An alternative approach to medical genetics based on modern evolutionary biology. Part 1: mutation and symbiogenesis

Frank P Ryan

Sheffield Primary Care Trust – CPC, Sheffield, UK

E-mail: F.Ryan@sheffield.ac.uk

DECLARATIONS

Competing interests
None declared

Funding
None

Ethical approval
Not applicable

Guarantor
FPR

Contributorship
FPR is the sole contributor

Acknowledgements
The author would like to thank Oliver Quarrell, Sheffield Children's Hospital, for his assistance with this paper

Introduction

The need to understand the genetic underpinning of common diseases, such as cancer, diabetes and autoimmune disorders, is a major challenge facing medical genetics. While many disorders involve a complex interaction between genetic traits and environmental influences, a significant proportion involve specific genetic causes that are diagnosable and thus amenable to therapeutic approaches, including family counselling, specific lines of chemically-directed therapy and even the possibility of addressing the underlying genetic causation through specific genetic therapy. The genetic approach to medicine has always found commonalities with the disciplines of evolutionary biology: only when we understand the evolutionary origins of the human genome can we aspire to understand how 'normality', in terms of genetics and epigenetics, becomes subverted in disease. Since the 1930s, mutation has been assumed to be the exclusive explanation of the hereditary basis of disease. But advances in evolutionary biology over the last two decades have introduced three additional, and important, sources of hereditary variation, namely symbiogenesis, particularly symbiogenesis involving viruses, epigenetics and hybridogenesis, all of which are now seen to have played important roles in the evolution of biodiversity. These additional mechanisms have played a major role in the evolution of the human genome – indeed they help to explain what would otherwise appear to be some paradoxical aspects of its construction and make-up. In this paper, and the four linked papers that follow it, I shall explain the evolutionary advances and show how, together with the knowledge already accrued from the study of mutation, they provide a more comprehensive understanding of the hereditary basis of human disease.

The evolutionary background

Modern evolutionary theory began, in 1859, with the publication by Darwin of his groundbreaking book, *The Origin of Species*, in which he proposed that natural selection offered a logical explanation for the diversity of life on Earth, including humanity. Darwin was well aware that natural selection depended in turn on some mechanism, or mechanisms, of hereditary variation for selection to choose from. But given the scant knowledge of his day, he was in no position to offer a scientific explanation for this. It was only with the discovery of mutation[1] – first understood as the result of errors in the copying of genes during cell division – coupled with a growing understanding of Mendelian genetics, that modern Darwinism, often referred to as 'neo-Darwinism', was born. The bringing together of natural selection, mutation and Mendelian genetics gave rise to the 'synthesis theory' of the 1930s, and this proved so useful it came to dominate evolutionary, and medical, thinking. Even today, the original synthesis theory remains highly important to evolutionary biology, but it no longer embraces all genomic potential for hereditary change. Powerful additional mechanisms for genetic, and epigenetic, change, such as symbiogenesis, hybridogenesis and epigenetics, are now seen to provide important sources of evolutionary novelty. These driving forces can be gathered together under the non-prejudicial umbrella concept of 'genomic creativity',[2] which, in tandem with Darwinian natural selection,[3] offers a more comprehensive explanation of evolution. These same driving forces also offer a more comprehensive understanding of the genetic basis of human disease.

Mutation

Traditionally mutation has been defined as any hereditable change in a gene, or more generally, in the genetic material, arising through copying errors during cell division, also extended to external causes, such as radiation or chemical insult.[4,5] This gives rise to a complex variety of random genetic change, which can affect single genes, whether translational or developmental, or bigger fractions of chromosomes, for example through deletions and insertions. When mutations affect germ cells they have hereditary, and thus evolutionary, potential. The majority of mutations will be silent or have deleterious effects, but occasionally a mutation, or more likely a cumulative series of mutation, will bring about sufficient genetic change to give rise to a new species. This linear-with-branching concept of evolution is the basis of modern Darwinism, in which mutation is considered random, meanwhile natural selection, in choosing those hereditary mutations that enhance survival, is widely promoted as the essential creative force.

