Article:
Bonthron, DT orcid.org/0000-0001-8132-8179 and Foulkes, WD (2017) Genetics meets Pathology - an increasingly important relationship. Journal of Pathology, 241 (2). pp. 119-122. ISSN 0022-3417

https://doi.org/10.1002/path.4849

Genetics meets Pathology – an increasingly important relationship

David T. Bonthron[1,*] and William D. Foulkes[2,*]

1 Section of Genetics, Leeds Institute of Biomedical and Clinical Sciences, Faculty of Medicine and Health, University of Leeds, Leeds, UK

2 Departments of Human Genetics, Oncology and Medicine, McGill University, Montreal, QC, Canada

Correspondence to: David Bonthron, Centenary Professor of Molecular Medicine, Level 9 Wellcome Trust Brenner Building, St James's University Hospital, Beckett Street, Leeds LS9 7TF, UK. E-mail: d.t.bonthron@leeds.ac.uk. William Foulkes, James McGill Professor, Departments of Human Genetics, Oncology and Medicine, McGill University, Montreal, QC, H3T 1E2, Canada. E-mail: william.foulkes@mcgill.ca

Conflicts of interest statement: DTB is a Senior Editor and WDF an Associate Editor of The Journal of Pathology.

Running head: Genetics meets Pathology – an increasingly important relationship

Abstract

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/path.4849

The analytical power of modern methods for DNA analysis has outstripped our capability to interpret and understand the data generated. To make good use of this genomic data in a biomedical setting (whether for research or diagnosis), it is vital that we understand the mechanisms through which mutations affect biochemical pathways and physiological systems. This lies at the centre of what genetics is all about, and it is the reason why genetics and genomics should go hand in hand whenever possible. In this Annual Review Issue of the Journal of Pathology, we have assembled a collection of 16 expert reviews covering a wide range of topics. Through these, we illustrate the power of genetic analysis to improve our understanding of normal physiology and disease pathology, and thereby to think in rational ways about clinical management.

KEY WORDS genetics, genomics, DNA sequencing, precision medicine, complex traits, rare diseases, molecular oncology, evolution, genome stability, clonality, mTOR, inflammasome, polyposis, colon cancer, prostate cancer, brain tumour, PARP, DNA repair, carcinoma in situ, tuberous sclerosis, Parkinson's disease, PINK1, Crispr/Cas9, stem cells, muscular dystrophy, ciliopathy, vascular malformation, pulmonary hypertension, TGF-beta, autoimmunity, atopy, inflammatory bowel disease, breast cancer, mitochondrial disease

We live at a time when medical and scientific strategies often seem to struggle to establish themselves without the aid of fashionable new titles. The term "precision medicine" – only recently evolved from its apparently extinct ancestor "stratified medicine" – enjoys a particular vogue at the moment. The more pedantic of us may be inclined to regard the adjectival noun as superfluous; after all, what has medicine been developing towards all these many years, if not greater precision? Nonetheless, this phrase carries the implication that today there exist powerful new tools, with which the classification of disease can be refined. Better classification (*i.e.* better pathology) should of course translate into better treatment and outcome.

Principal among these new tools is genetics – but should that perhaps be genomics? These two words are sometimes used interchangeably, which is a good enough reason of itself for pausing to consider the distinction. Genomics is to genetics as anatomy is to physiology; they are intimately related, but the former is more "what", and the latter "how". The powerful modern technologies that have captured the imaginations of the biomedical policy-makers are largely "genomic" – in particular, the ability to determine virtually complete human DNA sequences at trivial cost. For diagnostic purposes, it is assumed that the very large amount of genomic variability revealed in this way can and will be interpreted in some way and deployed for clinical benefit.

