Incompatibility and Interchangeability in Molecular Evolution

Daniel B. Sloan[1,*], Jessica M. Warren[2], Alissa M. Williams[3], Shady A. Kuster[1], Evan S. Forsythe[1]

[1]Department of Biology, Colorado State University. Fort Collins, CO, USA

[2]Center for Mechanisms of Evolution, Biodesign Institute and School of Life Sciences, Arizona State University, Tempe, AZ, USA

[3]Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA

*Corresponding author: dan.sloan@colostate.edu

Abstract

There is remarkable variation in the rate at which genetic incompatibilities in molecular interactions accumulate. In some cases, minor changes – even single nucleotide substitutions – create major epistatic incompatibilities when hybridization forces new variants to function in a novel genetic background from an isolated population. In other cases, genes or even entire functional pathways can be horizontally transferred between anciently divergent evolutionary lineages that span the tree of life with little evidence of incompatibilities. In this review, we explore whether there are general principles that can explain why certain genes are prone to epistatic incompatibilities while others maintain interchangeability. We summarize evidence pointing to four genetic features that may contribute to greater resistance to functional replacement: 1) function in multisubunit enzyme complexes and protein-protein interactions, 2) sensitivity to changes in gene dosage, 3) rapid rate of sequence evolution, and 4) overall importance to cell viability, which creates sensitivity to small perturbations in molecular function. We discuss the relative levels of support for these different hypotheses and lay out future directions that may help explain the striking contrasts in patterns of incompatibility and interchangeability throughout the history of molecular evolution.

KEYWORDS

Cytonuclear, Epistasis, Horizontal Gene Transfer, Hybridization, Protein-Protein Interactions

Introduction

A casual scan of the literature could yield radically different – but equally justifiable – conclusions about the robustness of genetic systems, depending on which corners of biology a reader happens to stumble into. On one hand, mutations that alter a single nucleotide can inactivate an entire gene and even produce lethal effects (Eyre-Walker and Keightley 2007), illustrating the fragility of many genetic systems. In other cases, organisms are astonishingly tolerant of major changes, such as genome-wide modifications to the genetic code (Mukai, et al. 2017) or addition of entire genomes (Itaya, et al. 2005; Tagwerker, et al. 2012). This contrast is especially evident in molecular interactions between gene products. For example, a single nucleotide substitution in a mitochondrial tRNA gene present in natural populations of the fruit fly *Drosophila simulans* has been shown to produce major incompatibilities when paired with a single amino-acid substitution in an interacting aminoacyl-tRNA synthetase (aaRS) enzyme from *Drosophila melanogaster* (Meiklejohn, et al. 2013); and yet, aaRSs undergo widespread horizontal gene transfer (HGT) across disparate domains of life and functionally replace counterparts that are highly divergent in sequence (Figure 1) (Woese, et al. 2000). Similarly, a few amino acid substitutions in interacting subunits within the mitochondrial NADH dehydrogenase complex of swordtail fishes (*Xiphophorus*) appear to be responsible for a lethal incompatibility (Moran, et al. 2021); and yet, subunits within the mitochondrial ribosome, another mitonuclear (alphaproteobacterial-like) enzyme complex, have been entirely replaced by anciently divergent counterparts from plastid (cyanobacterial-like) or cytosolic (archaeal-like) ribosomes in some plant lineages (Adams, et al. 2002). Such observations lead us to ask whether there are general principles to explain why certain systems are prone to rapid evolution of epistatic incompatibilities while others remain interchangeable even after billions of years of divergence.

We specifically selected the foregoing examples from the field of mitochondrial biology because the endosymbiotic history of eukaryotes may be especially valuable for disentangling the mechanisms that preserve interchangeability or lead to incompatibilities. The repeated merging of evolutionary lineages associated with the acquisition of mitochondria, plastids, and other bacterial endosymbionts creates redundancies between genetic systems and ample supply of material for HGT (which is also known as endosymbiotic or intracellular gene transfer in this context) (Timmis, et al. 2004; Sloan, et al. 2018). Mitochondria and plastids retain their own genomes (albeit highly reduced ones) while also importing thousands of nuclear-encoded proteins. As a result, organellar functions depend on direct molecular interactions between gene products encoded in different genomes. For example, the major OXPHOS enzyme complexes responsible for cellular respiration are composed of both nuclear- and mitochondrial-encoded protein subunits (Rand, et al. 2004; Burton, et al. 2013). Even though they are found within the same cell, nuclear and cytoplasmic genomes can differ in key biological properties such as mode of inheritance, mutation rate, genome copy number, and expression level (Lynch, et al. 2006; Smith and Keeling 2015; Forsythe, et al.

2022). Such asymmetries can help test hypotheses regarding the evolutionary forces that contribute to genetic incompatibilities.

Here, we review biological examples that illustrate the broad spectrum that ranges from incompatibility to interchangeability at the molecular level, pointing to four general principles that may explain where specific genes and functional pathways are placed along this spectrum.

Epistatic incompatibilities exposed by hybridization and HGT

One of the central goals of evolutionary biology is to identify the genetic and molecular basis of reproductive barriers that lead diverging populations to eventually evolve into isolated species. Some common themes about the genomic architecture of reproductive isolation have emerged from analysis of natural and lab-generated hybrids, including the effect of inversions and other recombination suppressors (Schumer, et al. 2018; Schluter and Rieseberg 2022) and the disproportionate role of sex chromosomes (Presgraves 2008; Presgraves 2018).

Studies have also been increasingly successful in pinpointing examples of specific genes involved in postzygotic reproductive isolation in the form of so-called Bateson-Dobzhansky-Muller incompatibilities (BDMIs; Table 1) (Johnson 2010; Bozdag and Ono 2022). The growing list of these "speciation genes" is enriched for certain functional categories. We have already noted examples of mitonuclear incompatibilities associated with direct physical interactions between mitochondrial gene products and imported nuclear-encoded proteins (Meiklejohn, et al. 2013; Moran, et al. 2021). These and similar examples have suggested that mitochondrial genes are frequent contributors to reproductive isolation and speciation (Burton and Barreto 2012; Hill 2016; Sloan, et al. 2017; Postel and Touzet 2020; Bozdag and Ono 2022). Meanwhile, many of the nuclear genes that have been implicated in BDMIs are involved in various forms of genomic conflict and antagonistic coevolution, including centromere binding, transposable element activity, male sterility, testis-specific functions, and pathogen defense (Johnson 2010; Crespi and Nosil 2013; Sankararaman, et al. 2014; Serrato-Capuchina and Matute 2018; Postel and Touzet 2020; Schluter and Rieseberg 2022). These recurring functional themes suggest that certain genes are more prone than others to developing epistatic incompatibilities.

Although hybridization and introgression studies have been highly informative in identifying genetic incompatibilities, they are inherently limited to recent histories of divergence because they depend on lineages that remain at least partially interfertile. The history of HGT between more anciently divergent lineages provides an alternative avenue to determine which genes preferentially build up incompatibilities and which remain highly interchangeable. Comparative studies have been valuable in identifying biological features associated with genes that are especially likely or unlikely to undergo HGT (Rivera, et al. 1998; Sorek, et al. 2007; Cohen, et al. 2011; Creevey, et al. 2011; Baltrus 2013; Nagies, et al. 2020). Although most of this HGT work has focused on the gain of novel

functions, HGT can also result in the replacement of homologous genes and existing functions (Koonin, et al. 2001; Andam and Gogarten 2011; Creevey, et al. 2011; Huang and Yue 2013; Nagies, et al. 2020). Such examples of direct functional replacement via HGT are particularly relevant to the subject of this review because they inform our understanding of interchangeability.

Laboratory experiments have complemented comparative analyses of HGT by allowing for more controlled and systematic tests of gene transferability (Table 1). In one classic study, Sorek, et al. (2007) took advantage of the fact that early genome projects involved cloning shotgun gene libraries into *E. coli*. The authors reasoned that gaps in genome assemblies that required closing by PCR could be used to identify genes that hindered *E. coli* growth and viability. More generally, heterologous expression and mutant rescue experiments in systems such as yeast and *E. coli* are commonly employed to test hypothetical gene functions that have been inferred from sequence homology (Minet, et al. 1992; Sweasy and Loeb 1993; Perkins, et al. 1999; Osborn and Miller 2007; Hamza, et al. 2015). An implicit assumption of such approaches is that gene function is largely portable and interchangeable even when donor species come from radically different parts of the tree of life. Conversely, failure of such experiments may reflect incompatibilities between a donor gene and the recipient species (Dick and Trumpower 1998; diCenzo, et al. 2017).

More targeted studies have also directly tested for genetic incompatibilities by generating chimeric enzyme complexes with subunits derived from two different species or complexes with an altered mix of paralogous subunits (Kanevski, et al. 1999; Kim, et al. 2009; Abdel-Ghany, et al. 2022). Likewise, cytoplasmic hybrid (cybrid) experiments, in which the nuclear genome of one species must function with the cytoplasmic genomes of another species, have documented incompatibilities associated with divergence between lineages (Kenyon and Moraes 1997; Schmitz-Linneweber, et al. 2005). Overall, this array of comparative and experimental approaches has provided extensive examples of genetic incompatibility, which we will draw on in this review.

