

1 Short Communication

2

3 **Epigenetic sampling effects: nephrectomy modifies the clear cell**
4 **renal cell cancer methylome**

5

6

7 Christophe Van Neste^{1,2}, Alexander Laird^{3,4,5}, Fiach O'Mahony^{3,4}, Wim Van Criekinge⁶, Dieter
8 Deforce¹, Filip Van Nieuwerburgh¹, Thomas Powles^{7,8}, David J. Harrison^{4,9,*}, Grant D.
9 Stewart^{3,4,§} and Tim De Meyer^{6,§}

10

11 ***Correspondence:**

12 david.harrison@st-andrews.ac.uk

13 9 School of Medicine,

14 University of St Andrews,

15 KY16 9AJ St Andrews, UK

16

17 § These authors contributed equally to this work

18

19 **Affiliations:**

20

21 ¹ Laboratory of Pharmaceutical Biotechnology, Ghent University, Ottergemsesteenweg 460,
22 9000 Gent, Belgium.

23 ² Center for Medical Genetics, Ghent University, De Pintelaan 185, 9000 Gent, Belgium.

24 ³ Edinburgh Urological Cancer Group, Institute of Genetics and Molecular Medicine, University
25 of Edinburgh, EH4 2XU, Edinburgh, UK.

26 ⁴ Scottish Collaboration On Translational Research into Renal Cell Cancer (SCOTRRCC).

27 ⁵ MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of
28 Edinburgh, EH4 2XU Edinburgh, UK.

29 ⁶ Biobix: Laboratory of Bioinformatics and Computational Genomics, Department of
30 Mathematical Modeling, Statistics and Bioinformatics, Ghent University, Coupure Links 653,
31 9000 Ghent, Belgium.

32 ⁷ Renal Cancer Unit, The Royal Free Hospital, NW3 2QG London, UK.

33 ⁸ Centre for Experimental Cancer Medicine, Bart's Cancer Institute, Queen Mary University of
34 London, EC1A 7BE London, UK.

35 ⁹ School of Medicine, University of St Andrews, KY16 9AJ St Andrews, UK.

36

37

38 **Abstract**

39

40 Currently, it is unclear to what extent sampling procedures affect the epigenome. Here, this
41 phenomenon was evaluated by studying the impact of artery ligation on DNA methylation in
42 clear cell renal cell cancer. To this end, DNA methylation profiles between vascularised tumour
43 biopsy samples and devascularized nephrectomy samples from two individuals were
44 compared. The relevance of significantly altered methylation profiles was validated in an
45 independent clinical trial cohort. We found that six genes were differentially methylated in
46 the test samples, of which four were linked to ischaemia or hypoxia (*REXO1L1*, *TLR4*, *hsa-mir-*
47 *1299*, *ANKRD2*). Three of these genes were also found to be significantly differentially
48 methylated in the validation cohort, indicating that the observed effects are genuine. Based
49 on these results, we conclude that the impact of sampling procedures in clinical epigenetic
50 studies should be considered, particularly after inducing hypoxia/ischemia, which occurs in
51 many oncological surgery procedures through which tissues are harvested for translational
52 research.

53

54 **Keywords:** Hypoxia; Cancer epigenetics; DNA methylation; Sampling effects

55

56

57

58 **1 Introduction**

59

60 The use of clinical samples in epigenetics research has become routine [1–3]. However, the
61 nature of the sampling procedures may significantly compromise the resulting epigenetic
62 profiles, leading to an “epigenetic observer effect”. In renal cell carcinoma (RCC), dynamic
63 molecular changes occur over time and with therapy, which require serial tissue samples for
64 elucidation [4,5]. Despite the development of standard operating procedures for tissue
65 acquisition and biobanking, less attention is paid to ensuring constant, robust pre-collection
66 conditions, such as warm ischaemia time, than to post-collection handling and processing
67 procedures, thus failing to avoid variation due to pre-analytical factors [6]. There are
68 recommended tissue sampling guidelines from organisations such as the Confederation of
69 Cancer Biobanks, which advise that warm ischaemia time should be minimised as much as
70 possible prior to freezing of fresh tissue samples, but this is very difficult to measure and
71 standardise [7]. Several studies have looked at the procurement conditions on RNA biomarker
72 expression and the effect on previously identified cancer biomarkers [8], but so far there have
73 been no such studies on DNA methylation.

