MOLECULAR ECOLOGY

Molecular Ecology (2013) **22**, 3437–3443

NEWS AND VIEWS

OPINION

What is a genome?

A. STENCEL* and B. CRESPI†

Institute of Environmental Sciences, Jagiellonian University, Gronostajowa 7, 30-387 Kraków, Poland; †Department of Biological Sciences, Simon Fraser University, Burnaby, British Columbia, V5A 1S6 Canada

The field of genomics is expanding rapidly, yet the meanings of the word 'genome' have yet to be conceptualized in explicit, coherent and useful frameworks. We develop and apply an evolutionary conceptualization of the genome, which represents a logical extension of the evolutionary definition of a gene developed by George C. Williams. An evolutionary genome thus represents a set of genetic material, in a lineage, that due to common interests tends to favour the same or similar phenotypes. This conceptualization provides novel perspectives on genome functions, boundaries and evolution, which should help to guide theoretical and empirical genomics research.

Keywords: definition, evolution, gene, genome

Received 24 February 2013; revised 8 April 2013; accepted 16 April 2013

The ongoing explosion of genomic data in studies of molecular ecology, evolution and disease is transforming the biological sciences. This progress has been driven by advances in technology – have conceptual frameworks kept pace? This question is important because how we think about – conceptualize – a gene or genome will often structure what questions we imagine and address, what data we collect, how we analyse our data and what conclusions we draw.

In this article, we draw attention to fundamental ambiguity and lack of consensus concerning the meaning of the term 'genome' and suggest an evolutionary conceptualization of this idea that parallels the evolutionary idea of a 'gene' developed by George C. Williams. We then illustrate the usefulness of these frameworks for theory and empirical research.

The term 'genome' was developed, apparently via merging of 'gene' with 'chromosome', in 1920 by the German botanist Hans Winkler (Lederberg & McCray 2001). He referred to a haploid set of chromosomes as the 'material basis' of the species. Currently, the term is defined, for example, as 'the entire genetic complement of a living organism' (Brown 2007), all DNA in the chromosomes of a eukaryote, all DNA in the chromosome of prokaryotes, or all DNA or RNA in a virus.

These perspectives on the genome can be considered incomplete and ambiguous for two main reasons. First, do we include nonchromosomal genetic material (such as cytoplasmic elements that harbour their own DNA) or elements in chromosomal DNA that are independently mobile (e. g., Zilber-Rosenberg & Rosenberg 2008)? Yes, but only if they are vertically inherited? But some elements, such as *Wolbachia*, transposable elements, viruses, supernumary chromosomes and plasmids, can be inherited horizontally as well as vertically (Table 1). Some such elements are parasitic, some mutualistic, some shift between the two, some can integrate into chromosomes and some develop degrees of mutual interdependencies with other DNA inhabiting the same cell. Some viruses, transposons and other elements have integrated into the DNA of former 'hosts' to such a degree that their nonhost origins are no longer, or are barely, recognizable. Chromosomes, parts of chromosomes and genes introgress between species and divergent populations. For a large suite of genetic elements and situations, the boundaries of 'genomes' are unclear (Table 1). Under what circumstances, if any, should any of these elements be usefully or logically considered to belong to a particular genome?

Second, is the genome a property of a cell, an individual, a population, a species or all of these? In GenBank or EMBL, we find 'the genome' for each of many species. For individualized medicine, we refer to our own, personal genome – or the aggregate of DNA from our *Homo sapiens* legacy plus our diverse intra- and extracellular microbiota, some of which we inherit vertically, some horizontally, some both. In cancer biology, we discuss the genome of a tumour, a cancer cell or a cancer cell lineage. All of these hierarchical conceptualizations of genome are meaningful, and potentially misleading, in their own ways. Can they be unified into a coherent framework?