Functional interchangeability can be maintained across ancient timescales

The preceding section emphasized that genetic incompatibilities can have severe effects on molecular interactions and sometimes emerge over short timescales. However, comparisons across the tree of life have revealed contrasting examples, in which genes with core cellular functions have been exchanged across anciently divergent lineages and still retained their functions (Table 2). In addition, laboratory experiments have been able to reconstitute complex molecular machinery with components from diverse donor species (McClintock, et al. 2018). In this section, we overview some of the striking examples of interchangeability in molecular evolution.

As noted above, aaRS enzymes have undergone extensive HGT among all domains of life (Woese, et al. 2000). Such patterns of interchangeability are also observed in tRNAs themselves. Mitochondria inherited tRNA genes from their bacterial progenitor, and some eukaryotes have

retained a minimally complete set of these genes in the mitochondrial genome, but multiple lineages have lost many or all of them (Adams and Palmer 2003; Pett and Lavrov 2015; Salinas-Giegé, et al. 2015). There are no known cases in which these tRNA genes have been transferred to the nucleus and targeted back to the mitochondria. Instead, mitochondrial tRNA gene loss has been accompanied by the import of the nuclear-encoded tRNAs that normally function in the cytosol, meaning bacterial-like tRNAs were replaced by their anciently divergent eukaryotic counterparts (Salinas-Giegé, et al. 2015; Warren, et al. 2021).

In other cases, the establishment and integration of endosymbiotic bacteria and organelles into eukaryotic hosts cells has depended on gene transfer to the nucleus. Surprisingly, however, many such transfers have not come directly from the endosymbiont but instead originated from other bacterial donors, suggesting replacement of machinery originally contributed by the endosymbiont. For example, peptidoglycan is one of the defining features of the bacterial cell wall, and peptidoglycan biosynthesis in some plastids and endosymbiotic bacteria is now controlled by nuclear genes. But phylogenetic analyses have traced these peptidoglycan biosynthesis genes to disparate bacterial lineages (Husnik, et al. 2013; Sato and Takano 2017; Dowson, et al. 2022), meaning that the native enzymes originally present in the endosymbionts have been functionally replaced by homologs from entirely different phyla. Such examples support the broader argument that establishment of endosymbiotic relationships may often involve a series of multiple relationships that leave genetic footprints (Larkum, et al. 2007; Bennett and Moran 2015; Gray 2015).

The history of interchangeability in molecular evolution also extends to arguably the most fundamental processes of life – the replication and transcription of nucleic acids. For example, the DNA polymerase responsible for replication of mitochondrial DNA in animals, fungi, and other opisthokonts is not bacterial-like, contrary to what might be expected given the origins of mitochondria. Instead, the ancestral DNA polymerase has been functionally replaced by a viral-like polymerase; likewise, all eukaryotes appear to use viral-like machinery for helicase activity and transcription in their mitochondria (Shutt and Gray 2006), and the plastid genome of the cryptophyte *Rhodomonas salina* CCMP1319 was found to have acquired a gene encoding a putative DNA polymerase subunit from an unrelated bacterial lineage (Khan, et al. 2007).

The foregoing examples highlight the widespread history of functional replacement between homologous genes across the tree of life (Creevey, et al. 2011; Nagies, et al. 2020). However, in even more extreme cases, native machinery can be replaced by a non-homologous molecular system that plays a similar functional role (Table 2). Such replacements are possible because many enzymes that catalyze the same reaction have evolved independently (e.g., the multiple structurally distinct superoxide dismutases distributed across the tree of life) (Omelchenko, et al. 2010; Sutherland, et al. 2021).

A striking example of non-homologous replacement involves the key roles of mitochondria in production of iron-sulfur clusters, which are so essential that parasitic eukaryotes that lose the ability to generate ATP through cellular respiration still retain mitochondrion-related organelles to perform this function (Tovar, et al. 2003). The only known exception is the oxymonad *Monocercomonoides*, which appears to have lost mitochondria entirely. This loss was likely facilitated by HGT and the acquisition of a bacterial-like sulfur mobilization system (SUF) system as a non-homologous alternative to produce iron-sulfur clusters (Karnkowska, et al. 2016).

Above, we highlighted tRNAs and aaRSs as striking examples of homologous functional replacement. However, lysine aaRSs have also been involved in non-homologous replacement events. Lysine is the only aaRS with representatives in both of the (evolutionarily unrelated) Class I and Class II families, and these two alternative forms have undergone numerous functional replacement via HGT (Shaul, et al. 2006). The enzyme responsible for processing the 5′ ends of tRNAs (RNase P) provides another example of interchangeability in tRNA metabolism, involving machinery that is functionally analogous but non-homologous. The discovery that the catalytic activity of RNase P was conferred by an RNA and not a protein was a groundbreaking advance in the history of molecular biology, illustrating that RNAs can have enzymatic activity (ribozymes) (Guerrier-Takada, et al. 1983). As such, it came as a great surprise when it was later shown that RNase P activity in plant and animal mitochondria is mediated by a protein-only enzyme (Holzmann, et al. 2008; Gobert, et al. 2010). It has since become clear that both the ribozyme and protein-only versions of RNase P were ancestrally present in eukaryotes, and the subsequent history of differential gene retention and loss across lineages has determined which of these interchangeable versions now plays the functional role in tRNA processing (Lechner, et al. 2015).

A striking case of interchangeable but non-homologous machinery arises from the challenge of maintaining telomeres at the linear ends of chromosomes. Most eukaryotes extend their telomeres using the ribonucleoprotein telomerase complex, which relies on reverse transcription of a non-coding RNA to synthesize telomeric DNA (Podlevsky and Chen 2016). However, in several lineages, this function has been replaced by alternative mechanisms. For example, in *Drosophila*, telomeres are extended via a transposon-mediated system (Biessmann, et al. 1990; Levis, et al. 1993; Louis 2002) and a similar transition from telomerase-mediated to transposon-mediated telomere maintenance appears to have evolved independently multiple times in insects (Fujiwara, et al. 2005; Mason, et al. 2016). Mosquitos use yet another mechanism – one based on recombination – to extend telomeres (Roth, et al. 1997). In addition, yeast lacking functional telomerase as well as certain human cancer lines have also been shown to perform recombination-mediated telomere elongation (Lundblad 2002; van Mourik, et al. 2016; Zhang and Zou 2020), and *Myotis* bats also appear to use an alternative to the standard telomerase mechanism (Foley, et al. 2018). Collectively, such examples illustrate the incredible extent to which evolution has produced alternative systems to

solve the same problems and how such systems can sometimes be transferred across disparate branches in the tree of life.

Genetic principles that determine balance between incompatibility and interchangeability

How is it that some molecular systems rapidly evolve genetic incompatibilities while others remain interchangeable over deep evolutionary timescales? The answer to this question is undoubtedly complex and multifaceted, but below we point to four hypothesized genetic features that may contribute to where molecular systems fall on the incompatibility-interchangeability spectrum (Figure 2).

1. Multisubunit complexes and extent of protein-protein interactions. The "complexity hypothesis" and derivations thereof have suggested that interactions within stable multisubunit complexes as well as more transient protein interactions represent barriers to functional replacement (Jain, et al. 1999). There is extensive evidence that interacting proteins coevolve (Clark and Aquadro 2010; de Juan, et al. 2013; Forsythe, et al. 2021; Neverov, et al. 2021). Accordingly, disruption of these coevolved relationships through hybridization or HGT has the potential to produce incompatibilities (Swamy, et al. 2021). This concept has been supported by a number of systematic and genome-wide tests, most of which have identified a negative relationship between a gene's number of protein-protein interactions and its propensity to undergo HGT (Jain, et al. 1999; Sorek, et al. 2007; Wellner, et al. 2007; Lercher and Pál 2008; Creevey, et al. 2011; Acar Kirit, et al. 2020; Burch, et al. 2022).

The ribosome is probably the most extensively documented example of a molecular system that is recalcitrant to functional replacement events. Because this massive, multisubunit enzyme complex appears to be largely (although not entirely) resistant to HGT, ribosomal gene trees are generally viewed as representative of species relationships even at deep phylogenetic scales (Adams, et al. 2002; Ciccarelli, et al. 2006; Creevey, et al. 2011; Burch, et al. 2022). In addition, the diverse range of interactions within the ribosome have facilitated more nuanced analyses. For example, ribosomal protein subunits with larger amounts of surface area in contact with ribosomal RNAs are more likely to produce incompatibilities (Sorek, et al. 2007). Therefore, the intimacy and not just the quantity of molecular interactions is likely important in restricting interchangeability.

Another set of multisubunit complexes that have long been predicted to be a source of incompatibilities even over short timescales of divergence are the OXPHOS enzymes found in mitochondria (Rand, et al. 2004; Burton and Barreto 2012; Hill 2016). This prediction arises from the following line of argument: 1) OXPHOS complexes are generally composed of both mitochondrial- and nuclear-encoded subunits, 2) mitochondrial genomes experience higher mutation rates and more rapid sequence evolution than in the nucleus in many eukaryotes, and 3) nuclear genes may

experience selection for coevolutionary responses to changes in interacting mitochondrial genes, resulting in co-adapted mitonuclear genotypes that are sensitive to disruption by hybridization. Analyses of evolutionary rates and signatures of selection have found indirect evidence of coevolution between mitochondrial- and nuclear-encoded subunits in these complexes (Osada and Akashi 2012; Havird, et al. 2015; Neverov, et al. 2021), and a number of nuclear-encoded proteins that function in other aspects of mitochondrial biology have been implicated in BDMIs (Table 1) (Sloan, et al. 2017; Bozdag and Ono 2022). However, specific examples of incompatibilities arising from interactions within OXPHOS complexes have remained somewhat limited (Burton 2022). Some of the most direct evidence with experimental support has come from examples of disrupted function in mitonuclear OXPHOS complexes in marine copepod hybrids (Ellison and Burton 2006; Harrison and Burton 2006) and the recently identified example of a lethal interaction within OXPHOS Complex I in hybrid swordtail fish (Moran, et al. 2021). As the tools to pinpoint such incompatibilities improve, it should become clear whether these examples are generalizable.

