.' In *The Textual History of the Greek New Testament. Changing Views in Contemporary Research.*, edited by Klaus Wachtel and Michael W. Holmes. Atlanta: SBL, 2011.

111. ———. 'Editing and Genealogical Studies: The New Testament'. *Literary and Linguistic Computing* 15 (2000): 51–56.

112. ———. 'Eine umfassende Genealogie der neutestamentlichen Überlieferung'. *New Testament Studies* 39 (1993): 481–99.

113. ———. 'Guide to "Genealogical Queries"', 2008. http://intf.uni-muenster.de/cbgm/guide_en.html.

114. ———. 'Problems of a Highly Contaminated Tradition: The New Testament – Stemmata of Variants as a Source of a Genealogy for Witnesses.' In *Studies in Stemmatology II*, edited by Pieter van Reenen, August den Hollander, and Margot van Mulken, 13–85. Amsterdam ; Philadelphia: John Benjamins Pub, 2004.

115. ———. 'The Coherence-Based Genealogical Method – What Is It About?', 2002. http://www.uni-muenster.de/INTF/Genealogical_method.html.

116. ———. 'The Coherence-Based Genealogical Method (CBGM) – Introductory Presentation, 1.0'. 2009. http://egora.uni-muenster.de/intf/bindata/CBGM_Presentation.zip.

117. ———. 'Was verändert sich in der Textkritik durch die Beachtung genealogischer Kohärenz?' In *Recent Developments in Textual Criticism, New Tesament, other Early Christian and Jewish Literature*, edited by Wilhelmus J. C. Weren and Dietrich-Alex Koch, STAR 8:39–68. Assen: Royal can Gorcum, 2003.

118. ———. 'Zur Stemmatisierung neutestamentlicher Handschriften'. In *Bericht der Hermann Kunst-Stiftung zur Förderung der neutestamentlichen Textforschung für die Jahre 1979-1981*, 110–14. Münster: INTF, 1982.

119. Morrill, Michael Bruce. 'A Complete Collation and Analysis of All Greek Manuscripts of John 18'. PhD, University of Birmingham, 2012. http://etheses.bham.ac.uk/3733/.

120. Moule, Charles Francis Digby. *An Idiom Book of New Testament Greek*. 2. ed., repr. Cambridge: Univ. Press, 1971.

121. 'MrBayes Tutorial, v3.2'. Accessed 4 August 2016. http://mrbayes.sourceforge.net/wiki/index.php/Tutorial_3.2.

122. M.W. Robinson, Peter, and Robert O'Hara. 'Report on the Textual Criticism Challenge 1991'. *Bryn Mawr Classical Review* 3 (1992): 331–337.

123. 'NTVMR: "New Testament Virtual Manuscript Room"'. Accessed 17 January 2014. http://ntvmr.uni-muenster.de/en_GB/home.

124. O'Hara, Robert, and Peter Robinson. 'Computer-Assisted Methods of Stemmatic Analysis'. *Occasional Papers of the Canterbury Tales Project* 1 (1993): 53–74.

125. Parker, David C. *An Introduction to New Testament Manuscripts and Their Texts*. Cambridge: CUP, 2008.

126. ———. *Codex Bezae: An Early Christian Manuscript and its Text*. Cambridge: CUP, 1992.

127. ———. *Codex Sinaiticus*. London: The British Library, 2010.

128. ———. 'Making Editions of the New Testament Today'. In *Jesus, Paul, and Early Christianity: Studies in Honour of Henk Jan de Jonge.*, edited by Harm W. Hollander, Johannes Tromp, and Rieuwerd Buitenwerf. Leiden: Brill, 2008.

129. ———. 'New Testament Textual Traditions in Byzantium'. In *The New Testament in Byzantium*, edited by Robert S. Nelson and Derek Krueger. Washington: Dumbarton Oaks Research Library and Collection, 2016.

130. ———. 'Review of Alexanderson, Problems in the New Testament: Old Manuscripts and Papyri, the New Genealogical Method (CBGM) and the Editio Critica Maior (ECM)', Unpublished at time of writing.

131. ———. *Textual Scholarship and the Making of the New Testament*. Oxford: OUP, 2012.

132. ———. ‘Through a Screen Darkly: Digital Texts and the New Testament’. *Journal for the Study of the New Testament* 25, no. 4 (June 2003): 395–411. https://doi.org/10.1177/0142064X0302500401.