Mutation is, of course, a major cause of inborn errors of metabolism and it is a contributory factor to a great diversity of diseases.[6] At its simplest, a copying error in the DNA sequence of a gene during cell division may alter the amino acid sequence of the protein coded by the gene, with loss of function of the protein. Phenylketonuria (an autosomal recessive condition) is associated with more than 400 different mutations all of which damage the enzyme phenylalanine hydroxylase. Such single-gene mutations will usually be inherited along Mendelian lines, such as the dominantly-inherited achondroplasia and Huntington's disease, the recessively inherited cystic fibrosis and the sex-linked disorders, which can themselves be dominant, and thus affect either sex (e.g. Vitamin-D resistant rickets), or recessive, when they almost exclusively affect the male sex (e.g. haemophilia and Duchenne muscular dystrophy).

To date geneticists have linked more than 5000 single-gene disorders in humans to causative mutations. Other mutations can change the number of chromosomes, as in Down's syndrome, or delete, duplicate, fragment or otherwise damage the structure of chromosomes, giving rise to a variety of syndromes. Mutation is also a common feature of cancers, which usually arise in fully developed tissues long after embryogenesis. Other chromosomal abnormalities affect the germ cells, where they give rise to a wide range of disorders, including aberrant embryological development, with resultant congenital abnormality, as well as a great many inborn errors of metabolism. In all such cases, a clear understanding of the genetic cause, or causes, is the basis for medical prevention and therapy.

It is beyond the remit of this paper to describe the many different types of mutation that play an important part in both evolution and human disease, or to explain the methods currently employed to locate and identify a mutated gene – I would refer interested readers to the standard texts for a fuller explanation.[4,7] The medical approach to mutation includes genetic counselling, for example enabling couples at risk of particular disorders to have essential information so they can make their own decisions on matters of reproduction, and public education about the risks of increasing maternal age, avoidance of risk factors such as radiation of the germ cells and fetus, caution with respect to drug and chemical exposure, such as thalidomide, and vaccination against rubella. Newer measures, such as *in vitro* fertilization and genetic screening of the fetus at the 16- or 32-cell stages, can be offered to high-risk families. Such preimplantation genetic diagnosis (PGD) may be efficacious in a variety of diseases, including sex-linked disorders, single gene defects and chromosomal disorders. The potentially amenable sex-linked disorders include haemophilia, fragile X syndrome, most of the neuromuscular disorders and hundreds of other diseases, including the sex-linked dominant disorders, Rett syndrome and incontinentia pigmenti. The potentially amenable single gene defects include the cystic fibrosis (at present limited to the most common causative mutations), Tay-Sachs disease, sickle-cell anaemia and Huntington's disease. The potentially amenable chromosomal disorders include a wide variety of translocations, inversions and deletions. Amenability requires a genetic abnormality that is predictable, and the availability of a suitable screening test in isolated embryological cells. In some of these cases PGD not only removes the risk of an affected offspring but also eliminates the risk in future generations.

One in every 2500 babies born to Caucasian parents suffers from cystic fibrosis, making it one

of the commonest of hereditary diseases. It is caused by a variety of mutations affecting a regulator gene, which is known as the cystic fibrosis transmembrane regulator gene, or CFTR, located in the region q31-32 of human chromosome 7, and which codes for an ion channel involved in transport across membranes.[8] Cystic fibrosis is perhaps the most familiar example of an autosomal recessive condition. There are many such recessive genetic disorders that might potentially be cured by the addition of a single 'normal' gene, and these conditions, including cystic fibrosis, are the subject of intensive current investigation aimed at 'gene therapy'.