In some cases, the coevolved interactions among subunits within enzyme complexes may be discriminating enough to preclude any opportunity for functional replacement by horizontally transferred homologs. For example, the bacterial acetyl-CoA carboxylase (ACCase) enzyme consists of multiple subunits and catalyzes the conversion of acetyl-CoA to malonyl-CoA, a key early step in fatty acid biosynthesis (Salie and Thelen 2016). Experimentally transferring genes encoding one of the ACCase subunits from divergent bacterial donors into *E. coli*, which encodes its own native copies of these subunits, had negligible effects on measured growth rates; however, the reason for these limited fitness consequences appeared to be that the foreign subunits were too divergent to even assemble or interact with the native subunits at all (Wellner and Gophna 2008). Thus, there does not appear to be any potential to functionally replace the native gene with one of these foreign copies.

Protein-protein interactions and multisubunit complexes are thought to represent a barrier to functional replacement because preservation of coevolved interactions in these cases would necessitate simultaneous exchange and subsequent retention of multiple genes. Even though such multi-gene replacements may occur (Waller, et al. 2006; Monier, et al. 2009; Karnkowska, et al. 2016), they are expected to be less probable than single-gene replacements (Keeling and Palmer 2008), which may explain some observed patterns of asymmetry in interchangeability. For example, plants typically have two distinct ACCase enzymes: 1) a typical eukaryotic multi-domain homomeric ACCase that is encoded by a single gene and functions in the cytosol and 2) an endosymbiotically acquired bacterial-like heteromeric ACCase that consists of four different subunits and functions in the plastids. However, in multiple independent angiosperm lineages, the homomeric ACCase has been duplicated and now functions in both the cytosol and the plastids, in some cases leading to the loss of the heteromeric complex altogether (Konishi, et al. 1996; Parker, et al. 2014; Park, et al.

2017; Williams, et al. 2022). In contrast, the subunits of the heteromeric ACCase have not been found to be duplicated and retargeted to the cytosol. Similarly, mitochondria use one of two different systems to perform heme attachment as part of cytochrome c maturation. Many eukaryotes retain the ancestral bacterial-like enzyme, which consists of subunits encoded by six or more genes; however, this heteromeric complex has been lost and replaced by a single-gene system (the holocytochrome c synthase or HCCS) many times throughout eukaryotic evolution (Babbitt, et al. 2015), a process which has likely included a history of HGT among eukaryotes (Allen, et al. 2008). These recurring histories of replacement supports the notion that transitions from multi-gene to single-gene systems are easier than the reverse process.

The history of functional replacement of mitochondrial aaRSs by their cytosolic counterparts also provides evidence for limitations imposed by multisubunit complexes in these replacement events. As described above, many lineages have lost some or all of their bacterial-like mitochondrial tRNA genes in favor of importing eukaryotic-like (nuclear) tRNAs from the cytosol (Salinas-Giegé, et al. 2015). In such cases, it is common for the corresponding mitochondrial aaRSs to also be lost and replaced by retargeted cytosolic aaRSs, preserving the ancestral aaRS-tRNA charging relationship. However, the most notable and consistent exception to this appears to be the cytosolic phenylalanine aaRS. This enzyme is the only of the cytosolic aaRSs to be expressed as two different subunits, which likely hinders retargeting and functional replacement of its mitochondrial aaRS counterpart (Pett and Lavrov 2015; Warren and Sloan 2022). Therefore, in cases of mitochondrial tRNA-Phe loss, the native mitochondrial phenylalanine aaRS is retained and presumably must adapt to charge the newly imported cytosolic tRNA.

The idea that aaRSs could readily evolve to charge a novel tRNA substrate (see above) or undergo HGT across divergent lineages that span the tree of life (Woese, et al. 2000) may seem surprising given the need for faithful aaRS-tRNA recognition in translation, but such evolutionary events may reinforce the hypothesized effects of molecular interactions in functional replacement. Accurate tRNA charging is generally achieved through the interaction between just two molecular components (the tRNA and the aaRS), and this interaction itself relies on a very small number of "identity elements" within the tRNA (Giegé, et al. 1998). As such, the limited scope of molecular interactions may make aaRSs a relatively "modular" enzyme class and, thus, explain why they seem so amenable to HGT and functional replacement. The contrasting histories of plant and animal mitochondrial tRNAs offer some support for this interpretation. Plant mitochondrial tRNA genes have shown an extensive history of interchangeability and functional replacement (Small, et al. 1999; Warren and Sloan 2020), which may indicate that the slow rate of sequence evolution in these genomes (Wolfe, et al. 1987) has led to conserved tRNA sequences and structures that retain similarities with other translation systems. In contrast, animal mitochondrial tRNAs often have highly divergent sequences and non-canonical structures (Watanabe 2010; Salinas-Giegé, et al. 2015;

316 Warren and Sloan 2021), which may have resulted in highly coevolved and “locked in” relationships
317 with their dedicated aaRSs. The very specific but limited basis of tRNA recognition may also help
318 resolve the apparent paradox that we highlighted in the Introduction. Whereas interchangeability
319 may be maintained as long as the key tRNA identity elements are present, even small changes in
320 sequence could lead to severe effects if they happen to disrupt this basis of recognition (Giegé, et al.
321 1998; Meiklejohn, et al. 2013).
322 The hypothesis that functional replacement is more likely to occur for proteins with limited
323 molecular interactions is also supported by examples such as the extensive HGT in the
324 peptidoglycan biosynthesis pathway for endosymbiotic bacteria/organelles (Husnik, et al. 2013; Sato
325 and Takano 2017; Dowson, et al. 2022). The enzymes in this pathway catalyze individual reactions
326 in series and do not assemble into large multisubunit complexes (Lovering, et al. 2012). Likewise,
327 the enzymes that act sequentially in the glycolysis pathway of eukaryotes are of
328 endosymbiotic/bacterial origin and replaced the ancestral host machinery (Bártulos, et al. 2018).
329 More generally, the complexity hypothesis was initially conceived based on observations that
330 “operational genes” (i.e., those involved in metabolic and housekeeping functions) are more likely to
331 undergo HGT and less likely to be involved in extensive protein-protein interactions (Jain, et al.
332 1999). As we have described in this section, subsequent studies in the last two decades have
333 produced growing evidence that multisubunit complexes and protein-protein interactions can
334 accelerate the accumulation of genetic incompatibilities and, thus, limit interchangeability.
335
336 *2. Sensitivity to changes in gene dosage*. Genes that are sensitive to changes in dosage (i.e., gene
337 copy number and/or expression level) are often toxic when experimentally introduced into a host
338 (Sorek, et al. 2007; Acar Kirit, et al. 2020). As such, dosage sensitivity may be a natural barrier to
339 functional replacement because such replacements can entail a period of redundancy between
340 native and foreign gene copies and, thus, changes in total expression level. Even in cases where
341 direct homologous replacements have been engineered, expression levels can change with
342 detrimental effects on fitness (Lind, et al. 2010; Bershtein, et al. 2015). Dosage sensitivity is a
343 widespread biological phenomenon and has been linked to the concept of gene “balance” (Papp, et
344 al. 2003). Specifically, shifts in gene copy number or expression levels may disrupt molecular
345 interactions that most occur at specific stoichiometric ratios. This phenomenon is thought to explain
346 why whole-genome duplication (polyploidy) is often better tolerated than partial-genome duplication
347 (aneuploidy) in many eukaryotes because the former generally maintains the same ratio of gene
348 copy numbers, whereas the latter perturbs these ratios (Birchler and Veitia 2012).
349 One prediction arising from this dosage hypothesis is that genes that exhibit frequent
350 functional replacement events can also readily be found in transitional states in which both copies
351 are functional, implying that dosage effects of expressing two copies are not prohibitively costly. For

example, as described above, the plastid heteromeric ACCase has been replaced in some taxa by importing the homomeric cytosolic ACCase, and species with both versions functioning in the plastid simultaneously have also been identified (Konishi, et al. 1996; Parker, et al. 2014; Park, et al. 2017; Williams, et al. 2022). Similarly, functional replacement of mitochondrial tRNAs by import of their cytosolic counterparts has been a common theme in eukaryotic evolution (Salinas-Giegé, et al. 2015), and this replacement process appears to involve a phase of functional redundancy in which both types of tRNAs are simultaneously present in the mitochondria (Warren, et al. 2021). More generally, this dosage hypothesis is supported by findings from genomic comparisons that genes that are preferentially maintained as single copy tend to be more resistant to HGT (Sorek, et al. 2007).