74 Prolonged ischaemia leads to higher levels of tissue hypoxia, which has been shown to
75 induce DNA demethylation in e.g. hepatoma cells [9]. This demethylation has been shown to
76 be induced by methionine adenosyltransferase 2A (*MAT2A*), of which the expression is
77 positively regulated by *HIF-1α* [9], which in turn is negatively regulated by *VHL* [10]. *VHL* plays
78 a crucial role in RCC development [11]. Hypoxia, methylation and RCC form, therefore, an
79 intricate network of which the components cannot be studied separately. Yet, different RCC
80 sampling procedures might obfuscate the results by influencing the hypoxic conditions. More
81 than most tumours, RCC sampling is challenging as the majority of the extirpative procedures
82 deal with minimally invasive approaches. It is well established that changes occur in mRNA
83 levels with increasing time after renal artery clamping in RCC [12]. Here, we assessed the effect
84 of renal artery clamping on the RCC methylome.

85

86 **Patient samples and methods**

87

88 Patient-matched sample sets were obtained from two patients who, at the time of open
89 cytoreductive nephrectomy for metastatic clear cell RCC (ccRCC), had fresh ccRCC tumour

90 biopsies taken prior to ligation of the renal artery after which matched fresh frozen tumour
91 samples were harvested following ligation and division of the renal artery and removal of the
92 kidney as per a previously described approach [5]. A total of 12 samples was collected, i.e., 3
93 biopsy and 3 nephrectomy samples from each patient. These samples were obtained as part
94 of the Scottish Collaboration On Translational Research into Renal Cell Cancer (SCOTRRCC)
95 study (East of Scotland Research Ethics Service REC 1: 10/S1402/33). For validation purposes,
96 but also to evaluate the potential impact on clinical epigenetics research, matched tumour
97 samples taken at the time of diagnostic renal tumour biopsy and subsequent nephrectomy
98 were obtained from 14 patients with metastatic ccRCC. Following primary tumour biopsy,
99 these patients were treated with three cycles of sunitinib (18 weeks) followed by
100 cytoreductive nephrectomy after 2 weeks of sunitinib as part of the Upfront Sunitinib
101 (SU011248) Therapy Followed by Surgery in Patients with Metastatic Renal Cancer: a Pilot
102 Phase II Study (SuMR; ClinicalTrials.gov identifier: NCT01024205) [13]. All samples used in this
103 study are listed in Supplementary file 1.

104 Extraction of genomic DNA was performed using a Qiagen DNeasy Blood and Tissue
105 (Qiagen, UK) kit following the manufacturer's instructions. DNA methylation analyses using
106 MBD sequencing was carried out as described previously [14]. A MethylCap kit (Diagenode,
107 Belgium) was used for capturing methylated fragments from 500 ng starting material.
108 Massively parallel sequencing of these fragments was subsequently performed on an Illumina
109 HiSeq 2000 machine (Illumina, San Diego, CA, USA).

110 Raw data files were mapped using BOWTIE to the human reference genome
111 Hg19/GRCh37, and summarized using an in-house developed Map of the Human Methylome
112 [15] consisting of a putative genome-wide overview of potentially methylated loci
113 ("methylation cores"). Further data analyses were performed using Python 3.4.3 and R 3.2.1.
114 The Bioconductor quantro software package (1.2) was used to assess quantile normalization
115 assumptions [16], whereas the limma software package (3.24.15) was used to identify regions
116 featured by differential methylation patterns. Prior to limma analysis, the samples were
117 quantile normalized and transformed using the voom function [17]. Only methylation cores
118 that referred to annotated promoter regions (including exon1), and that had at least an
119 average coverage of one mapped fragment per sample, were withheld for analysis. Low
120 coverage loci are featured by a too low power to be detected as differentially methylated,
121 thereby unnecessarily inflating the amount of hypotheses tested, which justifies their

122 unsupervised removal from the dataset. Regions with a False Discovery Rate (FDR) < 0.05 were
123 selected for further assessment using the sample dataset [13]. For the validation dataset, the
124 FDR estimation (Benjamini-Hochberg) was based on the amount of loci to be validated.

