Recent advances in understanding the roles of whole genome duplications in evolution [version 2; peer review: 2 approved]

Carol MacKintosh\textsuperscript{1}, David E.K. Ferrier\textsuperscript{2}

\textsuperscript{1}Division of Cell and Developmental Biology, University of Dundee, Dundee, Scotland, DD1 5EH, UK
\textsuperscript{2}The Scottish Oceans Institute, University of St Andrews, Scotland, KY16 8LB, UK

**Abstract**

Ancient whole-genome duplications (WGDs)—paleopolyploidy events—are key to solving Darwin's 'abominable mystery' of how flowering plants evolved and radiated into a rich variety of species. The vertebrates also emerged from their invertebrate ancestors via two WGDs, and genomes of diverse gymnosperm trees, unicellular eukaryotes, invertebrates, fishes, amphibians and even a rodent carry evidence of lineage-specific WGDs. Modern polyploidy is common in eukaryotes, and it can be induced, enabling mechanisms and short-term cost-benefit assessments of polyploidy to be studied experimentally. However, the ancient WGDs can be reconstructed only by comparative genomics: these studies are difficult because the DNA duplicates have been through tens or hundreds of millions of years of gene losses, mutations, and chromosomal rearrangements that culminate in resolution of the polyploid genomes back into diploid ones (rediploidisation). Intriguing asymmetries in patterns of post-WGD gene loss and retention between duplicated sets of chromosomes have been discovered recently, and elaborations of signal transduction systems are lasting legacies from several WGDs. The data imply that simpler signalling pathways in the pre-WGD ancestors were converted via WGDs into multi-stranded parallelised networks. Genetic and biochemical studies in plants, yeasts and vertebrates suggest a paradigm in which different combinations of sister paralogues in the post-WGD regulatory networks are co-regulated under different conditions. In principle, such networks can respond to a wide array of environmental, sensory and hormonal stimuli and integrate them to generate phenotypic variety in cell types and behaviours. Patterns are also being discerned in how the post-WGD signalling networks are reconfigured in human cancers and neurological conditions. It is fascinating to unpick how ancient genomic events impact on complexity, variety and disease in modern life.
Comparative plant genomics help solve Darwin’s abominable mystery

Darwin was vexed. While natural selection could explain gradual evolutionary transitions, the apparent sudden appearance of diverse flowering plants in the Cretaceous fossil record of approximately 130 million years ago (Mya) was his ‘abominable mystery’ 1. Fast-forward to today’s exciting era of high-throughput genome sequencing, and phylogenomic maps assembled from multiple whole-genome sequences tell the evolutionary story with revised timelines:

From the Carboniferous to early Cretaceous periods (approximately 360 to 130 Mya), the land was dominated by gymnosperms (literally ‘naked seeds’) including the cycads, Ginkgo, and conifers that still flourish in subarctic forests 2. The nuclear genomes of several gymnosperms have been sequenced recently, a heroic undertaking, given their exceptional size (10 to 40 gigabases) and high density of long terminal repeat (LTR)-retrotransposon repeats 3-4. Within these genomes many non-overlapping duplicated chromosomal regions were identified that display gene synteny, meaning that their gene contents are similar to those of other chromosomal blocks within the same genome and across gymnosperm genomes. These gene synteny patterns and complementary transcriptome data support the hypothesis that the gymnosperms emerged from their common ancestor via a WGD named ε that occurred an estimated 390 Mya during the Devonian period 5-6. Tell-tale traces of further lineage-specific WGDs have also been discovered in the genomes of Norway spruce, Sequoia and Ginkgo, and in the unusual two-leaved Namibian Welwitschia, suggesting that multiple WGDs contributed to the diversity of these gymnosperms, which include the longest-living and largest organisms on Earth 3-4.5.