Dosage effects may also apply to nonhomologous replacement. For example, it has been hypothesized that maintaining two distinct siderophore biosynthesis pathways (desferrioxamine or salinichelin) in *Salinispora* bacteria is harmful, explaining why the two pathways are never found in the same strain (Bruns, et al. 2018). It is unclear whether such a cost is mediated by dosage effects, but it at least indicates any selective advantages from higher dosage and expression of two distinct pathways are insufficient to select for retention of both pathways. In this case, however, any barriers imposed by harmful redundancy have not (fully) prevented functional replacement, because multiple independent replacement events have been observed for these siderophore pathways.

Overall, these lines of evidence indicate that dosage sensitivity is a significant contributor to incompatibilities. As such, it is not just the nature of physical interactions that limits interchangeability but also the balance associated with levels of gene expression.

3. Evolutionary rate. Genes can evolve at remarkably different rates due to variation in the strength and efficacy of selection, the balance between positive and purifying selection, and differences in the underlying mutation rate (Bromham 2009). Because sequence divergence is expected to drive the accumulation of genetic incompatibilities (Presgraves 2010), genes with faster evolutionary rates may be less interchangeable. This hypothesis is supported by observations that the level of sequence divergence between taxa is negatively correlated with frequencies of HGT (Popa, et al. 2011; Skippington and Ragan 2012; Williams, et al. 2012; Slomka, et al. 2020) and the ability of genes to functionally replace their homologs (Lind, et al. 2010; Kacar, et al. 2017). However, the overall level of sequence divergence confounds differences in divergence time with the effects of variation in evolutionary rate per se. Burch, et al. (2022) recently differentiated between these effects by comparing the transferability of orthologous genes from the same pairs of donor and recipient bacterial species. As such, divergence time is held constant so any differences in sequence divergence can be attributed to variation in evolutionary rates. This analysis found that genes with

high rates of sequence divergence were indeed less amenable to HGT and that this relationship is stronger for genes involved in large numbers of protein-protein interactions.

In eukaryotes, cytonuclear interactions have been particularly useful in testing for rate effects because there are often systematic differences in evolutionary rates between the mitochondrial (or plastid) genome and the nucleus (Wolfe, et al. 1987). For example, animal mitochondrial genomes often evolve substantially faster than the nuclear genome; thus, the accumulation of mitochondrial changes has been predicted to drive the coevolutionary process and select for compensatory responses in nuclear-encoded proteins that are targeted to the mitochondria (Rand, et al. 2004; Burton, et al. 2013). Osada and Akashi (2012) tested for this predicted asymmetry using primate sequence data for proteins in the mitochondrial cytochrome c oxidase complex, showing that substitutions in mitochondrial-encoded subunits tended to precede substitutions at nearby sites in nuclear-encoded subunits. This apparent selection for compensatory or coevolutionary changes is one explanation for the observation that proteins targeted to the mitochondria often evolve faster than other nuclear-encoded proteins (Barreto and Burton 2013). Taxa in which the rate of mitochondrial or plastid sequence evolution show large variation among closely related species have been especially useful for tests of these coevolutionary principles. Such tests have found strong correlations between evolutionary rates of cytoplasmic genomes and interacting nuclear-encoded proteins (Zhang, et al. 2015; Weng, et al. 2016; Havird, et al. 2017; Yan, et al. 2019; Forsythe, et al. 2021).

Although accelerated rates and coevolutionary signatures from comparative genomic studies are often assumed to be associated with a faster buildup of incompatibilities between divergent taxa, direct functional tests of this assumption have been rare. Nonetheless, some more targeted functional studies have engineered chimeric enzyme complexes or interaction networks by substituting in genes from donor species with varying levels of sequence divergence (Asai, et al. 1999; Lind, et al. 2010; Bershtein, et al. 2015; Kacar, et al. 2017). For example, Kanevski, et al. (1999) engineered a rubisco enzyme complex in tobacco consisting of the native nuclear-encoded small subunit and a plastid-encoded large subunit that had been transferred from sunflower. This chimeric enzyme was able to successfully maintain partial rubisco functionality. However, the same was not true for attempts using a large subunit gene from a more distant (cyanobacterial) donor, supporting the expectation that the age of divergence between donor and recipient lineages contributes to accumulation of genetic incompatibilities. More recently, experiments used flowering plants that differed dramatically in their historical rates of sequence evolution for the plastid-encoded ClpP1 protein as donors to replace the native tobacco copy in another plastid-nuclear enzyme complex (the caseinolytic protease), finding that a history of accelerated sequence divergence hindered functional replacement (Abdel-Ghany, et al. 2022). By using donors from the same genus

(*Silene*), this experiment controlled for divergence time, isolating effects of evolutionary rate variation.

While cytonuclear interactions have been valuable in testing and teasing apart effects of evolutionary rate, such effects are also expected to pertain to nuclear-nuclear interactions. For example, the *PRDM9* gene is the best characterized example of a locus contributing to reproductive incompatibilities in mammals, and it undergoes unusually fast rates of sequence evolution (Mihola, et al. 2009; Oliver, et al. 2009). This gene is involved in determining hotspots for meiotic recombination by recognizing specific DNA sequence motifs, and its rapid evolution may reflect perpetual selection to recognize new motifs to counterbalance the predicted depletion of existing hotspots through recombinational mechanisms (Ponting 2011; Paigen and Petkov 2018). More generally, the antagonistic coevolution that is often associated with genomic conflict can often lead to rapid rates of sequence evolution, which may explain why genes involved in such conflict are often involved in BDMIs and reproductive isolation (Johnson 2010; Crespi and Nosil 2013; Sankararaman, et al. 2014; Serrato-Capuchina and Matute 2018; Postel and Touzet 2020; Schluter and Rieseberg 2022). Therefore, differences in rates of sequence evolution appear to affect the balance between incompatibility and interchangeability in disparate evolutionary lineages.

4. Overall functional importance. Perhaps the simplest and most intuitive hypothesis to explain observed variation in interchangeability is that the molecular systems that are especially important to cell viability and sensitive to disruption may be the most resistant to functional replacement. The rationale would be that the process of functional replacement inevitably involves some degree of perturbation to molecular systems, which would create more severe “fitness valleys” when they affect highly important genes. There is clear evidence that introduction of foreign genes and other forms of functional replacement can be disruptive through changes in protein homeostasis, increased cytotoxicity, and inefficient gene expression (Park and Zhang 2012; Baltrus 2013; Bershtein, et al. 2015; Bedhomme, et al. 2019). Even though subsequent evolution can lead to “amelioration” of such effects (Lawrence and Ochman 1997), the immediate harmful consequences may present too great a barrier to overcome for long-term functional replacement to occur, especially in the most constrained molecular systems.

Multiple observations support the hypothesis that functionally constrained genes are more resistant to replacement. For example, highly expressed genes are generally more conserved and have been shown to be less likely to undergo HGT (Park and Zhang 2012). In these cases, the barriers imposed by high expression may be associated with cytotoxic effects of inefficient translation and protein misfolding (Drummond, et al. 2005; Zhang and Yang 2015).

Many of the core components of molecular biology were present in the common ancestor of all extant cellular organisms and are near-universally conserved across the tree of life. Such

458 systems are likely among the most important to cell function, and many of these appear to undergo
459 lower rates of HGT and functional replacement than the rest of the genome (Jain, et al. 1999;
460 Fournier and Gogarten 2010; Koonin 2016). Indeed, the genealogical histories of proteins such as
461 elongation factors G and Tu, RNA polymerase β chain, DNA polymerase III, signal recognition
462 particle protein, and many ribosomal proteins closely resemble the structure of the tree of life with
463 little history of reticulation (Brown, et al. 2002).
464 A more direct measure of a gene's functional importance is the fitness effects associated
465 with mutating it or knocking it out. At the extreme, many genes are considered essential because
466 disrupting their function results in lethality (Glass, et al. 2006; Wang, et al. 2015). As noted above,
467 proteins that have extensive molecular interactions are more resistant to functional replacement.
468 Under what is known as the centrality-lethality rule, these genes that encode highly interacting
469 proteins are also more likely to be essential (Jeong, et al. 2001; Hahn and Kern 2005; Wellner, et al.
470 2007; Zotenko, et al. 2008). The relatively rare cases where functional replacement of these
471 essential molecular systems does occur may also be informative. For example, turnover of some
472 core biochemical and molecular genetic machinery has been documented for mitochondria, plastids,
473 and other bacterial endosymbionts (Hess and Börner 1999; Adams, et al. 2002; Shutt and Gray
474 2006; Husnik, et al. 2013; Gray 2015). In all these cases, the history of endosymbiosis has likely
475 resulted in extreme bottlenecks and relaxation of selection pressures (McCutcheon and Moran
476 2012), which may have created a more permissive environment for functional replacement events
477 that would have otherwise been too harmful. In the extreme, genetic degeneration in endosymbionts
478 may be so severe that functional replacement events are not only tolerated but actually promoted by
479 selection as a form of genetic "rescue" (Bennett and Moran 2015).
480 Overall, these lines of evidence all point to a role of functional importance in determining the
481 balance between interchangeability and incompatibility.
482
483 **Open questions and future directions**
484
485 In this concluding section, we point to five areas where there may be opportunities to build on recent
486 progress in our understanding of evolutionary forces that shape the process of functional
487 replacement.
488
489 *Multifunctional proteins: the role of pleiotropy in evolution of incompatibilities.* One intuitive prediction
490 is that genes that have multiple functions and affect multiple phenotypes (i.e., pleiotropy) will be
491 more prone to genetic incompatibilities. However, to our knowledge, the relationship between
492 pleiotropy and a gene's amenability to functional replacement has not been directly tested. It has
493 long been suspected that pleiotropy could act as a constraint on evolution (Fisher 1930; Orr 2000;

Ngo, et al. 2022). There is evidence that pleiotropic genes occupy central positions in protein-protein interaction networks (Promislow 2004). As we have discussed, such interactions are expected to directly affect a gene's interchangeability. In addition, genes with extensive protein-protein interactions also exhibit slower sequence evolution (Fraser 2005; Hahn and Kern 2005; Ngo, et al. 2022) and more constrained gene expression (Lemos, et al. 2004; Papakostas, et al. 2014), which may also affect interchangeability. Likewise, pleiotropic genes appear to have more substantial phenotypic effects even when measured on a per-trait basis (Wang, et al. 2010). These patterns suggest that pleiotropy will affect the rate at which genetic incompatibilities arise, and with the establishment of genotype-phenotype maps on genome-wide scales (Wagner and Zhang 2011), resources are increasingly available to test for such an effect.