125

126 **Results and discussion**

127

128 A quantro test was performed to check the suitability of the samples for quantile
129 normalization, which was used for limma-voom. A p -value of 0.53 was obtained after 1000
130 permutations, implying that there were no global differences in the distributions between the
131 non-ischaeamic biopsy and ischaemic nephrectomy samples. A limma-voom data analysis was
132 subsequently performed using quantile normalization. In Table 1 seven regions are listed that
133 were found to be differentially methylated up to an FDR of 0.05 (Fig 1). Only two of these
134 seven (*AC232323.1* and *ANKRD2*) exhibited a relative hypomethylation in the ischaemic
135 nephrectomy samples compared to the non-ischaeamic samples. In one of these (*AC232323.1*)
136 a different region (8325757) exhibited a significant relative hypermethylation.

137 The differentially methylated regions with a FDR < 0.05 were subsequently subjected
138 to validation using methylation data obtained from matched biopsy and nephrectomy
139 samples from the SuMR clinical trial. Three of the seven regions identified in the test set, were
140 again significantly altered at the same level within the validation cohort (*AC232323.1* region
141 8325757, *REXO1L1* and *OR6Q1*). A binomial test – using the FDR threshold as expected
142 probability – rejected the null hypothesis that this would have occurred randomly ($p = 0.004$).
143 We therefore hypothesise that the shared results between the two studies are caused by
144 ischaemic conditions. Of the 7 loci identified in the test set, 6 were found to have the same
145 fold change direction in the validation cohort. However, assuming a 50% random chance of
146 having the same fold change direction, the number of similar direction of change was not
147 found to be significant ($p = 0.125$). A lower FDR cut-off of 10% yielded 36 significant regions,
148 yet this cut-off was deemed not sufficiently conservative as the fraction that could be
149 validated in the validation cohort was too low to reject the null hypothesis of significant
150 overlap ($p = 0.085$, data not shown).

151 Methylation meta-analyses have shown that DNA methylation is a critical event in
152 tumorigenesis [18]. It is, therefore, surprising that an analysis of the effect of tissue
153 procurement on DNA methylation has so far not been performed, as it has in other molecular

154 analyses. From our test and validation results it is clear that ischaemic conditions, induced as
155 part of the surgical procedure, may lead to differential methylation.

156 The results of the study presented here indicate that the RCC methylome may be
157 modulated following renal artery ligation. Global gene demethylation in samples was not
158 observed, but consistent demethylation of at least one individual gene (*ANKRD2*) was found.
159 These results hold considerable significance for translational methylation research for solid
160 tumours obtained by extirpative surgery, especially where minimally invasive surgical
161 approaches are used. The most significantly affected gene, *AC232323.1*, encodes a long non-
162 coding RNA (lncRNA) product. According to LNCipedia [19] this lncRNA is linked to the second
163 most significantly affected gene, *REXO1L1*, and has potential direct biological relevance
164 (<http://www.lncipedia.org/db/transcript/lnc-REXO1L2P-1:1>). *REXO1L1* was found to be
165 hypermethylated following renal artery clamping. *REXO1L1* deletions have been linked to
166 increased apoptosis under certain conditions [20], such as intense hypoxia, which may
167 underlie the results presented above [21]. The third significantly affected gene, *TLR4*, plays a
168 crucial role in kidney ischemia/reperfusion injury [22], and *Hsa-mir-1299* also has a possible
169 role in apoptosis through interaction with *PIM1* [23]. The *OR6Q1* gene encodes an olfactory
170 receptor and is, therefore, an unlikely candidate, although it should be noted that there is
171 evidence for a limited expression in other tissues as well [24]. Finally, *ANKRD2* belongs to the
172 conserved muscle ankyrin repeat protein (MARP) family. Expression of MARPs has been found
173 to be induced in response to physiologic stress, injury, hypoxia and hypertrophy [25,26] and
174 the *ANKRD2* mRNA level has indeed been found to be upregulated under specific hypoxic
175 conditions. Our results indicate that this upregulation may be brought about by demethylation
176 of a specific methylation core.

177 Despite the test and validation cohorts used, this study should be considered as proof
178 of principle with a necessarily low power, as it was deemed unethical to expose subjects to an
179 extra pre-operative percutaneous biopsy procedure without clinical benefit. Also, in the
180 validation cohort, patients were treated with sunitinib, the time between biopsy and
181 nephrectomy was longer (20 weeks) and the sequencing depths were generally more shallow,
182 implying that additional discrepancies between both datasets may have been caused by
183 treatment, temporal or other technical reasons.