It is at this point that the pluripotent champions of genomic technology tend to start differentiating. For some, genomic data are there to be used empirically. In this camp, DNA sequence variation is no more special than any other data type. Correlations between genotypes and phenotypes, if robust, are valid endpoints in themselves,

which may be used for clinical management. Though it may be forthcoming, there need not be any understanding of the "how". This approach tends of necessity to be applied to the study of "complex" or non-Mendelian traits, since the sequence variants that are found to be statistically associated with particular phenotypes may not be susceptible to experimental analysis or mechanistic understanding. This issue is discussed further below.

"Genetics", though, is a bit different, and it is this that we have mostly chosen to focus on in the present collection of review articles. The word feels much older than "genomics" (the coining of which most members of the editorial team can still remember). However, in the sense of "the study of heredity", as presently used, "genetics" is little more than a century old, the Oxford English Dictionary crediting its first use to William Bateson (1905). Bateson himself was a great proponent of the idea that important biological lessons can be learnt from the study of rare genetic mutants. Brought forward a century, this idea is still fresh, and indeed remains a driving force in research. The repeatedly fulfilled belief is that human genetic pathology will instruct us as to normal gene function, and in favourable cases may allow us to infer the likely consequences of modulating the activity of the gene product.

Even die-hards now concede that humans are important subjects for genetic study, complementary to more tractable experimental organisms [1]. Whether in the germline or in somatic tissues, animal models of genetic disease suffer from the twin limitations that (a) they may fail to replicate a human pathology faithfully and (b) it may be infeasible to study an allelic series comprising more than a few mutations.

Here lie both of the great strengths and hence attractions of the human as a genetic model; the allelic diversity observable in naturally occurring genetic accidents in humans, and also the level of phenotypic detail that can be characterised, both vastly exceed that of any laboratory animal.

In human rare disease genetics, we have now reached a stage at which most recognized human Mendelian phenotypes have had their genes identified. As a consequence, the reviews here can focus less on genes *per se*, and more on mechanisms and pathways that have been illuminated by these discoveries. In his often-cited Cambridge inaugural lecture, Bateson compared his newly-defined science of genetics to astronomy, and exalted it for showing the way to "*novelty and adventure…hardly to be excelled*" [2]. However, even he might have been surprised by the diversity and complexity of the processes that have revealed their secrets under genetic scrutiny.

We begin this issue with a prime example of such complexity: McDermott and colleagues summarize a huge array of phenotypes that relate to impairment of the mechanisms regulating host responses to pathogens [3]. The physiological mechanisms that have been implicated in these disorders range from signalling through familiar cytokine axes, to aspects that were quite unknown before human genetic studies – such as the intimate links between viral disease, innate immunity and maintenance of host genome integrity [4]. One of the frequently vaunted (but less easily achieved) goals of "genomic medicine" is of course the advent of new, rationally designed therapies. Insights from these autoinflammatory disorders have allowed highly effective therapeutic approaches based on interleukin-1 blockade to be

deployed not only in rare monogenic autoinflammatory disorders but also in commoner diseases such as gout, in which similar disease mechanisms are at work.

Another major success of genetics in unravelling a common but enigmatic disorder is described by Brown [5] in her review of atopic eczema. The identification of filaggrin as a key genetic contributor to the pathogenesis of this disorder has not only provided a mechanistic explanation for the development of atopy (previously regarded primarily as an immunological multi-system disorder), but has also focused attention on integrity of the epidermal barrier as a key therapeutic target. Given population prevalences of 10% or more for atopic disease, even marginal effects on prevention and treatment have enormous potential in this area.

Inflammatory bowel disease (IBD) comprises a further important group of disorders in which immunological mechanisms play important pathogenic roles, and in which the genetic contributions appear complex. Uniken Venema *et al*. [6] directly address the important technical challenge referred to above; that of bridging the gap between statistics (genome-wide association studies) and mechanistic understanding. The successful recapitulation of IBD features in animal models based on several genetic loci identified by GWAS in IBD gives cause for optimism in this regard.