Decoupling confounded variables: separating correlated genetic features and the phylogenetic distribution of donor genes. Many of the genetic features we have discussed are not independent of each other, resulting in confounding effects that are difficult to disentangle. For example, as noted above, the functional importance of genes is associated with their degree of integration into protein-protein interaction networks (Jeong, et al. 2001; Wellner, et al. 2007; Zotenko, et al. 2008). In other cases, contributing factors are negatively correlated (e.g., functional importance and evolutionary rate) and may mask each other's effects. Although some attempts have been made to distinguish the contributions of correlated variables (Cohen, et al. 2011; Burch, et al. 2022), separating such effects remains a pressing challenge. For example, we hypothesize that genes that are widespread across the tree of life would have a higher chance of functional replacement given the ample supply of potential donors. However, at face value, the available data do not appear to support this hypothesis, as the most anciently conserved and widely distributed genes exhibit *less* HGT (Jain, et al. 1999; Brown, et al. 2002; Fournier and Gogarten 2010; Koonin 2016). However, this clearly remains an open question, as it is possible that donor availability positively contributes to the probability of replacement once the confounded effects of functional importance are controlled for. More generally, addressing the challenge of correlated features may require experimental manipulations to complement existing comparative and statistical approaches. For example, altering environmental conditions or modifying gene regulatory systems could be means to control gene expression levels during environmental transfers.

Beyond E. coli: expanding the taxonomic scope of experimental interchangeability studies. Functional wet-lab analyses have provided a key complement to comparative-genomic and phylogenetic approaches in understanding the mechanisms of molecular incompatibility and interchangeability. Most of these groundbreaking studies have relied on the power of *E. coli* as a model system for high-throughput transgenic analyses to systematically screen the effects of gene

transfer and functional replacement (Asai, et al. 1999; Sorek, et al. 2007; Bershtein, et al. 2015; Kacar, et al. 2017; Acar Kirit, et al. 2020). However, there are many reasons to expect that the principles dictating the outcome of functional replacement may depend on the recipient genome and cellular environment. With the growing resources available for engineering the genomes of yeast and multicellular eukaryotes, there are exciting prospects to expand this field of functional studies beyond *E. coli*.

Retracing the steps: use of ancestral protein reconstructions in functional assays. A rapidly growing approach in the field of molecular evolution involves the use of phylogenetics to infer the sequence of ancestral protein-coding genes, which can then be synthesized and expressed (Hochberg and Thornton 2017). Such reconstructed ancestral proteins can then be used for functional assays both *in vitro* and *in vivo* (Smith, et al. 2013; Kacar, et al. 2017; Hochberg, et al. 2020). This approach addresses a fundamental limitation of conventional molecular incompatibility/interchangeability studies, which are typically restricted to analysis of extant proteins. Instead, inclusion of ancestral proteins presents the exciting opportunity to recreate the order and timing of the step-wise evolutionary process by which incompatibilities emerge and to determine how this evolutionary process plays out on complex epistatic fitness landscapes.

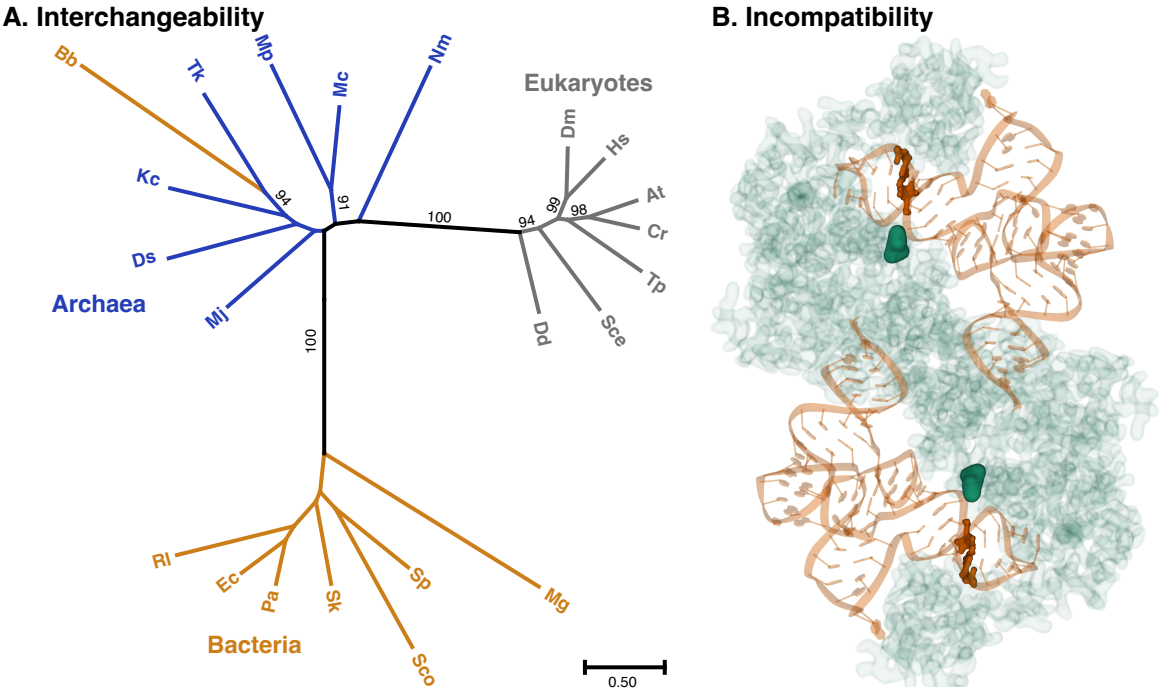
Experimental evolution: capturing the functional replacement process on laboratory timescales. An exciting recent development is the increasing use of experimentally evolved bacterial populations and whole-genome sequencing to track the effects of HGT across generations in the lab (Chu, et al. 2018; Slomka, et al. 2020; Woods, et al. 2020; Power, et al. 2021; Nguyen, et al. 2022). These studies grow bacterial populations in the presence of various sources of donor DNA in the media or allow bacteria to evolve with other strains and potentially exchange DNA. As such, the outcomes of genetic exchange and functional replacements can be directly assessed under more realistic conditions of population growth and competition. Such approaches should create the opportunity to strategically manipulate donor and recipient genomes to further develop and test hypotheses about genetic features that affect the balance between incompatibility and interchangeability in molecular evolution.

ACKNOWLEDGEMENTS

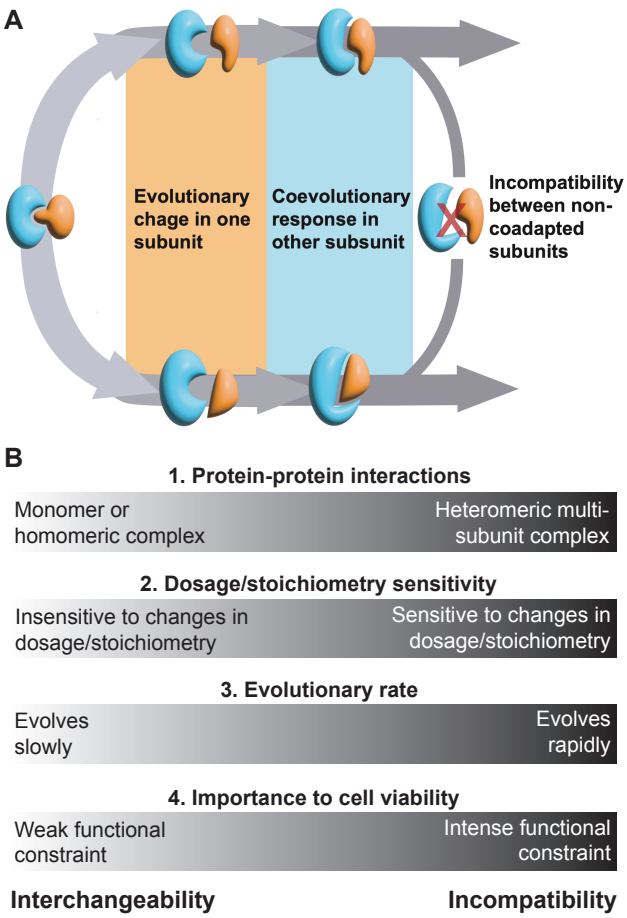
We thank Charleston Ducote for contributions to the design of Figure 2A. Our work on molecular coevolution and mutation is supported by the National Science Foundation (NSF; IOS-2114641 and MCB-2048407) and the National Institutes of Health (NIH; R01 GM118046). SAK is supported by a predoctoral training fellowships from NIH (T32 GM132057) and the NSF Graduate Research

Fellowship Program. JMW is supported by a postdoctoral fellowship from the Howard Hughes Medical Institute Hanna H. Gray Fellows Program.