184 Taken together, our data indicate that in both a test and a validation cohort renal
185 artery ligation modulates gene methylation in a biologically relevant fashion. As with any

186 research using surgically resected clinical samples, future methylation studies must be
187 designed to include robust and well-documented sample procurement techniques in order to
188 take these findings into account.

189

190 **Author's contributions**

191 Conception and design: DJH, AL, GDS, TDM, TP. Development of methodology: CVN, GDS, DJH,
192 TP, TDM. Acquisition of data (acquired and managed patients, provided facilities, etc.): GDS,
193 FOM, AL, DJH, TP. Analysis and interpretation of data (e.g., statistical analysis, biostatistics,
194 computational analysis): CVN, TDM, GDS. Writing, review, and/or revision of the manuscript:
195 CVN, AL, FOM, WVC, DD, FVN, TP, DJH, GDS, TDM. Study supervision: GDS, DJH, TDM, WVC.

196

197 **Funding**

198 This work was supported by the Chief Scientist Office, Scotland (ETM37; GDS, DJH), Cancer
199 Research UK (Experimental Cancer Medicine Centre; TP, London, DJH, Edinburgh), Medical
200 Research Council (AL, DJH), Royal College of Surgeons of Edinburgh Robertson Trust (AL),
201 Melville Trust AL), Renal Cancer Research Fund (GDS), Kidney Cancer Scotland (GDS) and an
202 educational grant from Pfizer (TP). CVN and TDM were funded by Ghent University
203 Multidisciplinary Research Partnership 'Bioinformatics: from nucleotides to networks'.

204

205 **Compliance with ethical standards**

206

207 **Conflict of Interest:** The authors declare that they have no conflict of interest.

208

209 **Informed consent:** Informed consent was obtained from all participants included in the study.

210

211

212 **Legend tot the Figure**

213

214 **Figure 1** Sampling procedure differences. For each of the regions with a FDR < 0.05, boxplots
215 of the methylation counts are shown. The methylation core region is shown under the gene
216 label.