Flowering plants (angiosperms, ‘seed born in a vessel’) are the most abundant plant group today, having a rich diversity of some 400,000 species from bananas to water lilies, grasses to beech trees 6. Many angiosperm genomes have been selected for sequencing to discover genes for special agronomic traits. The resulting genome assemblies have been used collectively to extrapolate back in time, tracking the genome evolution of the monocot and eudicot angiosperms, as well as more primitive flowering plants, and converging on a reconstructed genome of the most recent common ancestor of all angiosperms 7. These new phylogenomic maps solve part of Darwin’s dilemma by confirming that the first flowering plants evolved between 140 and 250 Mya after an unknown gymnosperm went through a WGD (named the ε event) an estimated 300 Mya during the Carboniferous period 8-9. The clearest support for the ε WGD comes from multiple gene synteny blocks in the genome of the primitive angiosperm Amborella trichopoda, whose ancestral lineage diverged early on from other flowering plants, experiencing no further post-ε WGDs 10.

The antiquity of the ε WGD means that angiosperms had the entire Triassic and Jurassic periods to evolve and diversify into the species richness reflected in the fossils of approximately 130 Mya that were known to Darwin. Aligning with the recalibrated timelines, older angiosperm microfossils have been discovered, though claims that the earliest ones date to the Triassic are controversial 11-12. In any case, angiosperm evolution was not as sudden an explosion as Darwin thought, and now a comprehensive phylogenomic framework exists to mine for answers about how angiosperm complexity and variety evolved after the ε WGD.

Contributions of whole-genome duplications to the origin and diversity of flowers

In the 1970s, Susumo Ohno had the prescience to propose that evolutionary leaps could occur by WGDs because one of each gene pair may continue to do what it was doing before, giving freedom for the other to either be lost from the genome or to evolve new characteristics (neofunctionalise) 13. Members of gene families generated via WGDs are named ohnologues in his honour 14. Another scenario is for the duplicates to each retain subsets of the functions of the ancestral gene (subfunctionalisation) 15. Moreover, the en masse diversification of many gene duplicates after a WGD would be expected to create selective pressures between gene families and opportunities for new interactions among them, so that large duplicated gene and protein sets co-evolve as complex systems.

Floral organs provide a canonical example of how interacting sets of diversified ohnologues can make variant structures. While many gene pairs lost one duplicate after the ε WGD, the retained ohnologue pairs include MADS-box transcription factors. Some of these were characterised more than 20 years ago for their ability to interact with each other in different combinations to specify the floral organs of Antirrhinum (snapdragon) and Arabidopsis. In these seminal studies, homeotic floral mutants were used to deduce an elegant ‘ABC model’, later elaborated into an ‘ABCDE model’ that explains how the four concentric whorls of sepals, petals, male stamens and female carpels develop in these flowers 16-19. A-function genes specify sepals; B, C and E are needed to make petals; B, C and E male stamens; C female carpels; and D for ovules. In homeotic mutants lacking B-function, for example, sepals replace petals and carpels replace stamens 20. Most of the A, B, C, D and E functions are performed by combinations of MADS-box transcription factors that operate as homodimers and heterodimers and tetramers with different selectivities for binding to variants of a common motif in the promoters of target genes 16-19,21. Hence, they target overlapping but distinct sets of floral identity genes many of which are themselves ohnologues 14.

The recent genome comparisons indicate that male ‘BC’ and female ‘C’ systems already existed to specify reproductive cells in gymnosperm cones, and they were duplicated via the ε WGD, after which the C duplicates diversified into angiosperm C and D genes. An A/E gymnosperm pro-orthologue gave rise to angiosperm A and E genes, and further duplicated A genes were also retained after the ε WGD 22-24. These duplicated and diversified gene sets organised to generate the first now-extinct flowers, and recent reconstructions suggest that these were bisexual with petal-like tepals and pollen-bearing stamens arranged in multiple concentric whorls, and
female carpels in a central spiral. Among living angiosperms, the Amborella lineage evolved a ‘fading borders’ programme, such that the whole flower is a spiral that gradually transitions from bracts to outer then inner tepals (specified by ABc combinations), from inner tepals to stamens (Aa) then carpels (abC), in which upper case indicates functions of greatest influence in the respective organs. Only in later-evolving flowers such as Arabidopsis and Antirrhinum did the tepals subdivide into sepal and petals, by restricting the boundaries of expression of floral identity genes. For example, the transcription of A and C genes became mutually exclusive. Further evolutionary diversity in flower form occurred by mechanisms that include shifts in the spatial expression of ABC functions across flowers, and by further WGDs that elaborated and extended the ABC regulatory network. For example, in stylised orchid flowers, subfunctionalisation of duplicated B genes underpins the development of three types of petals: three outer tepals, two inner tepals and a modified lip. Recent case studies implicate additional ancient WGDs – including one at the base of the eudicots, the γ genome triplication in the Pentapetalae (five-parted, the largest flower clade), and ρ, σ and τ polyploidisations for monocots – in evolution of the phenomenal variety of architectural form and size in pollen, fruits and seeds, in diversification of plant defence metabolites, and in the co-evolution of angiosperms with pollinators and symbiotic bacteria.