Cancer is another arena in which mutated genes are known to play an important role. But here the genetic abnormalities are more complex than in the inherited diseases and very often involve multiple mutations as well as important links to environmental factors. At genetic level, cancer involves a series of steps that include multiple mutations that deregulate signalling pathways. New lines of research suggest that these mutations must cooperate with each other for the cancer to develop, so that research aimed at determining the precise nature of the cooperating mutations and the signalling pathways they affect is a major challenge.[9] Other research focuses on three groups of genes that are frequently mutated in cancer: oncogenes, where mutations promote inappropriate cell proliferation; tumour-suppressor genes, where mutations remove the normal suppression of inappropriate cell proliferation; and mutator genes, which normally control the repair of damaged genetic regions, and where mutation impairs this reparative function. The decoding of the human genome has highlighted the genetic alterations that underlie cancers in such unprecedented detail that it has led two American oncologists, Vogelstein and Kinzler, to declare that 'cancer is, in essence, a genetic disease'.[10] The same authorities have summarized the mutated genes responsible for various cancers, together with the ways in which these mutations have subverted the normal genetic mechanisms to induce tumours. For example 15–20% of women with breast cancer have a family history of the condition and 5% of all breast cancers have been linked to mutations in the genes BRCA1 and BRCA2.[11] Geneticists can further predict that women who carry these mutations have an 80% risk of developing breast cancer during their lifetime, so that there are various options that help to reduce the risk, including prophylactic oophorectomy, regular breast screening and the potential of pre-emptive surgery.

In 2006, a systematic multicentre American study pioneered the screening of more than 13,000 genes taken from human breast and colon cancer cells.[12] Given the 'normal' human genome, they were in a position to compare the genes they found in the two cancers with the normal, revealing that individual tumours accumulate an average of 90 mutant genes. It seems that a much smaller number of these actually play a part in the cancer process, in their estimation perhaps 11 mutations for each of breast and colon cancer. Encouraged by these findings, the US National Institutes of Health is drawing up an atlas of cancer genomes – the Cancer Genome Atlas Project, or TCGA.[13] The aim is to decode the genomes of every human cancer and, by comparing these to the normal, extrapolate the genetic abnormalities that underlie all cancers. A pilot study has begun with cancers of the lung, brain and ovary.

As the foundation of knowledge grows, it is inevitable that preventive and therapeutic aspects will be extrapolated from it. However, a caveat needs to be drawn on the use of the term 'mutation' in such laudable enterprises since mutation should not be extrapolated as the exclusive genetic explanation of disease, including cancer.

Symbiogenesis

Symbiosis was first defined by de Bary in 1878 as 'the living together of differently-named organisms' – in modern terms a significant interaction between different life-forms.[14] The definition implied that symbiosis was a force in evolution, and this was subsequently defined as 'symbiogenesis' by Merezhkovskii in 1910.[15] It is a common mistake to equate symbiosis with mutualism, but only one of the partners needs to benefit for the association to be regarded as a symbiosis. In symbiological terminology, the partners are called 'symbionts', and the partnership of two or more symbionts is known as a 'holobiont'. From the very beginning, de Bary's definition included parasitism and commensalism as well as mutualism. Mutualistic symbiosis, which implies that two or more symbionts benefit, often begins with parasitism and it can be difficult to determine where, in the evolutionary trajectory, parasitism ends and mutualism begins. Even in the most virulent forms of parasitism, such as pandemic

plagues, the interaction culls the host species genotype, and particularly with persistent plague viruses, may lead to co-evolution of parasite and host in an important evolutionary progression I have labelled 'aggressive symbiosis'.[16] It appears likely that the current large-scale population crash among koalas in Australia is a classical example of aggressive symbiosis, with plague culling already progressing to 'endogenization' of the lethal virus, which will result in a permanent symbiotic union of host and viral genomes.[17,18] From this perspective, AIDS is an evolutionary phenomenon, with the present pandemic the latest in a long series of similar retroviral pandemics that have played a brutal, yet paradoxically creative, role in the evolution of vertebrates, particularly mammals, and notably humans.