217

218

219

220 **References**

- 221 1. Ferraro, Altered, Primary chromatin structures and their implications in cancer
222 development. *Cell. Oncol.* **39**, 195 (2016) doi:10.1007/s13402-016-0276-6
223
- 224 2. Taucher, V., Mangge, H., Haybaeck, J., Non-coding RNAs in pancreatic cancer: challenges
225 and opportunities for clinical application. *Cell. Oncol.* **39**, 295 (2016) doi: 10.1007/s13402-016-
226 0275-7
227
- 228 3. Vitiello, M., Tuccoli, A., Poliseno, L., Long non-coding RNAs in cancer: implications for
229 personalized therapy. *Cell. Oncol.* **38**, 17 (2015) doi: 10.1007/s13402-014-0180-x
230
- 231 4. Sharpe, K., Stewart, G.D., Mackay, A., Van Neste, C., Rofe, C., Berney, D., Kayani, I., Bex, A.,
232 Wan, E., O'Mahony, F.C., O'Donnell, M., Chowdhury, S., Doshi, R., Ho-Yen, C., Gerlinger, M.,
233 Baker, D., Smith, N., Davies, B., Sahdev, A., Boleti, E., Meyer, T.D., Crieckinge, W.V., Beltran, L.,
234 Lu, Y.-J., Harrison, D.J., Reynolds, A.R., Powles, T., The effect of VEGF-targeted therapy on
235 biomarker expression in sequential tissue from patients with metastatic clear cell renal cancer.
236 *Clin. Cancer Res.* **19**, 6924-6934 (2013) doi:10.1158/1078-0432.ccr-13-1631
237
- 238 5. Stewart, G.D., O'Mahony, F.C., Laird, A., Eory, L., Lubbock, A.L.R., Mackay, A., Nanda, J.,
239 O'Donnell, M., Mullen, P., McNeill, S.A., Riddick, A.C., Berney, D., Bex, A., Aitchison, M.,
240 Overton, I.M., Harrison, D.J., Powles, T., Sunitinib treatment exacerbates intratumoral
241 heterogeneity in metastatic renal cancer. *Clin. Cancer Res.* **21**, 4212-4223 (2015)
242 doi:10.1158/1078-0432.ccr-15-0207
243
- 244 6. Stewart, G.D., O'Mahony, F.C., Powles, T., Riddick, A.C.P., Harrison, D.J., Faratian, D., What
245 can molecular pathology contribute to the management of renal cell carcinoma? *Nature Rev.*
246 *Urol.* **8**, 255-265 (2011) doi:10.1038/nrrol.2011.43
247
- 248 7. Confederation of Cancer Biobanks: Biobank Quality Standard - Collecting, Storing and
249 Providing Human Biological Material and Data for Research (2014) [http://ccb.ncri.org.uk/wp-](http://ccb.ncri.org.uk/wp-content/uploads/2014/03/Biobank-quality-standard-Version-1.pdf)
250 [content/uploads/2014/03/Biobank-quality-standard-Version-1.pdf](http://ccb.ncri.org.uk/wp-content/uploads/2014/03/Biobank-quality-standard-Version-1.pdf)
251
- 252 8. Freidin, M.B., Bhudia, N., Lim, E., Nicholson, A.G., Cookson, W.O., Moffatt, M.F., Impact of
253 collection and storage of lung tumor tissue on whole genome expression profiling. *J. Mol.*
254 *Diagn.* **14**, 140-148 (2012) doi:10.1016/j.jmoldx.2011.11.002
255
- 256 9. Liu, Q., Liu, L., Zhao, Y., Zhang, J., Wang, D., Chen, J., He, Y., Wu, J., Zhang, Z., Liu, Z., Hypoxia
257 induces genomic DNA demethylation through the activation of HIF-1a and transcriptional
258 upregulation of MAT2A in hepatoma cells. *Mol. Cancer Ther.* **10**, 1113 (2011)
259 doi:10.1158/1535-7163.mct-10-1010
260
- 261 10. Kaelin Jr, W.G., The von Hippel-Lindau tumour suppressor protein: O2 sensing and cancer.
262 *Nat. Rev. Cancer* **8**, 865-873 (2008) doi:10.1038/nrc2502
263
- 264 11. Motzer, R., Michaelson, M., Redman, B., Hudes, G., Wilding, G., Figlin, R., Ginsberg, M.,
265 Kim, S., Baum, C., DePrimo, S., Li, J., Bello, C., Theuer, C., George, D., Rini, B., Activity of
266 SU11248, a multitargeted inhibitor of vascular endothelial growth factor receptor and

267 platelet-derived growth factor receptor, in patients with metastatic renal cell carcinoma. *J.*
268 *Clin. Oncol.* **24**, 16-24 (2006) doi:10.1200/JCO.2005.02257

269

270 12. Liu, N.W., Sanford, T., Srinivasan, R., Liu, J.L., Khurana, K., Aprelikova, O., Valero, V.,
271 Bechert, C., Worrell, R., Pinto, P.A., Yang, Y., Merino, M., Linehan, W.M., Bratslavsky, G. Impact
272 of ischemia and procurement conditions on gene expression in renal cell carcinoma. *Clin.*
273 *Cancer Res.* **19**, 42-49 (2013) doi:10.1158/1078-0432.ccr-12-2606

274

275 13. Stewart, G., Powles, T., Van Neste, C., Meynert, A., O'Mahony, F., Laird, A., Deforce, D.,
276 Van Nieuwerburgh, F., Trooskens, G., Van Criekinge, W., De Meyer, T., Harrison, D.J. Dynamic
277 epigenetic changes to VHL occur with sunitinib in metastatic clear cell renal cancer.
278 *Oncotarget* **7**, 25241–25250 (2016) doi:10.18632/oncotarget.8308

279

280 14. De Meyer, T., Mampaey, E., Vlemmix, M., Denil, S., Trooskens, G., Renard, J.-P., De
281 Keulenaer, S., Dehan, P., Menschaert, G., Van Criekinge, W., Quality evaluation of methyl
282 binding domain based kits for enrichment DNA-methylation sequencing. *PLoS ONE* **8**, 59068
283 (2013) doi:10.1371/journal.pone.0059068

284

285 15. Biobix: Map of the Human Methylome. [http://www.biobix.be/map-of-the-human-](http://www.biobix.be/map-of-the-human-methylome/mhm-version-2/)
286 [methylome/mhm-version-2/](http://www.biobix.be/map-of-the-human-methylome/mhm-version-2/)