133. Parker, David C., Klaus Wachtel, Bruce Morrill, and Ulrich Schmid. ‘The Selection of Greek Manuscripts to Be Included in the International Greek New Testament Project’s Edition of John in the Editio Critica Maior’. In *Studies on the Text of the New Testament and Early Christianity*, 287–328. Brill, 2015. https://doi.org/10.1163/9789004300026_015.

134. Perrin Jr. Jac D. ‘Family 13 in Saint John’s Gospel’. PhD, University of Birmingham, 2012. http://etheses.bham.ac.uk/4482/1/Perrin13PhD.pdf.

135. Platnick, Norman I., and H. Don Cameron. ‘Cladistic Methods in Textual, Linguistic, and Phylogenetic Analysis’. *Systematic Zoology* 26, no. 4 (1977).

136. Porter, Stanley E. *Fundamentals of New Testament Textual Criticism*. Grand Rapids, Michigan: William B. Eerdmans Publishing Company, 2015.

137. ———. *How We Got the New Testament*. Grand Rapids: Baker, 2013.

138. Racine, Jean-François, and Thomas C. Geer. ‘Analyzing and Categorizing New Testament Greek Manuscripts’. In *The Text of the New Testament in Contemporary Research*, edited by Bart D. Ehrman and Michael W. Holmes, 497–518. Brill, 2012. https://doi.org/10.1163/9789004236554_020.

139. Rambaut, Andrew. ‘How to Read a Phylogenetic Tree’, 2013. http://epidemic.bio.ed.ac.uk/how_to_read_a_phylogeny.

140. ———. ‘How to Read a Phylogenetic Tree (Text Only)’, 2013. https://github.com/rambaut/mol-evo-epi/blob/master/docs/how_to_read_a_tree.md.

141. Reenen, Pieter van, August den Hollander, and Margot van Mulken, eds. *Studies in Stemmatology II*. Amsterdam: John Benjamins Publishing Company, 2004. https://doi.org/10.1075/z.125.

142. Reenen, Pieter van, and Margot van Mulken, eds. *Studies in Stemmatology*. Amsterdam: John Benjamins Publishing Company, 1996. https://doi.org/10.1075/z.79.

143. Robinson, Maurice A., and William G. Pierpont, eds. *The New Testament in the Original Greek: Byzantine Textform, 2005*. 1st ed. Southborough, Mass: Chilton Book Pub, 2005.

144. Robinson, Peter. ‘Computer-Assisted Stemmatic Analysis and “Best-Text” Historical Editing’. In *Studies in Stemmatology*, edited by Pieter van Reenen and Margot van Mulken, 71. Amsterdam: John Benjamins Publishing Company, 1996. https://doi.org/10.1075/z.79.04rob.

145. ———. ‘Four Rules for the Application of Phylogenetics in the Analysis of Textual Traditions’. *Digital Scholarship in the Humanities* 31, no. 3 (September 2016): 637–51. https://doi.org/10.1093/llc/fqv065.

146. Robinson, Peter, and Robert O’Hara. ‘Cladistic Analysis of an Old Norse Manuscript Tradition’. *Research in Humanities Computing* 4 (1996): 115–137.

147. Ronquist, Fredrik, and Andrew R. Deans. ‘Bayesian Phylogenetics and Its Influence on Insect Systematics’. *Annual Review of Entomology* 55, no. 1 (January 2010): 189–206. https://doi.org/10.1146/annurev.ento.54.110807.090529.

148. Ronquist, Fredrik, and John P. Huelsenbeck. ‘MrBayes 3: Bayesian Phylogenetic Inference under Mixed Models’. *Bioinformatics (Oxford, England)* 19, no. 12 (12 August 2003): 1572–74.

149. Roos, Teemu, and Tuomas Heikkilä. ‘Evaluating Methods for Computer-Assisted Stemmatology Using Artificial Benchmark Data Sets’. *Literary and Linguistic Computing* 24, no. 4 (1 December 2009): 417–33. https://doi.org/10.1093/llc/fqp002.

150. Roos, Teemu, and Yuan Zou. ‘Analysis of Textual Variation by Latent Tree Structures’, 567–76. IEEE, 2011. https://doi.org/10.1109/ICDM.2011.24.

151. Saitou, N., and M. Nei. ‘The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees’. *Molecular Biology and Evolution* 4 (1987): 406–425.

152. Semple, Charles, and M. A. Steel. *Phylogenetics*. Oxford lecture series in mathematics and its applications 24. Oxford ; New York: Oxford University Press, 2003.