Mechanisms and cost-benefit analyses in recent natural and experimental polyploides

Evolutionarily recent polyploidy events are also prevalent in flowering plants. Many crops including coffee, bananas, peanuts, tobacco, kiwifruit and strawberries were unwittingly selected as polyploids for their exaggerated traits such as large fruits, seeds and leaves. For example, the durum wheat used to make pasta is a tetraploid resulting from hybrid doubling of the genomes of two diploid wild-grass ancestors approximately 0.5 Mya, and was selected for domestication much later by Neolithic farmers, during which time hexaploid bread wheat emerged by hybridisation of the tetraploid with a diploid followed by another WGD. Like wheat, many well-characterised crop polyploids are allopolyploid, which means that the genome became polyploid after a hybrid was formed between species, in which case the WGD resolved problems with meiotic pairing by providing each chromosome with a homeologous partner. However, autopolyploidy events, self-duplication within a species, are suspected in the ancestry of potatoes, bananas, poplar and soybean.

Recent statistical comparisons suggest that an individual polyploid plant has a higher risk of extinction than its still-diploid relatives. It makes sense that the odds are stacked against a newly tetraploid plant. Breeding with still-diploid relatives results in triploid progeny that cannot separate evenly into two gametes during meiosis, most often resulting in sterile offspring, and unless self-pollination can occur, chances may be low of finding compatible polyploid mates. Farmers get around such problems by cloning: grafting apple trees or propagating potato tubers instead of seeds, whilst sterility (for example in seedless bananas) is sometimes even preferred. However, the resulting monocultures may be susceptible to pathogens, as when the Fusarium oxysporum fungal pandemic brought the popular triploid Gros Michel banana to the brink of extinction in the 1960s.

Nevertheless, their prevalence suggests that once polyploids have beaten the early survival odds, with or without human intervention, their polyploid traits such as larger organs, stress tolerance and altered flowering time may improve fitness or allow them to adapt to new ecological niches. Experimental polyploidies show that having extra DNA can produce an immediate phenotypic change, attributable to gene dosage effects. For example, dwarfism in apple plants with colchicine-induced autotetraploidy correlates with increased expression of a microRNA that acts via a gene regulatory network to downregulate synthesis of auxin and brassinosteroid growth regulators. More generally, polyploid plants are notable for their increased cell and organ size, which is more than a passive consequence of increased nuclear DNA content: in experiments with Arabidopsis, increases in cell volume upon tetraploidisation were found to vary in different mutants and according to cell type, indicating a genetic contribution.

For allopolyploids, the relative contributions of the species hybridisations and the WGDs to subsequent evolution are interwoven. In a recent molecular dissection of the circadian clock in allotetraploids formed between diploids Arabidopsis thaliana (At) and Arabidopsis arenosa (Aa), biases in heterologous combinations of components were discovered: the Aa-derived CCA1 hiking expedition (CHE) transcription factor preferentially binds to the promoter of the At circadian clock associated 1 (CCA1) gene, elevating its expression over that of the AaCCA1. Such biased patterns of expression, and of protein–protein and protein–DNA interactions in the circadian regulatory network make the rhythm of the allotetraploid distinct from that of either parent.