The evolutionary gene

Let us step back and first discuss ideas concerning how the word 'gene' can be usefully considered in the contexts of biology in general and evolutionary biology in particular. A paper from members of the ENCODE team nicely reviews the history of the term 'gene' and describes how recent advances in molecular genetics, especially with regard to noncoding DNA, expression regulation, transcriptional diversity, and alternative splicing, have necessitated reconsideration of its working definition (Gerstein *et al.* 2007). These authors suggest that:

Correspondence: B. Crespi, Fax: 778 782 3496; E-mail: crespi@sfu.ca

© 2013 John Wiley & Sons Ltd

Table 1 Examples of genetic elements and situations for which the boundaries of the 'traditional' genome, conceptualized in terms of vertically inherited chromosomal material, become indistinct. Note also that the distinction between cytoplasmic elements and vertically inherited microbes is not absolute

Horizontal gene transfer	References
Horizontal transfers of genetic material between 'species' of bacteria	Didelot *et al.* (2009); Wiedenbeck & Cohan (2011)
Hybridization between species or genetically divergent populations of eukaryotes	Abbott *et al.* (2013); Shurtliff (2013)
Horizontal gene transfer into metazoans	Gladyshev *et al.* (2008); Boschetti *et al.* (2012); Yue *et al.* (2012)
Lateral transfer of DNA involving cancer cells	Bergsmedh *et al.* (2001); Goldenberg (2012); Trejo-Becerril *et al.* (2012)
Horizontal and vertical inheritance of elements	**References**
Vertical and horizontal inheritance and movement of *cytoplasmic elements and supernumary chromosomes*	Ballard & Whitlock (2004); Van Vugt *et al.* (2009); Boratynski *et al.* (2011); Curtis *et al.* (2012); Martis *et al.* (2012)
Vertical and horizontal inheritance and movement of *viruses*; gradual integration of viruses into host genomes	Holmes (2011); Doerfler (2012); Dupressoir *et al.* (2012)
Vertical and horizontal inheritance and movement of *transposable elements* with deleterious and beneficial effects on co-inherited genetic material; integration into host genomes	Venner *et al.* (2009); de Carvalho & Loreto (2012); Rebollo *et al.* (2012)
Vertical and horizontal inheritance and movement of *microbes* with, within and across hosts	Lederberg (2000); Ley *et al.* (2006); Moran *et al.* (2008); Zilber-Rosenberg & Rosenberg (2008); Koga *et al.* (2012)

> 'The gene is a union of genomic sequences encoding a coherent set of potentially overlapping functional products'.

This definition pivots on the phrase 'functional products' in a biochemical sense, in terms of genotype-to-phenotype mappings. It is intended to be backward compatible with previous and current definitions, including the myriad named genes in databases. Most importantly, it is 'fuzzy', yet empirical. The conceptualization thus tells us how to study DNA and its products in this era of genomic data – by focusing on functionality of gene products, whatever they may be. This view should change how we think about a 'gene' and how we analyse the processes of genes making biochemical phenotypes.

But the post-ENCODE perspective on gene is built for molecular biologists who usually focus on mechanisms – not for evolutionary biologists, who usually focus on the processes that generate and change genes over time. Williams (1966, pp. 22–25; 1992, pp. 10–13), with contributions from Dawkins (1976, p. 28; 1982a, pp. 83–89; 1982b), developed the first explicitly evolutionary conceptualization of the gene, which we rephrase as:

> 'A region of DNA or RNA that persists for long enough to have its frequency adjusted or maintained by natural selection'.

Strictly speaking, Williams (1992) emphasized the '*information*' in the DNA, not its material form, but this distinction need not directly affect our deployment.

This conceptualization of 'gene' from Williams and Dawkins shares notable properties with the post-ENCODE view described just above it. The evolutionary gene is fuzzy with regard to particular contiguous DNA or RNA sequences. As such, it resembles a haplotype, an indefinite length of DNA that persists for some indefinitely long period of time, at some frequency, as it is subject to mutation, recombination and natural selection. Mutation and recombination thus alter, fracture and rejoin evolutionary genes, purifying selection removes deleterious variants, and positive selection leads to relatively new variants increasing in frequency. Evolutionary genes may also usefully be viewed as particular alleles, of highly variable size, in some population. Thus, an evolutionary gene may be a small portion of a nominal gene (i.e. a named gene in GenBank), an open reading frame, a megabase-scale section of DNA that does not recombine, a whole chromosome, or in some asexuals, all chromosomal DNA. Whatever-length piece of genetic material sails successfully across multiple generations, as a long-term replicator on the radar of natural selection, may be an evolutionary gene. This conceptualization is based on 'selected effect' (natural selection based) definitions of functionality, in comparison with 'causal role' functionality, which is based on experimental or inferential assignment of molecular biological cause and effect (Doolittle 2013; Graur *et al.* 2013).