153. Soden, Hermann F. von. *Die Schriften des Neuen Testaments*. 6th ed. Göttingen: Vandenhoeck und Ruprecht, 1913.

154. Spencer, Matthew, Barbara Bordalejo, Peter Robinson, and Christopher J. Howe. ‘How Reliable Is a Stemma? An Analysis of Chaucer’s Miller’s Tale’. *Literary and Linguistic Computing* 18, no. 4 (1 November 2003): 407–22. https://doi.org/10.1093/llc/18.4.407.

155. Spencer, Matthew, Barbara Bordalejo, Li-San Wang, Adrian C. Barbrook, Linne R. Mooney, Peter Robinson, Tandy Warnow, and Christopher J. Howe. ‘Analyzing the Order of Items in Manuscripts of “The Canterbury Tales”’. *Computers and the Humanities* 37, no. 1 (2003): 97–109.

156. Spencer, Matthew, Elizabeth A Davidson, Adrian C Barbrook, and Christopher J. Howe. ‘Phylogenetics of Artificial Manuscripts’. *Journal of Theoretical Biology* 227, no. 4 (April 2004): 503–11. https://doi.org/10.1016/j.jtbi.2003.11.022.

157. Spencer, Matthew, Klaus Wachtel, and Christopher J. Howe. ‘Representing Multiple Pathways of Textual Flow in the Greek Manuscripts of the Letter of James Using Reduced Median Networks’. *Computers and the Humanities* 38, no. 1 (2004): 1–14.

158. ———. ‘The Greek Vorlage of the Syra Harclensis: A Comparative Study on Method in Exploring Textual Genealogy’. *TC – A Journal of Biblical Textual Criticism* 7 (August 2002). http://purl.org/TC/vol07/SWH2002/.

159. Streeter, B.H. *The Four Gospels: A Study of Origins*. London: Maxmillan, 1964.

160. Strutwolf, Holger. ‘Scribal Practices and the Transmission of Biblical Texts: New Insights from the Coherence-Based Genealogical Method’. In *Editing the Bible: Assessing the Task Past and Present*, edited by John S. Kloppenborg and Judith H. Newman. Atlanta: Society of Biblical Literature, 2012. http://www.jstor.org/stable/j.ctt32bzxf.

161. Strutwolf, Holger, Georg Gäbel, Annette Hüffmeier, Gerd Mink, and Klaus Wachtel. *Novum Testamentum Graecum. Editio Critica Maior: III Acts*. Stuttgart: Deutsche Bibelgesellschaft, 2017.

162. Sutton, Darrell A. 'CAVEAT EMPTOR. Remarks on the Nestle/Aland Novum Testamentum Graece'. *The Ds Commentary on Books*, no. 1 (Autumn/Fall 2014). https://www.researchgate.net/publication/279537091_The_Ds_Commentary_on_Books_1.

163. Swofford, David L. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10*, 2002.

164. Taylor Farnes, Alan. 'Scribal Habits in Selected New Testament Manuscripts, Including Those with Surviving Exemplars'. PhD, University of Birmingham, 2017. http://etheses.bham.ac.uk/8285/1/Farnes18PhD.pdf.

165. TC Editors. 'Special Feature: The Coherence-Based Genealogical Method. Editorial Introduction.' *TC – A Journal of Biblical Textual Criticism* 20 (2015). http://rosetta.reltech.org/TC/v20/TC-2015-CBGM-intro.pdf.

166. Tehrani, Jamshid, Quan Nguyen, and Teemu Roos. 'Oral Fairy Tale or Literary Fake? Investigating the Origins of *Little Red Riding Hood* Using Phylogenetic Network Analysis'. *Digital Scholarship in the Humanities* 31, no. 3 (September 2016): 611–36. https://doi.org/10.1093/llc/fqv016.

167. Timpanaro, Sebastiano. *The Genesis of Lachmann's Method*. Edited by Glenn W. Most. Chicago: University of Chicago Press, 2005.

168. Tischendorf, Constantin von. *Novum Testamentum Graece, Editio octava critica maior*. Leipzig, 1869.

169. Tregelles, Samuel P. *The Greek New Testament, Edited from Ancient Authorities, with Their Various Readings in Full, and the Latin Version of Jerome*. London:Bagster; Stewart, 1857.

170. Wachtel, K., and D.C. Parker. 'The Joint IGNTP/INTF Editio Critica Maior of the Gospel of John: Its Goals and Their Significance for New Testament Scholarship'. Halle, 2005. http://epapers.bham.ac.uk/754/1/2005_SNTS_WachtelParker.pdf.

171. Wachtel, Klaus. 'A Stemmatological Approach in Editing the Greek New Testament: The Coherence-Based Genealogical Method'. In *Advances in Digital Scholarly Editing: Papers Presented at the DiXiT Conferences in The Hague, Cologne, and Antwerp*, edited by Peter Boot. Leiden: Slidestone Press, 2017.