Common themes from recent reconstructions of ancient whole-genome duplications in plants, animals and fungi

Successful polyploidy is said to be less common in animals than in plants. Based on incidences of chromosomal anomalies in embryos that fail to develop, it appears that when two sperm fertilise one egg or when meiotic cell division fails the result is usually lethal in humans and birds. However, polyploidy is relatively common in ectothermic vertebrates. Also, synthetic fish and shellfish polyploids, generally sterile, have been created to increase food production. Moreover, helped by technical advances in deep sequencing, genome assembly and pattern-matching software, ancient WGDs have been identified in invertebrate and vertebrate animal lineages, including mammals. The most recent discovery was that the house spider Parasteatoda tepidariorum and bark scorpion Centruroides sculpturatus are common descendants of a WGD that occurred over 450 Mya, and was distinct from an ancestral WGD of horseshoe crabs.

Although few ancient WGDs have been identified thus far in unicellular eukaryotes, the diploid baker’s yeast Saccharomyces cerevisiae and five other fungal genera all stem from the same well-characterised ancestral allopolyploid WGD approximately 100 Mya, and further fungal WGDs have been identified.
recently. For example, the opportunistic honeybee fungal pathogen *Nosema ceranae*, which is spreading to beehives worldwide, is a suspected tetraploid⁶⁰.

In summary, ancient WGDs that were successful in leaving modern descendants have occurred in diverse eukaryotes across eons of time, in terrestrial and aquatic environments. Remarkably, despite their radically different contexts and a sparsity of data on ancient WGDs outside of laboratory models and domesticated species⁶³, common principles are emerging that tie disparate WGD events together:

- **Successive WGDs have occurred in multiple lineages**: As indicated, angiosperms have experienced multiple WGDs, and recursive WGDs among the *Brassica* crops that include cauliflower, broccoli and cabbages have been precisely mapped recently⁶⁸. The vertebrate animals emerged from the invertebrates approximately 500 Mya via two sequential rounds of WGD (2R-WGD)⁶⁹. In fish there was a further teleost-specific WGD (TSGD, 3R) approximately 300 Mya, followed by a salmonid-specific WGD (Ss4R) approximately 95 Mya; and certain *Xenopus* frog species and the red viscacha rat *Tympanoctomyos barraei* also result from lineage-specific WGDs⁶⁰-⁶⁴. The unicellular ciliate *Paramecium tetraurelia* has a history of three successive WGDs⁶⁵, and WGDs may have contributed to the record number of chromosomes (2n=1260) in the fern-like genus *Ophioglossum*⁶⁶.

- **Most, though not all, well-characterised WGDs were allopolyploid events**: In common with most polyploid plants, the tetraploid frog *Xenopus laevis* and diploid yeast *S. cerevisiae* were recently identified to be descendants of allopolyploidies⁶⁷-⁶⁸. In contrast, high similarity between homoeologous regions in salmonid genomes indicate that the TSGD/3R and salmonid Ss4R were allopolyploid events⁶⁹-⁷⁰. The mechanisms of the 2R-WGD at the origin of the vertebrates 500 Mya are unresolved, although early studies argued for two closely spaced autotetraploidies⁷¹.

- **Long lag periods may occur between WGDs and subsequent species radiations**: A WGD generates a new organism that is immediately distinct from the parental species, especially after inter-species hybridisation allopolyploidies. Intuitively, one would expect this new organism to lead to species radiations due to the availability of new genetic material for evolution to mould in different ways. In practice however, the mechanistic links between WGDs and species radiations are not so clear-cut. In many instances there is a time-lag between WGDs and species radiations, formalised as the WGD Radiation Lag-Time model⁷². For example, comparisons of post-TSGD fishes, including zebrafish and Japanese medaka, show that post-TSGD genome changes were biphasic. An initial period of bulk losses of chromosomal segments was overlaid by a more extended period of gradual gene losses by pseudogenisation and mutation of the retained ohnologues. The latter phase, after the initial bulk genome reshaping, correlates with radiation of bony fish species⁷⁳, though how or whether WGD is mechanistically linked to fish diversification is still an open question⁷⁴,⁷⁵. Among the fungi, comparisons of the six genera that share the same ancestral WGD as *S. cerevisiae* suggest that around 4000 genes still existed in duplicate when these genera were diverging from each other, with subsequent losses of different paralogues in different lineages, such that, for example, *S. cerevisiae* now has 551 pairs of ohnologues and *Candida glabrata* has 404 pairs⁷⁵-⁷⁷. The next two bulleted points further discuss how mechanisms of post-WGD genomic evolution, as well as environmental influences, steer the course of speciation and phenotypic diversity after a WGD.