Figure 1. The paradox of interchangeability and incompatibility illustrated with aaRS genes: A) An example of interchangeability between anciently divergent copies of phenylalanine aaRS via HGT from archaea to the bacterial lineage that includes spirochaetes, represented here by *Borrelia burgdorferi* (Bb) (Woese, et al. 2000). Amino-acid sequences for phenylalanine aaRS orthologs were recovered with SHOOT (Emms and Kelly 2022) using *B. burgdorferi* (ADQ30774) as a query sequence, aligned with MAFFT (Katoh and Standley 2013), and used for maximum-likelihood phylogenetic inference with IQ-TREE (Minh, et al. 2020). Bipartitions with greater than 90% support from ultrafast bootstrap pseudoreplicates are indicated. Aligned sequences with full taxon names are provided as supplemental material (File S1) . B) A contrasting example of aaRS-tRNA incompatibility based on only a single nucleotide substitution in the tRNA and a single amino-acid substitution in the aaRS. The structural model represents a tyrosine aaRS dimer (green) complexed with two tRNA-Tyr molecules (orange). The highlighted residues and base pairs indicate the positions that are homologous to sites where substitutions occurred in *Drosophila*, leading to an incompatibility (Meiklejohn, et al. 2013). The structural model is based on Protein Data Bank accession 1H3E (Yaremchuk, et al. 2002) and was visualized with Mol* (Sehnal, et al. 2021).



582 **Figure 2.** The origins of epistatic incompatibilities: A) Stylized representation of the coevolutionary
583 process leading to incompatibilities between isolated evolutionary lineages. Interacting subunits
584 (blue and orange) undergo evolutionary changes and coevolutionary responses, preserving
585 functional interactions within a lineage but leading to incompatibilities between subunits when
586 brought back together through hybridization or HGT. B) Summary of genetic principles that may
587 determine the balance between interchangeability and incompatibility in specific molecular systems.



588 **Table 1.** Examples of molecular genetic incompatibilities revealed by hybridization between recently
589 diverged lineage or by gene transfer (either natural or experimental) between more distantly related
590 taxa.

	Taxon	Description	Reference
Hybrid Incompatibility	*Drosophila*	tRNA-aaRS mitonuclear interaction	(Meiklejohn, et al. 2013)
	Drosophila	Lhr/Hmr heterochromatin interactions	(Brideau, et al. 2006)
	Xiphophorus	OXPHOS Complex I mitonuclear interaction	(Moran, et al. 2021)
	Mus	PRDM9 and recombination hotspots	(Mihola, et al. 2009)
	Homo	Testis-specific genes	(Sankararaman, et al. 2014)
	Saccharomyces	AEP2/OLI1 mitonuclear interaction	(Lee, et al. 2008)
	Oryza	S5 Proteases	(Chen, et al. 2008)
	Arabidopsis	NLR immune receptor genes	(Chae, et al. 2014)
Transfer Incompatibilities	Tree of Life	Ribosomal proteins	(Ciccarelli, et al. 2006; Sorek, et al. 2007)
	Angiosperms	Plastid Clp protease	(Abdel-Ghany, et al. 2022)
	Bacteria	ACCase	(Wellner and Gophna 2008)
	Sinorhizobium	BacA and plant nodulation coevolution	(diCenzo, et al. 2017)
	Bacteria	DNA replication machinery	(Jain, et al. 1999; Sorek, et al. 2007)
	Bacteria	Elongation Factor Tu	(Kacar, et al. 2017)
	Plants/Bacteria	Rubisco	(Kanevski, et al. 1999)
	Bacteria	Dihydrofolate reductase	(Bershtein, et al. 2015)

Table 2. Examples of interchangeability in molecular interactions including both homologous and non-homologous replacement events.

	Description	Reference
Homologous Replacement	Aminoacyl-tRNA synthetases (cellular tree of life)	(Woese, et al. 2000)
	Mitochondrial ribosomal proteins (cellular tree of life)	(Adams, et al. 2002)
	Mitochondrial tRNAs (cellular tree of life)	(Warren, et al. 2021)
	Endosymbiont peptidoglycan biosynthesis (bacteria)	(Husnik, et al. 2013)
	Plastid GAPDH (cellular tree of life)	(Keeling 2009)
	in vitro reconstitution of dynein motor complex (metazoans)	(McClintock, et al. 2018)
	Heteromeric and homomeric ACCase (cellular tree of life)	(Konishi, et al. 1996)
	Mitochondrial DNA polymerase (cellular-viral tree of life)	(Shutt and Gray 2006)
Non-homologous Replacement	SUF sulfur mobilization system	(Karnkowska, et al. 2016)
	Telomerase functions	Multiple (see text)
	Cytochrome c maturation	(Babbitt, et al. 2015)
	Siderophore biosynthesis	(Bruns, et al. 2018)
	Class I and II LysRS	(Shaul, et al. 2006)
	Ribozyme and Protein-Only Rnase P	(Lechner, et al. 2015)
	Fructose-6-phosphate aldolase (FBA)	(Patron, et al. 2004)
	Superoxide dismutase	(Sutherland, et al. 2021)

REFERENCES

Abdel-Ghany SE, LaManna LM, Harroun HT, Maliga P, Sloan DB. 2022. Rapid sequence evolution is associated with genetic incompatibilities in the plastid Clp complex. Plant Molecular Biology 108:277-287.

Acar Kirit H, Lagator M, Bollback JP. 2020. Experimental determination of evolutionary barriers to horizontal gene transfer. BMC Microbiology 20:326.

Adams KL, Daley DO, Whelan J, Palmer JD. 2002. Genes for two mitochondrial ribosomal proteins in flowering plants are derived from their chloroplast or cytosolic counterparts. The Plant Cell 14:931-943.

Adams KL, Palmer JD. 2003. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. Molecular phylogenetics and evolution 29:380-395.

Allen JW, Jackson AP, Rigden DJ, Willis AC, Ferguson SJ, Ginger ML. 2008. Order within a mosaic distribution of mitochondrial c-type cytochrome biogenesis systems? FEBS Journal 275:2385-2402.

Andam CP, Gogarten JP. 2011. Biased gene transfer in microbial evolution. Nature Reviews Microbiology 9:543-555.

Asai T, Zaporojets D, Squires C, Squires CL. 1999. An Escherichia coli strain with all chromosomal rRNA operons inactivated: complete exchange of rRNA genes between bacteria. Proceedings of the National Academy of Sciences 96:1971-1976.

Babbitt SE, Sutherland MC, San Francisco B, Mendez DL, Kranz RG. 2015. Mitochondrial cytochrome c biogenesis: no longer an enigma. Trends in biochemical sciences 40:446-455.

Baltrus DA. 2013. Exploring the costs of horizontal gene transfer. Trends in Ecology & Evolution 28:489-495.

Barreto FS, Burton RS. 2013. Evidence for compensatory evolution of ribosomal proteins in response to rapid divergence of mitochondrial rRNA. Molecular Biology and Evolution 30:310-314.

Bártulos CR, Rogers MB, Williams TA, Gentekaki E, Brinkmann H, Cerff R, Liaud M-F, Hehl AB, Yarlett NR, Gruber A. 2018. Mitochondrial glycolysis in a major lineage of eukaryotes. Genome Biology and Evolution 10:2310–2325.

Bedhomme S, Amorós-Moya D, Valero LM, Bonifaci N, Pujana M-À, Bravo IG. 2019. Evolutionary changes after translational challenges imposed by horizontal gene transfer. Genome Biology and Evolution 11:814-831.

621 Bennett GM, Moran NA. 2015. Heritable symbiosis: The advantages and perils of an evolutionary rabbit
622 hole. Proceedings of the National Academy of Sciences 112:10169-10176.

623 Bershtein S, Serohijos AW, Bhattacharyya S, Manhart M, Choi J-M, Mu W, Zhou J, Shakhnovich EI.
624 2015. Protein homeostasis imposes a barrier on functional integration of horizontally transferred genes in
625 bacteria. PLoS Genetics 11:e1005612.

626 Biessmann H, Carter SB, Mason JM. 1990. Chromosome ends in Drosophila without telomeric DNA
627 sequences. Proceedings of the National Academy of Sciences 87:1758-1761.

628 Birchler JA, Veitia RA. 2012. Gene balance hypothesis: connecting issues of dosage sensitivity across
629 biological disciplines. Proceedings of the National Academy of Sciences 109:14746-14753.

630 Bozdag GO, Ono J. 2022. Evolution and molecular bases of reproductive isolation. Current Opinion in
631 Genetics & Development 76:101952.