A large proportion of the many present-day industrial-scale deployments of genomic technology are directed not at inherited germ-line pathologies, but at the more pervasive problem of cancer. Actually, many instructive parallels can be recognized between genetics of the germ-line and tumour (somatic) genetics; not least the importance, both in the evolution of species and in tumorigenesis, of genome

rearrangements and of the sequential selection of mutations. As far as evolution is concerned, the accurate classification of species was an essential precursor to the work of Darwin and other evolutionary biologists, who changed how we think about biology forever. In pathology, classification of tumours has long been based on tissue of origin and morphological distinctions. While immunohistochemistry has had a major role in diagnostics, the molecular revolution has changed the entire rationale behind classification, and now no pathologist would attempt to classify a problematic tumour without some attention to the underlying genomic structure. Chiang and Ellison [7] exemplify this in the case of paediatric brain tumours, where not only is wholesale reclassification under way, but some distinct entities, such as CNS neuroblastoma with FOXR2 activation, have recently been identified solely on the basis of the molecular findings.

While paediatric tumours are in general rare, the same cannot be said for prostate cancer. De Bono and colleagues [8] summarise the extraordinary recent progress in the understanding of the molecular basis of advanced prostate cancer, whereby massively parallel sequencing of both tumour and constitutional DNA and RNA, together with methylation studies, have led to the development of numerous targeted therapies. These therapies, fully informed by the underlying genomic and epigenomic defects, are best exemplified by the use of PARP inhibitors in the treatment of men whose tumours lack normal BRCA2 protein [9]. Durable responses, sometimes measured in years, have been observed in men who previously would have had only months to live. Pathologists can play an important role in identifying candidates for genetic testing.

These advances would not have been possible without a deep understanding of the genomics, but many of these concepts will be new to pathologists. For example, Graham and Sottoriva [10] discuss neutral evolution and selection, as well as graduated and punctuated evolution, which are well-known to those working in evolution-related fields, but which have perhaps surprising relevance for the study of cancer genomes. One of the most recent examples of the relevance of understanding these evolutionary processes is the discovery that some tumours rapidly accumulate a very large number of focal rearrangements as a result of a one-time shattering of their genomes – "chromothripsis" [11]. Tumours that undergo chromothripsis at an early stage of their development may grow rapidly, foiling attempts at early diagnosis, and rendering some treatments ineffective [12].

Many chemotherapeutic agents work because they damage DNA. Dosages and scheduling determines that this damage occurs preferentially in tumour cells rather than in normal cells. O'Driscoll's detailed review [13] focus on defects in components of the DNA replication machinery. Loss of control of replication results in a wide spectrum of largely congenital disorders, such as Seckel and Meier-Gorlin syndrome, emphasizing its critical role in normal development. Cancer is also a consequence of mutations in the genes encoding these proteins (as discussed in the final cancer review of this Annual Review Issue by Glaire *et al*. [14]). One area of increasing interest is disorders of replication stress, and therapies are now being developed to promote this, hoping to result in mitotic catastrophe and tumour cell death.

Vastly better outcomes for patients with breast cancer have been achieved by improvements in treatment and to a certain extent, by earlier diagnosis. The latter,

though, has come at the price of over-diagnosis [15], particularly of early stage ("stage 0") breast cancer, also known as ductal carcinoma in situ (DCIS). Casasent *et al* [16] take on the challenge of using modern genomic technology to answer a critical question - which patients with DCIS are most likely to progress to invasive ductal carcinoma (IDC)? They apply three models to the available data - independent evolution, evolutionary bottlenecks and multiclonal invasion - to try to answer this critical question. In so doing, they outline the clinical consequences of each model. In the independent model, targeting any one biomarker would be unlikely to be helpful, whereas in the bottleneck and multiclonal models, identifying the key "truncal" mutations, that are common to DCIS and IDC could be of value, with the added complexity of the need to consider interactions in the multiclonal model. The authors posit that new technologies may help in distinguishing between these models, but consider the bottleneck model to be the most consistent with the existing data.