The evolutionary gene is thus, crucially, conceptualized in terms of the only process that directly produces and maintains function: natural selection. Our molecular biological gene and evolutionary gene agree on this most

© 2013 John Wiley & Sons Ltd

important dimension, provided that evidence for function comes from studies of purifying or positive selection. Research programmes may concentrate now on the officially designated genes in databases, but natural selection and evolution only see them, as entities, to the extent that they correspond to the evolutionary view.

The evolutionary genome

Armed with meaningful, useful and dovetailing conceptualizations of 'gene' from molecular biology and evolution, how should we define 'genome'? Put differently, a gene 'does' function – what does a genome do?

We suggest that a genome cooperatively programmes the production of an organism – an individual. This individual – a constellation of interacting phenotypes – survives and reproduces as a unit and functions, at least very substantially, as an integrated entity, in the same general manner that a gene functions in contributing to a biochemical or other phenotype. The genes in this conceptual genome cooperate because they share common evolutionary interests – they are in the same genomic boat, as it were. Such cooperation stems in part from vertical inheritance but, most fundamentally, from the shared, genomically predominant evolutionary genetic interests of autosomal genes (in eukaryotes) or linked genes (in prokaryotes). We thus suggest that an evolutionary genome may be conceptualized as:

> 'A set of genetic material, in a lineage, that due to a sufficient degree of vertical inheritance and other factors favouring cooperation among genes, tends to favour a common or broadly overlapping suite of phenotypes'.

Said another way, whereas a gene represents a functional 'replicator', a genome codes for producing a functional 'vehicle' of selection. To do so, its elements necessarily persist together for long enough, across generations, to have the frequencies and forms of the phenotypes that they cooperatively produce altered or maintained by selection in a coherent and adaptive way. The evolutionary genome can thus be considered as a logical extension of the evolutionary gene. It also resembles Dobzhansky's (1970) idea of the 'coadapted gene complex', extended to any set of genes that interacts. And as for the gene of Williams (1992), it is actually the information of the genome, rather than its material embodiment, that is important from functional and evolutionary perspectives.

The idea of the evolutionary genome is predicated on the distinction between the physical presence of genetic material (and its information content, if any) and functional presence in terms of cooperative contribution to common phenotypes. Evolutionary 'genes' and 'genomes' both change with time, in that their boundaries, composition and functional coherence are expected to alter under the influences of natural selection and other evolutionary genetic processes. The evolutionary gene thus changes in form under mutation and recombination and in frequency under selection. By contrast, evolutionary genomes change mainly in response to the three forces arrayed in Fig. 1:

1 Selection for joint, common interest of autosomal genes, the majority faction in eukaryotic genomes, all of which usually reproduce equally via fair meiosis. Such selection causes the genome to functionally cohere and cooperate, in producing a more or less optimal phenotye and in suppressing elements that conflict with autosomal gene interests (Frank 2003; Leigh 2010), as described in more detail below.
2 Genomic incompatibilities caused by hybridization or relatively strong outbreeding, whereby novel alleles from