- **Post-WGD chromosomal rearrangements culminate in a return to the diploid state (rediploidisation) as well as lineage divergence**: Allopolyploidy results in immediate rediploidisation (diploid pairing of homoeologous chromosomes during meiosis/mitosis) if the chromosomes from the two parental species are sufficiently distinct that chromosomes do not form quadrivalents during cell division. In contrast, autopolyploidy leads into a process of gradual rediploidisation⁷⁸, such that descendants of some WGDs that happened just tens of Mya are still polyploid, whereas species whose last WGD occurred hundreds of Mya (certain angiosperms and most vertebrates) have reverted to diploid. The post-Ss4R salmonids provide interesting snapshots of genomes in transition – some chromosomes are still functionally tetraploid whereas others have become diploid. Examining these genomes in different species reveals that the dynamic interplay of rediploidisation, ohnologue divergence, speciation and post-speciation evolution is complicated⁷⁹. For example, regions of salmonid genomes have been identified for which rediploidisation and evolutionary diversification of ohnologues occurs after speciation, such that functional divergence of ohnologues occurs in lineage-specific ways (Lineage-specific Ohnologue Resolution, LORe). LORe may facilitate adaptations of the distinct species to different ecological contexts⁸⁰.

- **WGDs and species radiations following WGDs have been linked with major climate change**: Many angiosperm WGDs date to the asteroid-triggered Cretaceous-Paleogene boundary events approximately 66 Mya, indicating that polyploid establishment may be favoured during times of environmental stress⁷⁷. After the Ss4R of approximately 88 Mya, the greatest species radiation occurred in fish that evolved the ability to migrate between fresh water and seawater following the climatic cooling of the Eocene–Oligocene transition 40 to 50 million years later⁸¹. Such correlations between polyploidisations and environmental changes strengthen the view that the two are linked, perhaps due to polyploid organisms being more robust to environmental change and stress⁸².

- **Knowledge of WGDs informs how laboratory animals are used as biomedical models**: Zebrafish and polyploid *Xenopus* frogs are valuable models for development and disease. However, they have been through lineage-specific WGDs that humans have not, which means that phenotypes may differ when ohnologues are mutated in zebrafish, polyploid frogs, and humans⁸³,⁷⁹. Moreover, families of sister ohnologues from the 2R-WGD may subfunctionalise or neofunctionalise in different ways along different vertebrate lineages, as has been found for
the neurogenin and snailslug genes\textsuperscript{82,83}. It is therefore important to consider the composite functions of all sister ohnologues when making cross-species comparisons.

- **Asymmetries in the fates of DNA duplicates occur at multiple levels after ancient allopolyploidies:** Patterns of gene loss, retention and mutation may differ markedly in the sub-genomes derived from each parent of the original hybrid\textsuperscript{84}. A striking example is the tetraploid *X. laevis* in which large- and small-scale losses of DNA differ to such an extent that chromosomes derived from one parental species are markedly shorter than chromosomes from the other\textsuperscript{85}. It has been proposed that biases in the fates of DNA duplicates may result from ancient allopolyploidies, but that more evenly balanced post-WGD patterns of gene loss, retention and differentiation may follow on from ancient autopolyploidies\textsuperscript{86}. The proposal is that initial differences in expression levels and in affinities of interactions of ohnologue proteins derived from two parental species\textsuperscript{84} could propagate into further knock-on biases such as preferential retention of highly-expressed genes\textsuperscript{85,86}. Another type of post-WGD bias occurred in the shared ancestry of *S. cerevisiae* and *Candida glabrata*. Their ancestral WGD occurred approximately 100 Mya, after one parent from the KLE (*Kluyveromyces, Lachancea, Eremothecium*) clade mated with one of the ZT (*Zygosaccharomyces, Torulaspora*) clade. However, for various reasons the mixed parentage of *S. cerevisiae* was not immediately obvious when its genome was sequenced. One reason is that the *S. cerevisiae* genome contains more ZT-than KLE-derived sequences, possibly due to biased gene conversion that replaced some KLE-derived sequences with homoeologous ZT-derived ones\textsuperscript{87,88}. Finding that *S. cerevisiae* results from an allopolyploidy, rather than an autopolyploidy as originally believed, means that calculations of the relative rates of evolution of its ohnologue pairs\textsuperscript{89} may need revision.