632 Brideau NJ, Flores HA, Wang J, Maheshwari S, Wang X, Barbash DA. 2006. Two Dobzhansky-Muller
633 genes interact to cause hybrid lethality in Drosophila. Science 314:1292-1295.

634 Bromham L. 2009. Why do species vary in their rate of molecular evolution? Biology Letters 5:401-404.

635 Brown JR, Italia MJ, Douady C, Stanhope MJ. 2002. Horizontal gene transfer and the universal tree of
636 life. In. Horizontal Gene Transfer: Elsevier. p. 305-349.

637 Bruns H, Crüsemann M, Letzel A-C, Alanjary M, McInerney JO, Jensen PR, Schulz S, Moore BS, Ziemert
638 N. 2018. Function-related replacement of bacterial siderophore pathways. ISME Journal 12:320-329.

639 Burch CL, Romanchuk A, Kelly M, Wu Y, Jones CD. 2022. Genome-wide determination of barriers to
640 horizontal gene transfer. bioRxiv:2022.2006.2029.498157.

641 Burton RS. 2022. The role of mitonuclear incompatibilities in allopatric speciation. Cellular and Molecular
642 Life Sciences 79:103.

643 Burton RS, Barreto FS. 2012. A disproportionate role for mtDNA in Dobzhansky-Muller incompatibilities?
644 Molecular ecology 21:4942-4957.

645 Burton RS, Pereira RJ, Barreto FS. 2013. Cytonuclear genomic interactions and hybrid breakdown.
646 Annual Review of Ecology, Evolution, and Systematics 44:281-302.

Chae E, Bomblies K, Kim S-T, Karelina D, Zaidem M, Ossowski S, Martín-Pizarro C, Laitinen RA, Rowan BA, Tenenboim H. 2014. Species-wide genetic incompatibility analysis identifies immune genes as hot spots of deleterious epistasis. Cell 159:1341-1351.

Chen J, Ding J, Ouyang Y, Du H, Yang J, Cheng K, Zhao J, Qiu S, Zhang X, Yao J. 2008. A triallelic system of S5 is a major regulator of the reproductive barrier and compatibility of indica–japonica hybrids in rice. Proceedings of the National Academy of Sciences 105:11436-11441.

Chu HY, Sprouffske K, Wagner A. 2018. Assessing the benefits of horizontal gene transfer by laboratory evolution and genome sequencing. BMC Evolutionary Biology 18:54.

Ciccarelli FD, Doerks T, Von Mering C, Creevey CJ, Snel B, Bork P. 2006. Toward automatic reconstruction of a highly resolved tree of life. Science 311:1283-1287.

Clark NL, Aquadro CF. 2010. A novel method to detect proteins evolving at correlated rates: identifying new functional relationships between coevolving proteins. Molecular Biology and Evolution 27:1152-1161.

Cohen O, Gophna U, Pupko T. 2011. The complexity hypothesis revisited: connectivity rather than function constitutes a barrier to horizontal gene transfer. Molecular Biology and Evolution 28:1481-1489.

Creevey CJ, Doerks T, Fitzpatrick DA, Raes J, Bork P. 2011. Universally distributed single-copy genes indicate a constant rate of horizontal transfer. PloS one 6:e22099.

Crespi B, Nosil P. 2013. Conflictual speciation: species formation via genomic conflict. Trends in Ecology & Evolution 28:48-57.

de Juan D, Pazos F, Valencia A. 2013. Emerging methods in protein co-evolution. Nature Reviews Genetics 14:249-261.

diCenzo GC, Zamani M, Ludwig HN, Finan TM. 2017. Heterologous complementation reveals a specialized activity for BacA in the Medicago–Sinorhizobium meliloti symbiosis. Molecular Plant-Microbe Interactions 30:312-324.

Dick FA, Trumpower BL. 1998. Heterologous complementation reveals that mutant alleles of QSR1 render 60S ribosomal subunits unstable and translationally inactive. Nucleic Acids Research 26:2442-2448.

Dowson AJ, Lloyd AJ, Cuming AC, Roper DI, Frigerio L, Dowson CG. 2022. Plant peptidoglycan precursor biosynthesis: Conservation between moss chloroplasts and Gram negative bacteria. bioRxiv.

675 Drummond DA, Bloom JD, Adami C, Wilke CO, Arnold FH. 2005. Why highly expressed proteins evolve
676 slowly. Proceedings of the National Academy of Sciences of the United States of America 102:14338-
677 14343.

678 Ellison CK, Burton RS. 2006. Disruption of mitochondrial function in interpopulation hybrids of *Tigriopus*
679 *californicus*. Evolution 60:1382-1391.

680 Emms DM, Kelly S. 2022. SHOOT: phylogenetic gene search and ortholog inference. Genome biology
681 23:85.

682 Eyre-Walker A, Keightley PD. 2007. The distribution of fitness effects of new mutations. Nature Reviews
683 Genetics 8:610-618.

684 Fisher RA. 1930. The Genetical Theory of Natural Selection. Oxford: Clarendon Press.

685 Foley NM, Hughes GM, Huang Z, Clarke M, Jebb D, Whelan CV, Petit EJ, Touzalin F, Farcy O, Jones G.
686 2018. Growing old, yet staying young: The role of telomeres in bats' exceptional longevity. Science
687 Advances 4:eaao0926.

688 Forsythe ES, Grover CE, Miller ER, Conover JL, Chavarro CF, Arick II MA, Peterson DG, Leal-Bertioli
689 SCM, Sharbrough J, Wendel JF, et al. 2022. Organellar transcripts dominate the cellular mRNA pool
690 across plants of varying ploidy levels. Proceedings of the National Academy of Sciences In Press.

691 Forsythe ES, Williams AM, Sloan DB. 2021. Genome-wide signatures of plastid-nuclear coevolution point
692 to repeated perturbations of plastid proteostasis systems across angiosperms. Plant Cell 33:980-997.

693 Fournier GP, Gogarten JP. 2010. Rooting the ribosomal tree of life. Molecular Biology and Evolution
694 27:1792-1801.

695 Fraser HB. 2005. Modularity and evolutionary constraint on proteins. Nature genetics 37:351-352.

696 Fujiwara H, Osanai M, Matsumoto T, Kojima KK. 2005. Telomere-specific non-LTR retrotransposons and
697 telomere maintenance in the silkworm, Bombyx mori. Chromosome research 13:455-467.

698 Giegé R, Sissler M, Florentz C. 1998. Universal rules and idiosyncratic features in tRNA identity. Nucleic
699 Acids Research 26:5017-5035.

700 Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR, Maruf M, Hutchison III CA, Smith HO,
701 Venter JC. 2006. Essential genes of a minimal bacterium. Proceedings of the National Academy of
702 Sciences 103:425-430.

703 Gobert A, Gutmann B, Taschner A, Gößringer M, Holzmann J, Hartmann RK, Rossmanith W, Giegé P.
704 2010. A single Arabidopsis organellar protein has RNase P activity. Nature structural & molecular biology
705 17:740-744.

706 Gray MW. 2015. Mosaic nature of the mitochondrial proteome: Implications for the origin and evolution of
707 mitochondria. Proceedings of the National Academy of Sciences 112:10133-10138.

708 Guerrier-Takada C, Gardiner K, Marsh T, Pace N, Altman S. 1983. The RNA moiety of ribonuclease P is
709 the catalytic subunit of the enzyme. Cell 35:849-857.

710 Hahn MW, Kern AD. 2005. Comparative genomics of centrality and essentiality in three eukaryotic
711 protein-interaction networks. Molecular Biology and Evolution 22:803-806.

712 Hamza A, Tammpere E, Kofoed M, Keong C, Chiang J, Giaever G, Nislow C, Hieter P. 2015.
713 Complementation of yeast genes with human genes as an experimental platform for functional testing of
714 human genetic variants. Genetics 201:1263-1274.

715 Harrison JS, Burton RS. 2006. Tracing hybrid incompatibilities to single amino acid substitutions.
716 Molecular Biology and Evolution 23:559-564.

717 Havird JC, Trapp P, Miller C, Bazos I, Sloan DB. 2017. Causes and consequences of rapidly evolving
718 mtDNA in a plant lineage. Genome Biology and Evolution 9:323-336.

719 Havird JC, Whitehill NS, Snow CD, Sloan DB. 2015. Conservative and compensatory evolution in
720 oxidative phosphorylation complexes of angiosperms with highly divergent rates of mitochondrial genome
721 evolution. Evolution 69:3069-3081.

722 Hess WR, Börner T. 1999. Organellar RNA polymerases of higher plants. International review of cytology
723 190:1-59.

724 Hill GE. 2016. Mitonuclear coevolution as the genesis of speciation and the mitochondrial DNA barcode
725 gap. Ecology and Evolution 6:5831-5842.

726 Hochberg GKA, Liu Y, Marklund EG, Metzger BP, Laganowsky A, Thornton JW. 2020. A hydrophobic
727 ratchet entrenches molecular complexes. Nature 588:503-508.

728 Hochberg GKA, Thornton JW. 2017. Reconstructing ancient proteins to understand the causes of
729 structure and function. Annual Review of Biophysics 46:247-269.

730 Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmanith W. 2008. RNase P without RNA:
731 identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. Cell
732 135:462-474.

733 Huang J, Yue J. 2013. Horizontal gene transfer in the evolution of photosynthetic eukaryotes. Journal of
734 Systematics and Evolution 51:13-29.