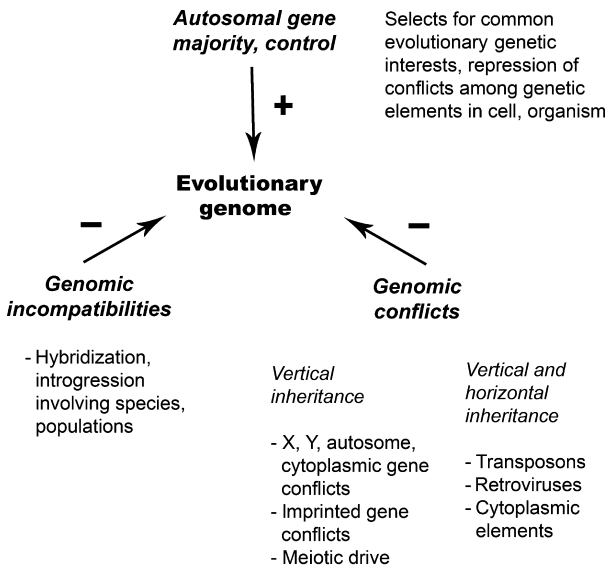

Fig. 1 The evolutionary genome codes for producing a functional 'vehicle' of selection – a suite of favourable phenotypes – to the extent that its different genetic elements cooperate. In eukaryotes, the shared reproductive interests, and usual strong genetic majority, of autosomes, select for cooperation among autosomal genes and repression of conflicts with non-autosomes (Frank 2003; Leigh 2010). Genetic integrity, and programming of functional phenotypes, may be reduced by two processes: (i) hybridization or relatively extreme outbreeding, which introduce novel alleles and genes, and (ii) genomic conflicts, which involve introduction of elements with effects that tend to be phenotypically deleterious (by horizontal transmission), or different phenotypic optima between elements that are inherited vertically. Higher levels of outbreeding and hybridization may contribute to both processes, given that they favour the spread of selfish elements (Hurst & Werren 2001) as well as leading to disruption of local genetically based adaptation. Among prokaryotes, the shared reproductive interests of linked, chromosomal DNA, co-inherited over many generations, is expected to favour cooperation among genetic elements in the cell. Extracellular genetic material (not shown on Figure), such as components of the microbiome, may influence levels of phenotypic adaptation either positively or negatively.

© 2013 John Wiley & Sons Ltd

divergent genomes are introduced. Such gene flow degrades the functional coherence of evolutionary genomes, but should be followed by selection against its phenotypic effects and introgressed alleles, with exceptions for alleles that confer benefits.

3 Genomic conflicts that involve elements with transmission that is either vertical, or vertical and horizontal. Elements with both vertical and horizontal inheritance, exerting selective impacts, are generally expected to reduce the cooperative, functional integration of genomes, given that they may function at least in part as genomic parasites with their own reproductive interests. As for alleles introgressed by hybridization, however, some may be co-opted by their 'host' and evolve individual-beneficial functions (e. g., Rebollo *et al.* 2012), thus joining our evolutionary genome.

What about conflicts involving elements that undergo only vertical transmission, such as meiotic drive, imprinted genes or sex chromosomes? Such conflicts exist for two main reasons, both of which reinforce and sharpen the framework described here. First, vertical inheritance need not always be 'fair' to alternative alleles in a genome, as meiosis can provide opportunities for one allele to increase its transmission relative to alleles and unlinked genes. Meiotic drive alleles remain in the genome, but cheat by getting themselves more 'room' in the boat. Similar considerations apply to elements that bias offspring sex allocation.

Second, alleles that mediate genomic conflict may programme the generation of phenotypes to facilitate their reproduction not just by favouring their increased vertical inheritance via more descendants, but also through positive effects on collateral kin, or other individuals who are relatively likely to harbour allele copies identical by descent. Genes with partially divergent interests, such as paternal versus maternal imprinted genes, or sex chromosomes versus autosomes, thus exhibit some mixture of cooperation and conflict with regard to other genetic elements. Such conflicts are expected to result in push and tug of quantitative phenotypes, representing attempts to steer the boat somewhat more in one direction than another, which may lead to deviations from autosomal interests.