Symbiosis operates at different levels. We are familiar with the 'behavioural symbioses' of the oceanic cleaner stations, where small fish and crustaceans remove parasites and detritus from the skins and mouths of predators. These behavioural patterns must be hardwired, through genetic or epigenetic inheritance, in the respective genomes. In metabolic symbioses there is an exchange of metabolic products between the symbionts, for example in the mycorrhizae that link fungi with the roots of 97% of all plants, or in lichens, which are intimate partnerships of fungi and photosynthetic algae. Other metabolic symbioses include the gut bacteria that play an important role in human metabolism.[19] Symbioses may also involve the union of partners at a genetic level, usually involving a microbe and host, where they enable the transfer of a gene, or a cluster of genes, from microbe to host. This differs from mutational changes in genes since it involves the movement of pre-evolved genes, or genetic pathways, from one evolutionary lineage into another. A good example is the symbiosis between rhizobial bacteria and legumes, where the bacteria donate genes essential for nitrogen fixation to the plant. In fact the genetic transfer, essential to the nitrogen cycle, involves a second layer of symbiosis, with a phage virus donating an integrase gene that enables the transfer of a 'symbiotic genetic island' from a symbiotic to a non-symbiotic rhizobium in the field, thus enabling the rhizobial-legume symbiosis.[20]

While it does not contradict or disagree with Darwinian natural selection, symbiosis does complicate the level at which selection operates. In mutualistic symbiosis selection will operate, to a significant degree, at the level of the partnership.[21,22] At its most powerful level, genetic symbiosis involves the union of entire genomes, usually very different genomes arising from disparate kingdoms, to create a novel 'holobiontic genome', with major evolutionary potential. Such holobiontic genomic fusions have played an important role in the evolution of life and its subsequent diversification. For example, the mitochondria that enable us to breathe oxygen are descended from the genomic union of a single-celled eukaryotic ancestor (formerly termed a 'protozoan' but nowadays termed a 'protist') and an oxygen-breathing bacterium.[23] This major symbiogenetic event only happened once, so it is part of the evolutionary inheritance of all animals, plants, fungi and the oxygen-breathing protists. Our human mitochondria still retain some of their bacterial genes contained within a typical bacterial-style ring genome and, remarkably, sequence analysis suggests that perhaps the closest known modern relative to the ancestor of all mitochondria is *Rickettsia prowazeki*, the cause of epidemic typhus.[24,25]

Diseases involving mitochondria

The evolutionary origin of human mitochondria, from an archaic free-living oxygen-breathing bacterium, has considerable implications for the genetics of mitochondrial disease. At the time of first symbiotic union, more than 1 billion years ago, the ancestral bacterium would have possessed some 2000 or more genes. Today, as a result of selection working at the level of holobiontic union, the genome of each human mitochondrium has been whittled down to a residuum of 37 genes, encoding 22 transfer RNAs, two types of ribosomal RNA and 13 proteins involved in cellular oxidative phosphorylation.[7] Meanwhile approximately 300 of the original bacterial genes have been transferred to the nucleus, where many continue to play a part in the nucleus-mitochondrial genetic linkage that is necessary for normal function.[26] The mitochondria are inherited from the ovum as part of the 'cytoplasmic inheritance' of the cell and they reproduce by bacterial-style budding independently of the mitotic reproduction of the nucleus. Thus, where the nuclear genome is biparental in origin and follows Mendelian laws of inheritance, the mitochondrial genome is exclusively maternal and follows a non-Mendelian pattern of inheritance.

Mitochondria fulfil multiple cellular functions, including energy production, generation of reactive oxygen species (toxic free radicals that are by-products of respiration) and the regulation of programmed cell death, or apoptosis, so that genetic abnormalities can lead to many different patterns of disease.[27] The bacterial inheritance also leads to significant differences when it comes to mitochondrial genetics. Since the mitochondrial genome is much smaller than the nuclear genome, we might anticipate fewer mutations and thus a low prevalence of genetically-induced disease. However, where most of our nuclear DNA is non-coding, so that mutations are less likely to cause pathogenesis, most mitochondrial DNA is coding and thus mutations are more likely to cause disease. Moreover, bacterial genes are more error-prone than vertebrate genes, so that mutations in mitochondrial genes are about 10 to 20 times more common than would be expected in nuclear genes. This is further complicated by the fact that mitochondrial disease can also result from mutations affecting the nuclear-based genes, added to which there are thousands of mitochondria within each cell, so that a mutation in a single mitochondrion may give rise to an abnormal subpopulation, so the individual ends up with a mixed population of normal and mutants in the same tissues. Thus it would be very difficult to understand the genetics of mitochondrial diseases without taking into account the symbiogenetic origins of mitochondria.