287

288 16. Hicks, S., Irizarry, R., Quantro: a data-driven approach to guide the choice of an appropriate
289 normalization method. *Genome Biol.* **16**, 117 (2015) doi:10.1186/s13059-015-0679-0

290

291 17. Law, C.W., Chen, Y., Shi, W., Smyth, G.K., Voom: precision weights unlock linear model
292 analysis tools for RNA-seq read counts. *Genome Biol.* **15**, 29 (2014) doi:10.1186/gb-2014-15-
293 2-r2

294

295 18. Esteller, M., Epigenetic gene silencing in cancer: the DNA hypermethylome. *Hum. Mol.*
296 *Genet.* **16**, 50-59 (2007) doi:10.1093/hmg/ddm018

297

298 19. Volders, P.-J., Verheggen, K., Menschaert, G., Vandepoele, K., Martens, L., Vandesompele,
299 J., Mestdagh, P., An update on LNCipedia: a database for annotated human lncRNA sequences.
300 *Nucleic Acids Res.* **43**, 174-180 (2014) doi:10.1093/nar/gku1060

301

302 20. D'Apice, M.R., Novelli, A., di Masi, A., Biancolella, M., Antoccia, A., Gullotta, F., Licata, N.,
303 Minella, D., Testa, B., Nardone, A.M., Palmieri, G., Calabrese, E., Biancone, L., Tanzarella, C.,
304 Frontali, M., Sangiuolo, F., Novelli, G., Pallone, F. Deletion of REXO1L1 locus in a patient with
305 malabsorption syndrome, growth retardation, and dysmorphic features: a novel recognizable
306 microdeletion syndrome? *BMC Med. Genet.* **16**, 20 (2015) doi:10.1186/s12881-015-0164-3

307

308 21. Greijer, A., van der Groep, P., Kemming, D., Shvarts, A., Semenza, G., Meijer, G., van de
309 Wiel, M., Belien, J., van Diest, P., van der Wall, E., Up-regulation of gene expression by hypoxia
310 is mediated predominantly by hypoxia-inducible factor 1 (hif-1). *J. Pathol.* **206**, 291-304 (2005)
311 doi:10.1002/path.1778

312

313 22. Wu, H., Chen, G., Wyburn, K.R., Yin, J., Bertolino, P., Eris, J.M., Alexander, S.I., Sharland,

314 A.F., Chadban, S.J., Tlr4 activation mediates kidney ischemia/reperfusion injury. *J. Clin. Invest.*
315 **117**, 2847-2859 (2007) doi:10.1172/jci31008
316

317 23. Liu, Z., He, W., Gao, J., Luo, J., Huang, X., Gao, C., Computational prediction and
318 experimental validation of a novel synthesized pan-pim inhibitor pi003 and its apoptosis-
319 inducing mechanisms in cervical cancer. *Oncotarget* **6**, 8019-8035 (2015)
320 doi:10.18632/oncotarget.3139
321

322 24. Safran, M., Dalah, I., Alexander, J., Rosen, N., Stein, T.I., Shmoish, M., Nativ, N., Bahir, I.,
323 Doniger, T., Krug, H., Sirota-Madi, A., Olender, T., Golan, Y., Stelzer, G., Harel, A., Lancet, D.,
324 GeneCards version 3: the human gene integrator. *Database (Oxford)* **2010**, baq020 (2010)
325 doi:10.1093/database/baq020
326

327 25. Miller, M.K., Bang, M.-L., Witt, C.C., Labeit, D., Trombitas, C., Watanabe, K., Granzier, H.,
328 McElhinny, A.S., Gregorio, C.C., Labeit, S., The muscle ankyrin repeat proteins: Carp,
329 ankrd2/arpp and darp as a family of titin filament-based stress response molecules. *J. Mol.*
330 *Biol.* **333**, 951-964 (2003) doi:10.1016/j.jmb.2003.09.012
331

332 26. Band, M., Joel, A., Avivi, A., The muscle ankyrin repeat proteins are hypoxia-sensitive: In
333 vivo mrna expression in the hypoxia-tolerant blind subterranean mole rat, spalax ehrenbergi.
334 *J. Mol. Evol.* **70**, 1-12 (2009) doi:10.1007/s00239-009-9306-6