For both manifestations of genomic conflict, close linkage will tend to unite the interests of sets of elements in conflict with others, but the effects of conflict elements are expected to be strongly countered by the predominant power of selection for favourable, phenotype-level and organism-level outcomes, and joint vertical inheritance. Such agreement stems from the clear common interests, derived from the fairness of meiosis, and the usual much larger numbers of autosomal genes than other factions in the genome (Leigh 1991, 2010; Frank 2003; Haig 2006). Across evolutionary time, intragenomic conflicts will thus be subject to continual selection for suppression, by genes in the autosomal majority, of organism-level deleterious effects – they may flare up periodically followed by repression or engage in longer-term tugs of war (Frank & Crespi 2011), but their weaker ability than autosomal genes to influence phenotypes suggests that their influences on phenotype-scale adaptation will be sharply countered and reduced by other forces. Such conflict elements may thus tend to favour somewhat different phenotypes than the autosomal majority, or one another, but vertical transmission constrains them to co-inheritance. They therefore belong in an evolutionary genome as conceptualized above, but are under selection to somewhat disagree, leading to less efficient and adaptive programming of complex phenotypes to some as yet unknown degree.

How does our conceptualization, and its dynamics as shown in Fig. 1, obviate the problems with traditional, less explicit definitions? Like the evolutionary gene, the evolutionary genome is fuzzy and quantitative, but simply because Mother Nature has made it so. At one extreme, all genes in a lineage always favour the same phenotypes – their interests coincide fully, due to vertical inheritance, fair meiosis, identical relatedness to offspring and other inclusive fitness mediating interactants, and any other factors favouring shared common interest. Such genes cooperate fully with one another, such that gene-level functions extend to genome-level, phenotype-scale functionality generated and maintained by selection. At the other extreme, some DNA in cells may be parasitic and inherited only horizontally, such as most viruses or pathogenic bacteria. This genetic material is under selection to modify 'host' phenotypes exclusively for their own benefit. In the region between these two extremes, we have a suite of genetic elements that can be inherited vertically as well as horizontally, each to some degree with some frequency (Table 1). To the extent that their interests in producing particular phenotypes tend to coincide with those of the other genetic material in the lineage – they will be, relatively more, elements of this evolutionary genome.

We use the term lineage because it may refer to any level in the hierarchy of life, from cells themselves (including evolving somatic and cancer cell lineages) to populations or species in phylogenies. This genome is, moreover, agnostic to the physical locations of genes in an organism: for example, if some lineage of prokaryotes in our guts was only vertically inherited from human mother to offspring, it deserves to belong – functionally – to the human genome, in the same way that formerly viral or transposing DNA may be purely vertically inherited, and, in this case, eventually become unrecognizable as to origins. Most generally, prokaryotes, viruses, transposons, supernumary chromosomes and other elements belong to our evolutionary genome – an extended, coadapted genotype – to the extent that they share interests in our survival and reproduction.

In most cases, the evolutionary genome will correspond closely to the genetic material inherited through gametes – the traditional conceptualization as developed by Hans Winkler. In some situations, however, as shown in Table 1, this original idea becomes insufficient to capture and reflect the causal processes that make a set of co-inherited

© 2013 John Wiley & Sons Ltd

genes function together – and may indeed become misleading. Consider several specific cases. In both *Homo* and *Macaca*, immune system genes appear to have directionally introgressed across species (Abi-Rached *et al.* 2011; Trask *et al.* 2013). If such alleles are subject to positive or purifying selection in their new genetic environments, and exert positive effects on some component of fitness, then they can be considered as components of their adopted evolutionary genome; however, some immune system region loci, such as t-alleles in mice (e. g., Ben-Shlomo *et al.* 2007), exert negative organism-level fitness effects, and transfer of such elements thus represents parasitism and a reduction in genome-level function. More broadly, hybridization commonly results in reduced hybrid fitness, but also commonly leads, across longer timescales, to positive effects via transfer of beneficial genes, indicating that it both reduces and enhances phenotypic adaptation, presumably via genome-scale interactions. A notable proportion of transposable elements have apparently evolved host-beneficial functions (Rebollo *et al.* 2012), transitioned from mobile parasites to stable mutualists and met criteria for inclusion by the evolutionary genome concept. Prokaryotes, such as some *E. coli* and species of *Campylobacter*, are subject to high levels of gene influx and loss, such that their evolutionary genomes are highly labile, but can be recognized as such through studies of convergent functions in transferred DNA across lineages (Welch *et al.* 2002) and studies of genome composition in relation to ecology (Sheppard *et al.* 2013). For all such examples, the evolutionary genome concept takes a many-generation perspective on the phenotypic contributions to organismal fitness of sets of co-inherited genes (in comparison with just considering all DNA in a cell as 'the genome'), just as the evolutionary gene concept focuses on selection and long-term persistence of particular alleles (in comparison with considering 'genes' simply as the named entities in databases such as GenBank).