172. ———. 'Constructing Local Stemmata for the ECM of Acts: Examples'. *TC – A Journal of Biblical Textual Criticism* 20 (2015). http://rosetta.reltech.org/TC/v20/TC-2015-CBGM-examples.pdf.

173. ———. 'Reconstructing the Initial Text of Acts: Principles and Criteria'. SBL, 2017. Slides: https://www.academia.edu/35232745/Reconstructing_the_Initial_Text_of_Acts_Principles_and_Criteria_-SBL2017.

174. ———. ‘Textual Changes in NA28’, 2013. http://intf.uni-muenster.de/NA28/files/TextChangesNA28.pdf.

175. ———. ‘The Coherence Method and History’. *TC – A Journal of Biblical Textual Criticism* 20 (2015). http://rosetta.reltech.org/TC/v20/TC-2015-CBGM-history.pdf.

176. ———. ‘The Coherence-Based Genealogical Method: A New Way to Reconstruct the Text of the Greek New Testament’. In *Editing the Bible: Assessing the Task Past and Present*, edited by John S. Kloppenborg and Judith H. Newman. Atlanta: Society of Biblical Literature, 2012. http://www.jstor.org/stable/j.ctt32bzxf.

177. ———. ‘Towards a Global Stemma of the Greek New Testament Textual Tradition: Methodological Approach’. Antwerp, 2016. http://dixit.uni-koeln.de/wp-content/uploads/wachtel_CBGM_DiXiT2016_2.pdf.

178. ———. ‘Towards a Redefinition of External Criteria: The Role of Coherence in Assessing the Origin of Variants’. In *Textual Variation: Theological and Social Tendencies?*, edited by Hugh A. G. Houghton and David C. Parker, 6:109–127. Texts and Studies 3. Piscataway: Gorgias Press, 2008.

179. Wachtel, Klaus, and Michael W. Holmes, eds. *The Textual History of the Greek New Testament. Changing Views in Contemporary Research*. Atlanta: SBL, 2011.

180. Wallace, Daniel B. ‘Challenges in New Testament Textual Criticism for the Twenty-First Century’. *Journal of the Evangelical Theological Society* 52, no. 1 (March 2009): 79–100.

181. Wasserman, Tommy. ‘Criteria for Evaluating Readings in New Testament Textual Criticism’. In *The Text of the New Testament in Contemporary Research*, edited by Bart D. Ehrman and Michael W. Holmes, 579–612. Brill, 2012. https://doi.org/10.1163/9789004236554_022.

182. ———. ‘Historical and Philological Correlations and the CBGM as Applied to Mark 1:1’. *TC – A Journal of Biblical Textual Criticism* 20 (2015). http://rosetta.reltech.org/TC/v20/TC-2015-CBGM-Wasserman.pdf.

183. ———. ‘The Coherence Based Genealogical Method as a Tool for Explaining Textual Changes in the Greek New Testament’. *Novum Testamentum* 57 (2015): 206–218.

184. ———. *The Epistle of Jude: Its Text and Transmission*. Stockholm: Almqvist & Wiksell, 2006.

185. Wasserman, Tommy, and Peter J. Gurry. *A New Approach to Textual Criticism: An Introduction to the Coherence-Based Genealogical Method*. Resources for Biblical Study, number 80. Atlanta : Stuttgart: SBL Press ; Deutsche Bibelgesellschaft, 2017.

186. Welsby, Alison S. ‘A Textual Study of Family 1 in The Gospel of John.’ PhD, University of Birmingham, 2011. http://etheses.bham.ac.uk/3338/1/Welsby12PhD.pdf.

187. West, Martin. *Textual Criticism and Editorial Technique*. Stuttgart: Teubner, 1973.

188. Westcott, Brooke Foss, and Fenton J. A. Hort. *The New Testament in the Original Greek*. Cambridge:Macmillan, 1881.

189. ———. *The New Testament in the Original Greek, Introduction and Appendix*. Cambridge:Macmillan, 1882.

190. Wettlaufer, Ryan D. *No Longer Written: The Use of Conjectural Emendation in the Restoration of the Text of the New Testament, the Epistle of James As a Case Study*. Leiden: Brill, 2013.

191. Wisse, Frederik. *The Profile Method for the Classification and Evaluation of Manuscript Evidence, as Applied to the Continuous Greek Text of the Gospel of Luke*. Grand Rapids: Eerdmans, 1982.