- **Retained ohnologue gene families are strikingly enriched in signalling and regulatory proteins in plants, fungi and animals\textsuperscript{90,91}.** For example, ancestral WGDs that were identified recently for the fungi *Mucor circinelloides* and *Phycomyces blakesleeanus* resulted in increased proportions of genes whose transcription is regulated by light\textsuperscript{91}. In humans, while only approximately 25% of genes are ohnologues stemming from the 2R-WGD, ~66% of protein kinases and nearly 90% of well-characterised 14-3-3-binding phosphoproteins are ohnologues; and developmental regulators and post-synaptic density (PSD) brain proteins are also highly enriched in ohnologues\textsuperscript{92,93}. Signalling in biology ranges from simple signal-response systems to the complex signalling networks of our brains that coordinate complicated actions, create memories and find meaning in patterns. How have WGDs shaped these networks?

### Post-whole-genome duplication evolution of parallel processing via duplicated signalling networks, and dysregulation in cancers and neurological disorders

Studies in *S. cerevisiae*, plants and mammals have shown that regulatory proteins that form oligomers, that interact transiently with multiprotein complexes, and catalyse consecutive steps in metabolic and regulatory pathways are enriched among duplicate pairs that are retained following a WGD\textsuperscript{94,95}. These findings are interpreted by the gene balance hypothesis, which states that copy numbers of genes encoding multi-protein structures and pathways must be kept in a constant ratio to avoid architectural disruption or metabolic imbalances\textsuperscript{89}, although stoichiometry can also be achieved by other mechanisms such as differential degradation of protein components\textsuperscript{93}.

Interestingly, the architectures of sister ohnologue proteins are generally conserved with respect to content and order of their domains, at least in vertebrates. Instead, functional divergence between sisters occurs via small-scale mutations that lead to differences in temporal and spatial patterns of expression, altered sites of regulatory post-translational modifications, and changes in specificities and affinities of catalytic domains and interaction interfaces\textsuperscript{89,92,96,97}.

Examples abound to illustrate how the resulting families of differentiated sister ohnologues contribute to phenotypic robustness, plasticity and complexity. For instance:

- **The ‘Gli code’ refers to how different combinations of sister Gli transcription factors—Gli1 and Gli2 that activate transcription and the repressor Gli3— influence tissue shape and size during vertebrate embryonic development.** The code can be changed via differential regulation of the sister Gli proteins by multiple signalling pathways\textsuperscript{98}.

- **Genetic and biochemical dissections indicate that sets of sister ohnologues within the PSD ‘supercomplexes’ of mammalian brains can differentially process the signals responsible for individual cognitive abilities and emotions\textsuperscript{99}.** For example, in mammals the discs large MAGUK scaffold protein 4 (Dlg4) has evolved a role in simple associative learning, whereas its sisters Dlg2 and Dlg3 have distinct and opposing functions in complex cognitive processes\textsuperscript{92}.

- **In *S. cerevisiae*, differential pathways can be created by co-ordinated expression of different combinations of ohnologues that have been largely partitioned into obligate sub-networks\textsuperscript{100}.** Switching from one sub-network to another allows the yeast to adapt to different stresses, and to reconfigure fluxes through metabolism to enable growth on different types and quantities of sugars\textsuperscript{100–104}.