Usefulness and implications

The evolutionary conceptualization of genome described here is not meant to be 'true' or 'literal', but is instead meant to be valuable for ongoing thought and research in evolutionary biology, genetics and genomics. Perhaps most importantly, its consideration raises a new question – how to usefully conceptualize genomes from an evolutionary (or indeed, any other) perspective, because concepts structure the collection and interpretation of data. Witness, as a closely related example, recent discussions of the meanings of gene 'function' (Doolittle 2013; Graur *et al.* 2013), which directly determine how function is ascertained from biochemical and evolutionary criteria and guide collection and analyses of data.

We stress that many other conceptualizations of genomes are possible and effective, for addressing other questions. The primary usefulness of our idea rests in drawing attention to, and motivating analyses of, the fundamental processes of selection and inheritance that make any interacting set of genes, and gene products, functionally more or less coherent, rather than competitive or parasitic (Fig. 1).

For many research programmes, whether or not the gene or genome is considered as evolutionary will not matter. But with regard to ongoing empirical research in genomics, several practical implications follow from consideration of the evolutionary genome concept. First, the high frequencies of gene flow across populations and species, across recent and older evolutionary time (e. g., Grant & Grant 1992; Abbott *et al.* 2013), suggest that the genome of a 'species' should be interpreted carefully with regard to characterizing a single monophyletic lineage. To the extent that such gene flow is relatively recent, and selection is less effective, genetic incompatibilities and deviations from phenotypic adaptation are expected. In this general regard, might 'the genome' of a species or lineage sometimes also usefully be considered as the set of selected-effect (or otherwise defined) functional elements that are shared across individuals and populations?

Second, analyses of functionality at the genome level, involving interactions between genes, will benefit from explicit consideration of both cooperation and conflicts between genetic factions and elements (Haig 2006). Within-genome interactions are expected to include mutualism, parasitism, competition, and cooperation to compete, all among genomic factions. An evolutionary genomic view compels attention towards the degree to which the genome, as a whole, functions as an integrated, fitness-maximizing unit, with direct implications for fitness in natural populations, as well as human health. Are particular molecular genetic mechanisms involved in enforcing, or otherwise being selected to maximize, genome-scale integration and common interests? Such processes would be expected to operate at both the somatic-evolution level (mediating effects such as carcinogenesis and senescence) and the level of germ-line evolution (mediating repression of genomic conflicts).

Third, and following from the considerations just above, analysis and annotation of genomes tends to centre on protein-coding and single-copy sequences in autosomal genes, which are expected to exhibit the highest levels of functional integrative coherence. How common are cooperation and conflict-mediated positive and negative interaction effects between (as well as within) the myriad different genetic elements within cells and organisms? How much of fitness variation, and heritable variation, can be attributed to such interactions?

Additional applications include studies of microbiome interactions with host genomes, prokaryotic gene transfer and the functions of genomes in this context, and viruses and endosymbionts at the interfaces of parasitism, commensalism and mutualism. Most generally, novel perspectives, founded in evolutionary biology, on the meanings of genomes as well as genes, should lead to useful insights derived from new questions, and new genome-scale data targetted to address them.

© 2013 John Wiley & Sons Ltd

Acknowledgements

We are grateful to Wieslaw Babik, Steve Frank, Nolan Kane, Egbert Leigh, Mika Mokkonen, Patrik Nosil, Tanja Schwander, Steve Springer, Corey Watson, Dominika Wloch-Salamon and Stuart West for helpful comments and discussions.