Mitochondrial diseases tend to be highly specific to the individual, or family, ranging in severity from mild to fatal, with the resulting manifestations depending on the organ, or organs most affected, and their oxygen requirements. It is hardly surprising that the complexity of the underlying genetics coupled with the patient-specific variation in disease presentation, can make the genetic basis of such diseases hard to diagnose and trace. Oxidative phosphorylation (OXPHOS) within the mitochondria involves enzyme-enabled electron transport along four chain reactions (complexes I-IV), meanwhile a fifth enzymic step (complex V) produces ATP. Roughly 1 in 7600 births are affected by genetic abnormalities affecting the OXPHOS pathways, both mitochondrial-based and nuclear-based, contributing a significant proportion of inborn errors of metabolism. Pathogenic mutations, leading to significant disease, have now been identified in more than 30 of the 37 mitochondrial genes and in more than 30 of the related nuclear genes.[28] 'Complex I deficiency' is the commonest within-mitochondria category, accounting for roughly one-third of 'respiratory chain deficiencies'. Often presenting at birth or in early childhood, affected individuals suffer a progressive neuro-degenerative disorder, accompanied by a variety of symptoms in organs and tissues that require high energy levels, such as brain, heart, liver and skeletal muscle. Clinical categories presenting in infancy include fatal infantile lactic acidosis, Leigh's syndrome, cardiomyopathy with cataracts, hepatomegaly with renal tubulopathy, cataracts and developmental delay, lactic acidaemia in the neonatal period followed by mild symptoms.[29,30]

Another complex-1 syndrome, presenting in adult life, is Leber's hereditary optic neuropathy, which is one of the commonest inherited forms of eye disease. This presents with bilateral, painless, subacute visual failure that develops during young adult life, leading to central loss of vision and atrophy of the optic nerves. Men are four times more likely to be affected than women. Most cases of Leber's syndrome are the result of one of three point mutations in mitochondrial genes, and while it is sometimes confused with X-linked genetic inheritance, it is inherited through non-Mendelian mitochondrial inheritance.[31] Other mitochondrial disorders include Leigh syndrome mentioned above, a subacute sclerosing encephalopathy that affects about 1 in 40,000 neonates, and NARP, a syndrome of neurogenic muscular weakness, ataxia and retinitis pigmentosa.[32] OXPHOS disorders are said to be the most frequent cause of metabolic abnormality in paediatric neurology, although they often present with non-neurological symptoms such as failure to thrive, or with hepatic, cardiac, renal, gastroenterological, endocrine or haematological symptoms, and geneticists have identified a wide spectrum of such diseases variously associated with abnormalities of mitochondrial genes, or the nuclear genes associated with mitochondrial function.[26,27] An interesting example of mitochondrial dysfunction arising from mutation in nuclear genes is Friedreich's ataxia, where the affected nuclear gene, known as FXN, codes for the mitochondrial protein, frataxin. Frataxin enables the removal of iron from the mitochondria, so that when it is defective or absent the mitochondria accumulate excess iron and

suffer free radical damage, with resulting OXPHOS dysfunction.

There is growing evidence that mitochondrial dysfunction may play a significant role in a much broader spectrum of diseases, including diabetes mellitus, cancers, cardiovascular disease, lactic acidosis, osteoporosis, Alzheimer's disease, Parkinson's disease, stroke and even the ageing process. Given the rapid advance in genetics, we may in time have an effective gene-based therapy for conditions such as Leber's hereditary optic neuropathy and the many OXPHOS dysfunctions, but it is clear that any such therapeutic approach will be obliged to consider the symbiotic evolutionary origins of mitochondria and the complex genetic and molecular dynamics that arise from such an inheritance.

Only recently have we realized that symbiotic viruses have also played an important role in the evolution of the human nuclear genome, with considerable implications for health and disease. This will be the subject of the next paper in this series.