- **The phosphoprotein-binding 14-3-3 proteins interact with hundreds of (phosphorylated) ohnologue proteins in mammalian cells, suggesting that regulated phosphorylation of 14-3-3 docking sites provides a large-scale mechanism for switching from one set of sister ohnologues to another.** Indeed, case studies indicate that one set of ohnologues within a cell may be phosphorylated and consequently bind to 14-3-3s when cells are stimulated by insulin for example, while other combinations of sister ohnologues bind to 14-3-3s in response to phosphorylations that are promoted by growth factors, nutrient stress and adrenalina. Partitioning of ohnologues into obligate sub-networks does not seem to be so clear-cut as in *S. cerevisiae* however, as mammalian ohnologues have been identified that are convergence points for regulated interactions with 14-3-3 proteins in response to multiple stimuli\textsuperscript{105–109}.  

---

Page 6 of 11
These few examples—and the aforementioned ABCDE model of floral development—indicate how the parallelised signalling networks of ohnologues generated via WGDs act as multiple-input multiple-output systems. Collectively, these systems generate different cell phenotypes via differential expression and post-translational switching among sets of ohnologues with different kinetic and regulatory properties.

Deeper understanding of how post-WGD signalling networks operate will underpin advances in understanding polygenic disorders. For instance, mutations in many ohnologue genes are associated with neurological and psychiatric diseases and developmental disorders such as RASopathies[107, 108], and there are many examples of heterogeneous patterns of overexpressions and mutations across ohnologue gene families in cancers[99-111]. For example, overexpression of insulin receptor substrate 4 (IRS4) drives a subset of breast cancers, while IRS1 and IRS2 are not oncogenic in these cancers, even though all three IRS proteins activate PI 3-kinase–Akt growth signalling. The critical difference is that IRS4 lacks a negative feedback mechanism by which its sisters IRS1 and IRS2 can switch off the pathway via the tyrosine phosphatase SHP2[112]. This example illustrates how the parallel signalling pathways generated via WGDs can evolve specific regulatory interconnections, which in principle enable these systems to integrate inputs from multiple sensory stimuli, buffer signal noise via responsive feedback loops, and generate a wider repertoire of phenotypic outcomes than would be possible with the original simple circuit[112,113].

Finally, it should be noted that WGDs do not underpin every evolutionary leap: A WGD was hypothesised to explain why cephalopods (squids, cuttlefish and octopuses) are behaviourally more sophisticated than other molluscs. However, the octopus genome shows no evidence of a WGD. Rather, sensory intelligence in cephalopods is attributed than other molluscs. However, the octopus genome shows no evidence of a WGD. Rather, sensory intelligence in cephalopods is likely to be underpinned by the massively expanded gene families of C2H2 zinc-finger transcription factors and protocadherins that are expressed in their neuronal and sensitive tissues[114]. It will be fascinating to compare the neural network topologies that have been built from gene families generated via WGDs versus those made of multiple small-scale duplications in humans and octopuses.

Competing interests
The authors declare that they have no competing interests.

Grant information
The author(s) declared that no grants were specifically involved in supporting this work.

References


Page 7 of 11
PubMed Abstract | Publisher Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text


PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

77. F1000Research 2018, 6(F1000 Faculty Rev):1623 Last updated: 30 MAR 2022


PubMed Abstract | Publisher Full Text | Free Full Text


PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text


PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text

PubMed Abstract | Publisher Full Text | Free Full Text


Open Peer Review

Current Peer Review Status: ✔️ ✔️

Editorial Note on the Review Process

Faculty Reviews are review articles written by the prestigious Members of Faculty Opinions. The articles are commissioned and peer reviewed before publication to ensure that the final, published version is comprehensive and accessible. The reviewers who approved the final version are listed with their names and affiliations.

The reviewers who approved this article are:

1. Daniel Macqueen
   Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, UK
   **Competing Interests:** No competing interests were disclosed.

2. Gavin Conant
   Department of Biological Sciences, North Carolina State University, Raleigh, North Carolina, USA
   **Competing Interests:** No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com