References

Abbott R, Albach D, Ansell S *et al*. (2013) Hybridization and speciation. *Journal of Evolutionary Biology*, **26**, 229–246.

Abi-Rached L, Jobin MJ, Kulkarni S *et al*. (2011) The shaping of modern human immune systems by multiregional admixture with archaic humans. *Science*, **334**, 89–94.

Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. *Molecular Ecology*, **13**, 729–744.

Ben-Shlomo R, Neufeld E, Berger D *et al*. (2007) The dynamic of the t-haplotype in wild populations of the house mouse *Mus musculus* domesticus in Israel. *Mammalian Genome*, **18**, 164–172.

Bergsmedh A, Szeles A, Henriksson M *et al*. (2001) Horizontal transfer of oncogenes by uptake of apoptotic bodies. *Proceedings of the National Academy of Sciences, USA*, **98**, 6407–6411.

Boratynski Z, Alves PC, Berto S *et al*. (2011) Introgression of mitochondrial DNA among *Myodes* voles: consequences for energetics? *BMC Evolutionary Biology*, **11**, 355.

Boschetti C, Carr A, Crisp A *et al*. (2012) Biochemical diversification through foreign gene expression in bdelloid rotifers. *PLoS Genetics*, **8**, e1003035.

Brown TA (2007) *Genomes 3*. Garland Science, New York.

de Carvalho MO, Loreto EL (2012) Methods for detection of horizontal transfer of transposable elements in complete genomes. *Genetics and Molecular Biology*, **35**, 1078–1084.

Curtis BA, Tanifuji G, Burki F *et al*. (2012) Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. *Nature*, **492**, 59–65.

Dawkins R (1976) *The Selfish Gene*. Oxford University Press, Oxford.

Dawkins R (1982a) *The Extended Phenotype*. Oxford University Press, Oxford.

Dawkins R (1982b) Replicators and vehicles. *Current Problems in Sociobiology* (eds King's College Sociobiology Group), pp. 45–64. Cambridge University Press, Cambridge.

Didelot X, Darling A, Falush D (2009) Inferring genomic flux in bacteria. *Genome Research*, **19**, 306–317.

Dobzhansky T (1970) *Genetics of the Evolutionary Process*. Columbia University Press, New York.

Doerfler W (2012) Impact of foreign DNA integration on tumor biology and on evolution via epigenetic alterations. *Epigenomics*, **4**, 41–49.

Doolittle WF (2013) Is junk DNA bunk? A critique of ENCODE. *Proceedings of the National Academy of Sciences USA*, **110**, 5294–5300.

Dupressoir A, Lavialle C, Heidmann T (2012) From ancestral infectious retroviruses to bona fide cellular genes: role of the captured syncytins in placentation. *Placenta*, **33**, 663–671.

Frank SA (2003) Perspective: repression of competition and the evolution of cooperation. *Evolution*, **57**, 693–705.

Frank SA, Crespi BJ (2011) Pathology from evolutionary conflict, with a theory of X chromosome versus autosome conflict over sexually antagonistic traits. *Proceedings of the National Academy of Sciences USA*, **108**(Suppl 2), 10886–10893.

Gerstein MB, Bruce C, Rozowsky JS *et al*. (2007) What is a gene, post-ENCODE? History and updated definition. *Genome Research*, **17**, 669–681.

Gladyshev EA, Meselson M, Arkhipova IR (2008) Massive horizontal gene transfer in bdelloid rotifers. *Science*, **320**, 1210–1213.

Goldenberg DM (2012) Horizontal transmission of malignancy by cell-cell fusion. *Expert Opinion in Biological Therapy*, **12**(Suppl 1), S133–S139.

Grant PR, Grant BR (1992) Hybridization of bird species. *Science*, **256**, 193–197.

Graur D, Zheng Y, Price N, Azevedo RB, Zufall RA, Elhaik E (2013) On the immortality of television sets: "Function" in the human genome according to the evolution-free gospel of ENCODE. *Genome Biology and Evolution*, **5**, 578–590.