References

1 De Vries H. *Species and Varieties, Their Origin by Mutation*. Chicago, IL: Open Court; 1906
2 Ryan FP. Genomic creativity and natural selection: a modern synthesis. *Biol J Linn Soc* 2006;**88**:655–72
3 Feder ME. Evolvability of physiological and biochemical traits: evolutionary mechanisms including and beyond single-nucleotide mutation. *J Experimental Biology* 2007;**210**:1653–60
4 Hartl D, Jones E. *Genetics: Analysis of Genes and Genomes*. 5th edn. Sudbury, MA: Jones and Bartlett; 2003
5 See http://en.wikipedia.org/wiki/Mutation
6 See http://en.wikipedia.org/wiki/Inborn_error_of_metabolism
7 See http://www.biochemistry.org/education/pdfs/basc/10/BASC10_full.pdf
8 See http://en.wikipedia.org/wiki/Cystic_fibrosis_transmembrane_conductance_regulator
9 Du Y, Spence SE, Jenkins NA, Copeland NG. Cooperating cancer gene identification through oncogenic-retrovirus-induced insertional mutagenesis. *Blood* 2005;**106**:2498–505
10 Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med* 2004;**10**:789–99
11 Caldas C, Edwards PA. A quantum leap in our knowledge of breast cancer mutations. *Breast Canc Res* 2006;**8**:304–6
12 Sjöblom T, Jones S, Wood LD, *et al* The consensus coding sequences of human breast and colorectal cancers. *Science* 2006;**314**:268–74
13 See http://cancergenome.nih.gov/
14 Sapp J. *Evolution by Association: A History of Symbiosis*. Oxford: Oxford University Press; 1994
15 Merezhkovskii C. La plante considérée comme un complexe symbiotique. *Bulletin de la Société Naturelles* 1910;**6**:17–98
16 Ryan F. *Virus X*. New York, NY: Little Brown; 1996
17 Tarlinton RE, Meers J, Young PR. Retroviral invasion of the koala genome. *Nature* 2006;**442**:79–81
18 Stoye JP. Koala retrovirus: a genome invasion in real time. *Genome Biology* 2006;**7**:241–3
19 Ryan F. *Darwin's Blind Spot*. Boston, MA: Houghton Mifflin; 2002
20 Sullivan J, Ronson CW. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene. *PNAS* 1998;**95**:5145–9
21 Smith JM. A Darwinian view of symbiosis. In: Margulis L, Fester R, eds. *Symbiosis as a Source of Evolutionary Innovation*. Cambridge, MA: MIT Press; 1991
22 Smith JM, Szathmary E. *The Origins of Life*. Oxford: Oxford University Press; 1999
23 Margulis L. *Origin of Eukaryotic Cells*. New Haven, CT: Yale University Press; 1970
24 Andersson GE, Zomorodipour A, Andersson JO, *et al* The genome sequence of *Rickettsia prowazeki* and the origin of mitochondria. *Nature* 1998;**396**:133–40
25 Gray MW. Rickettsia, typhus and the mitochondrial connection. *Nature* 1998;**396**:109–10
26 Timmis JN, Ayliffe MA, Huang CY, *et al* Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature Rev Genetics* 2004;**5**:123–35
27 Wallace DC. Mitochondrial diseases in man and mouse. *Science* 1999;**283**:1482–8
28 Thorburn DR. Mitochondrial disorders: prevalence, myths and advances. *J Inherit Metabolic Disease* 2004;**27**:349–62
29 Pitkänen S, Robinson BH. Mitochondrial complex I deficiency leads to increased production of superoxide radicals and induction of superoxide dismutase. *J Clin Invest* 1996;**98**:345–51
30 See http://www.umdf.org/atf/cf/%7B28038C4C-02EE-4AD0-9DB5-D23E9D9F4D45%7D/COMPLEX_1_DEFICIENCY.pdf
31 See http://www.ncbi.nlm.nih.gov/bookshelf/picrender.fcgi?book=gene&&partid=1174&blobtype=pdf
32 See http://www.ncbi.nlm.nih.gov/bookshelf/picrender.fcgi?book=gene&&partid=1173&blobtype=pdf