Lam *et al* [17] bring us up to date with an inherited cancer syndrome – tuberous sclerosis - that exemplifies some of the excitement, as well as the major challenges of clinical cancer genetics. For example, it is a remarkably pleiotropic syndrome and because a sizeable fraction of cases may have mosaic mutations not detected in blood DNA by Sanger sequencing, subtle manifestations may go unnoticed. But because TSC1 and TSC2 have inhibitory inputs to the mTOR signalling pathway, and loss of TSC1/2 results in its activation, drugs such as sirolimus, that inhibit mTORC1, have had remarkable success in treating some of the benign but serious tumours that occur in this syndrome.

Tomlinson and colleagues [14] round out the cancer section of this Annual Review Issue with their update on polymerase proofreading-associated polyposis (PPAP), a new polyposis syndrome identified by the groups of Richard Houlston and Ian Tomlinson several years ago. They use PPAP as an opportunity to point out the necessity to examine both the germline and the somatic variants present in a tumour and to include consideration of identified variants in the evolving field of precision medicine. While recognizing the potential power of this approach, they rightly point out the Achilles' heel – when individual variants are so rare, and therapies are based on "exceptional responders", how can one combine a desire for precision with the kind of evidence base that epidemiologists require and publicly-funded health systems will expect?

A key area of difficulty in somatic cell genetics, both in cancer and when considering other conditions caused by mosaic mutations [18], is the detection of mutations that may be at low level and spatially restricted. This is also an important consideration for disorders due to mutations of the mitochondrial genome, in which mutant genomes are often present in varying proportions in the cell (heteroplasmy). Alston *et al*. [19] survey this area of human genetic disease, which is of great complexity as a result not only of heteroplasmy, but of the interactions between mitochondrially-encoded proteins and the much larger number of mitochondrial proteins encoded by the nuclear genome. An in-depth review of the biology of one nuclear-encoded mitochondrial protein, PINK1, is given by Arena and Valente [20]. They not only highlight the limitations of mouse knockout models in recapitulating the autosomal recessive Parkinson's disease caused by *PINK1* germline mutations, but also emphasize the pleomorphic effects mediated by this one protein in different contexts.

We complete our eclectic survey of genetics in pathology with four reviews dealing with specific areas of development and organismal homeostasis. Crist [21] focuses on the problem of muscle regeneration and repair, which is central to a range of inherited muscular dystrophies as well as non-genetic myopathies. In addition to reviewing the genetic control of myogenic specification, he considers the challenges inherent in recapitulating and controlling the activity of muscle stem cells *in vitro*, including the prospects for targeted gene editing.

Two contributions deal with genetic insights into vascular biology; Ma and Chung [22] describe current knowledge of the genetic contributions to pulmonary hypertension, while Wetzel-Strong *et al*. [23] discuss the genetics of vascular malformations. A wide range of the latter, some inherited as Mendelian traits and others caused exclusively by somatic mutations, are caused predominantly by mutations in signalling pathways, rather than structural vascular components, providing unexpected information about the crucial role of cell-cell communication in the maintenance of normal vascular integrity.

Finally, Mitchison and Valente review the bewildering spectrum of human pathology that results from genetic abnormalities of the cilia, both motile and non-motile [24]. These “ciliopathies” can show single- or multi-system involvement, and clinical presentations as diverse as retinopathy or congenital heart disease. Disorders previously categorized as quite distinct from one another have been shown to result from mutations in the same gene; the authors thus emphasize the need to move away from disease classifications based on isolated phenotypic features, in favour of those

based on aetiology.

Summary

With some 7,000 rare inherited genetic disorders and a further vast range of genetic pathology that could have been sampled from any branch of oncology, we have inevitably had space here only to offer a few glimpses, rather than to display a clear and full perspective of our subject. Despite this, we hope that readers of *The Journal of Pathology* will find at least one article that helps to catalyse the further exploitation of genetics in Understanding Disease– which, of course, is our long-term aim.