Haig D (2006) Intragenomic politics. *Cytogenetic and Genome Research*, **113**, 68–74.

Holmes EC (2011) The evolution of endogenous viral elements. *Cell Host & Microbe*, **10**, 368–377.

Hurst GD, Werren JH (2001) The role of selfish genetic elements in eukaryotic evolution. *Nature Reviews Genetics*, **2**, 597–606.

Koga R, Meng X-Y, Tsuchida T *et al*. (2012) Cellular mechanism for selective vertical transmission of an obligate insect symbiont at the bacteriocyte–embryo interface. *Proceedings of the National Academy of Sciences, USA*, **109**, 1230–1237.

Lederberg J (2000) Infectious history. *Science*, **288**, 287–293.

Lederberg J, McCray AT (2001) 'Ome sweet 'omics: a genealogical treasury of words. *The Scientist*, **15**, 8.

Leigh E (1991) Genes, bees and ecosystems: the evolution of common interest among individuals. *Trends in Ecology and Evolution*, **6**, 257–262.

Leigh E (2010) The group selection controversy. *Journal of Evolutionary Biology*, **23**, 6–19.

Ley RE, Peterson DA, Gordon JI (2006) Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell*, **124**, 837–848.

Martis MM, Klemme S, Banaei-Moghaddam AM *et al*. (2012) Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. *Proceedings of the National Academy of Sciences, USA*, **109**, 13343–13346.

Moran NA, McCutcheon JP, Nakabachi A (2008) Genomics and evolution of heritable bacterial symbionts. *Annual Review of Genetics*, **42**, 165–190.

Rebollo R, Romanish MT, Mager DL (2012) Transposable elements: an abundant and natural source of regulatory sequences for host genes. *Annual Reviews of Genetics*, **46**, 21–42.

Sheppard SK, Didelot X, Jolley KA *et al*. (2013) Progressive genome-wide introgression in agricultural *Campylobacter coli*. *Molecular Ecology*, **22**, 1051–1064.

Shurtliff QR (2013) Mammalian hybrid zones: a review. *Mammal Review*, **43**, 1–21.

Trask JAS, Garnica WT, Smith DG *et al*. (2013) Single-nucleotide polymorphisms reveal patterns of allele sharing across the species boundary between Rhesus (*Macaca mulatta*) and Cynomolgus (*M. fascicularis*) Macaques. *American Journal of Primatology*, **75**, 135–144.

Trejo-Becerril C, Pérez-Cárdenas E, Taja-Chayeb L *et al*. (2012) Cancer progression mediated by horizontal gene transfer in an *in vivo* model. *PLoS ONE*, **7**, e52754.

Van Vugt JJ, de Jong H, Stouthamer R (2009) The origin of a selfish B chromosome triggering paternal sex ratio in the parasitoid wasp *Trichogramma kaykai*. *Proceedings of the Royal Society of London, Biological Sciences*, **276**, 4149–4154.

Venner S, Feschotte C, Biémont C (2009) Dynamics of transposable elements: towards a community ecology of the genome. *Trends in Genetics*, **25**, 317–323.

© 2013 John Wiley & Sons Ltd

Welch RA, Burland V, Plunkett G 3rd *et al*. (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proceedings of the National Academy of Sciences, USA*, **99**, 17020–17024.

Wiedenbeck J, Cohan FM (2011) Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiol Reviews*, **35**, 957–976.

Williams GC (1966) *Adaptation and Natural Selection*. Princeton University Press, Princeton, New Jersey.

Williams GC (1992) *Natural Selection: Domains, Levels, and Challenges*. Oxford University Press, New York.

Yue J, Hu X, Sun H *et al*. (2012) Widespread impact of horizontal gene transfer on plant colonization of land. *Nature Communications*, **3**, 1152.

Zilber-Rosenberg I, Rosenberg E (2008) Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiology Reviews*, **32**, 723–735.

A. Stencel and B. Crespi conceived and wrote the paper together.

---

doi: 10.1111/mec.12355

© 2013 John Wiley & Sons Ltd