Author contribution statement

Both authors contributed to the writing and editing of this introductory article.

References

1. Johnston M. Humans as a model organism: the time is now. *Genetics* 2014; **198:** 441.

2. Bateson, W. *The Methods and Scope of Genetics* (inaugural lecture). Cambridge [Eng.], University Press,1908

3. Peckham D, Scambler T, Savic S *et al*. The burgeoning field of innate immune-mediated disease and autoinflammation. *J Pathol* 2017; path.4812.

4. Crow YJ, Rehwinkel J. Aicardi-Goutieres syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity. *Hum Mol Genet* 2009; **18(R2):** R130-136.

5. Brown SJ. Molecular mechanisms in atopic eczema: insight gained from genetic studies. *J Pathol* 2017; path.4810.

6. Uniken Venema WT, Voskuil MD, Dijkstra G, *et al*. The genetic background of Inflammatory Bowel Disease: From correlation to causality. *J Pathol* 2017; path.4817.

7. Chiang JCH, Ellison DW. Molecular pathology of paediatric central nervous system tumours. *J Pathol* 2017; path.4813.

8. Rodrigues DN, Boysen G, Sumanasuriya S, *et al*. The molecular underpinnings of prostate cancer: impacts on management and pathology practice. *J Pathol* 2017; path.4826.

9. Mateo J, Carreira S, Sandhu S, *et al*. DNA-repair defects and olaparib in metastatic prostate cancer. *N Engl J Med* 2015; **373:** 1697-1708.

10. Graham TA, Sottoriva A. Measuring cancer evolution from the genome. *J Pathol* 2017; path.4821.

11. Stephens PJ, Greenman CD, Fu B, *et al*. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 2011; **144:** 27-40.

12. Notta F, Chan-Seng-Yue M, Lemire M, *et al*. A renewed model of pancreatic cancer evolution based on genomic rearrangement patterns. *Nature* 2016; **538:** 378-382.

13. O'Driscoll M. The pathological consequences of impaired genome integrity in humans; disorders of the DNA replication machinery. *J Pathol* 2017; path.4828.

14. Glaire MA, Brown M, Church DN, *et al*. Cancer predisposition syndromes: lessons for truly precision medicine. *J Pathol* 2017; accepted 4 Nov 16.

15. Welch HG, Passow HJ. Quantifying the benefits and harms of screening mammography. *JAMA Intern Med* 2014; **174:** 448-454.

16. Casasent A, Edgerton M, Navin NE. Genome evolution in ductal carcinoma in situ: invasion of the clones. *J Pathol* 2017; waiting for DOI.

17. Lam HC, Nijmeh JS, Henske EP. New developments in the genetics and pathogenesis of tumours in tuberous sclerosis complex. *J Pathol* 2017; path.4827.

18. Campbell IM, Shaw CA, Stankiewicz P, *et al.* Somatic mosaicism: implications for disease and transmission genetics. *Trends Genet* 2015; **31:** 382-392.

19. Alston CL, Rocha MC, Lax NZ, *et al*. The genetics and pathology of mitochondrial disease. *J Pathol* 2017; path.4809.

20. Arena G, Valente EM. PINK1 in the limelight: multiple functions of an eclectic protein in human health and disease. *J Pathol* 2017; path.4815.

21. Crist C. Emerging new tools to study and treat muscle pathologies: genetics and molecular mechanisms underlying skeletal muscle development, regeneration and disease. *J Pathol* 2017; path.4830.

22. Ma L, Chung WK. The role of genetics in pulmonary arterial hypertension. *J Pathol* 2017; path.4833.

23. Wetzel-Strong SE, Detter MR, Marchuk DA. The pathobiology of vascular malformations: insights from human (and model organism) genetics. *J Pathol* 2017; waiting for DOI.

24. Mitchison HM, Valente EM. Motile and non-motile cilia in human pathology: from function to phenotypes. *J Pathol* 2017; waiting for DOI.