735 Husnik F, Nikoh N, Koga R, Ross L, Duncan RP, Fujie M, Tanaka M, Satoh N, Bachtrog D, Wilson ACC.
736 2013. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested
737 mealybug symbiosis. Cell 153:1567-1578.

738 Itaya M, Tsuge K, Koizumi M, Fujita K. 2005. Combining two genomes in one cell: stable cloning of the
739 Synechocystis PCC6803 genome in the Bacillus subtilis 168 genome. Proceedings of the National
740 Academy of Sciences 102:15971-15976.

741 Jain R, Rivera MC, Lake JA. 1999. Horizontal gene transfer among genomes: the complexity hypothesis.
742 Proceedings of the National Academy of Sciences 96:3801-3806.

743 Jeong H, Mason SP, Barabási A-L, Oltvai ZN. 2001. Lethality and centrality in protein networks. Nature
744 411:41-42.

745 Johnson NA. 2010. Hybrid incompatibility genes: remnants of a genomic battlefield? Trends in Genetics
746 26:317-325.

747 Kacar B, Garmendia E, Tuncbag N, Andersson DI, Hughes D. 2017. Functional constraints on replacing
748 an essential gene with its ancient and modern homologs. mBio 8:e01276-01217.

749 Kanevski I, Maliga P, Rhoades DF, Gutteridge S. 1999. Plastome engineering of ribulose-1, 5-
750 bisphosphate carboxylase/oxygenase in tobacco to form a sunflower large subunit and tobacco small
751 subunit hybrid. Plant Physiology 119:133-142.

752 Karnkowska A, Vacek V, Zubáčová Z, Treitli SC, Petrželková R, Eme L, Novák L, Žárský V, Barlow LD,
753 Herman EK. 2016. A eukaryote without a mitochondrial organelle. Current Biology 26:1274-1284.

754 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in
755 performance and usability. Molecular Biology and Evolution 30:772-780.

756 Keeling PJ. 2009. Chromalveolates and the evolution of plastids by secondary endosymbiosis 1. Journal
757 of Eukaryotic Microbiology 56:1-8.

Keeling PJ, Palmer JD. 2008. Horizontal gene transfer in eukaryotic evolution. Nature Reviews Genetics 9:605-618.

Kenyon L, Moraes CT. 1997. Expanding the functional human mitochondrial DNA database by the establishment of primate xenomitochondrial cybrids. Proceedings of the National Academy of Sciences of the United States of America 94:9131-9135.

Khan H, Parks N, Kozera C, Curtis BA, Parsons BJ, Bowman S, Archibald JM. 2007. Plastid genome sequence of the cryptophyte alga Rhodomonas salina CCMP1319: lateral transfer of putative DNA replication machinery and a test of chromist plastid phylogeny. Molecular Biology and Evolution 24:1832-1842.

Kim J, Rudella A, Rodriguez VR, Zybailov B, Olinares PDB, van Wijk KJ. 2009. Subunits of the plastid ClpPR protease complex have differential contributions to embryogenesis, plastid biogenesis, and plant development in Arabidopsis. Plant Cell 21:1669-1692.

Konishi T, Shinohara K, Yamada K, Sasaki Y. 1996. Acetyl-CoA carboxylase in higher plants: most plants other than gramineae have both the prokaryotic and the eukaryotic forms of this enzyme. Plant and Cell Physiology 37:117-122.

Koonin EV. 2016. Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transitions. F1000Research 5:1805.

Koonin EV, Makarova KS, Aravind L. 2001. Horizontal gene transfer in prokaryotes: quantification and classification. Annual Reviews in Microbiology 55:709-742.

Larkum AW, Lockhart PJ, Howe CJ. 2007. Shopping for plastids. Trends in plant science 12:189-195.

Lawrence JG, Ochman H. 1997. Amelioration of bacterial genomes: rates of change and exchange. Journal of Molecular Evolution 44:383-397.

Lechner M, Rossmanith W, Hartmann RK, Thölken C, Gutmann B, Giegé P, Gobert A. 2015. Distribution of ribonucleoprotein and protein-only RNase P in Eukarya. Molecular Biology and Evolution 32:3186-3193.

Lee HY, Chou JY, Cheong L, Chang NH, Yang SY, Leu JY. 2008. Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. Cell 135:1065-1073.

Lemos B, Meiklejohn CD, Hartl DL. 2004. Regulatory evolution across the protein interaction network. Nature genetics 36:1059-1060.

787 Lercher MJ, Pál C. 2008. Integration of horizontally transferred genes into regulatory interaction networks
788 takes many million years. Molecular Biology and Evolution 25:559-567.

789 Levis RW, Ganesan R, Houtchens K, Tolar LA, Sheen F-m. 1993. Transposons in place of telomeric
790 repeats at a Drosophila telomere. Cell 75:1083-1093.

791 Lind PA, Tobin C, Berg OG, Kurland CG, Andersson DI. 2010. Compensatory gene amplification restores
792 fitness after inter-species gene replacements. Molecular Microbiology 75:1078-1089.

793 Louis EJ. 2002. Are Drosophila telomeres an exception or the rule? Genome biology
794 3:reviews0007.0001.

795 Lovering AL, Safadi SS, Strynadka NC. 2012. Structural perspective of peptidoglycan biosynthesis and
796 assembly. Annual Review of Biochemistry 81:451-478.

797 Lundblad V. 2002. Telomere maintenance without telomerase. Oncogene 21:522-531.

798 Lynch M, Koskella B, Schaack S. 2006. Mutation pressure and the evolution of organelle genomic
799 architecture. Science 311:1727-1730.

800 Mason JM, Randall TA, Capkova Frydrychova R. 2016. Telomerase lost? Chromosoma 125:65-73.

801 McClintock MA, Dix CI, Johnson CM, McLaughlin SH, Maizels RJ, Hoang HT, Bullock SL. 2018. RNA-
802 directed activation of cytoplasmic dynein-1 in reconstituted transport RNPs. Elife 7:e36312.

803 McCutcheon JP, Moran NA. 2012. Extreme genome reduction in symbiotic bacteria. Nature Reviews
804 Microbiology 10:13-26.

805 Meiklejohn CD, Holmbeck MA, Siddiq MA, Abt DN, Rand DM, Montooth KL. 2013. An Incompatibility
806 between a mitochondrial tRNA and its nuclear-encoded tRNA synthetase compromises development and
807 fitness in Drosophila. PLoS Genetics 9:e1003238.

808 Mihola O, Trachtulec Z, Vlcek C, Schimenti JC, Forejt J. 2009. A mouse speciation gene encodes a
809 meiotic histone H3 methyltransferase. Science 323:373-375.

810 Minet M, Dufour ME, Lacroute F. 1992. Complementation of Saccharomyces cerevisiae auxotrophic
811 mutants by Arabidopsis thaliana cDNAs. Plant Journal 2:417-422.

812 Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R. 2020.
813 IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Molecular
814 Biology and Evolution 37:1530-1534.

Monier A, Pagarete A, de Vargas C, Allen MJ, Claverie J-M, Ogata H. 2009. Horizontal gene transfer of an entire metabolic pathway between a eukaryotic alga and its DNA virus. Genome Research 19:1441-1449.

Moran BM, Payne CY, Powell DL, Iverson EN, Banerjee SM, Langdon QK, Gunn TR, Liu F, Matney R, Singhal K. 2021. A lethal genetic incompatibility between naturally hybridizing species in Mitochondrial Complex I. bioRxiv.

Mukai T, Lajoie MJ, Englert M, Söll D. 2017. Rewriting the genetic code. Annual Review of Microbiology 71:557-577.

Nagies FS, Brueckner J, Tria FD, Martin WF. 2020. A spectrum of verticality across genes. PLoS Genetics 16:e1009200.

Neverov AD, Popova AV, Fedonin GG, Cheremukhin EA, Klink GV, Bazykin GA. 2021. Episodic evolution of coadapted sets of amino acid sites in mitochondrial proteins. PLoS Genetics 17:e1008711.

Ngo TM, Williams AM, Tate AT. 2022. The effect of developmental pleiotropy on the evolution of insect immune genes. bioRxiv.

Nguyen AN, Woods LC, Gorrell R, Ramanan S, Kwok T, McDonald MJ. 2022. Recombination resolves the cost of horizontal gene transfer in experimental populations of Helicobacter pylori. Proceedings of the National Academy of Sciences 119:e2119010119.

Oliver PL, Goodstadt L, Bayes JJ, Birtle Z, Roach KC, Phadnis N, Beatson SA, Lunter G, Malik HS, Ponting CP. 2009. Accelerated evolution of the Prdm9 speciation gene across diverse metazoan taxa. PLoS Genetics 5:e1000753.

Omelchenko MV, Galperin MY, Wolf YI, Koonin EV. 2010. Non-homologous isofunctional enzymes: a systematic analysis of alternative solutions in enzyme evolution. Biology Direct 5:31.

Orr HA. 2000. Adaptation and the cost of complexity. Evolution 54:13-20.

Osada N, Akashi H. 2012. Mitochondrial-nuclear interactions and accelerated compensatory evolution: evidence from the primate cytochrome C oxidase complex. Molecular Biology and Evolution 29:337.

Osborn MJ, Miller JR. 2007. Rescuing yeast mutants with human genes. Briefings in Functional Genomics and Proteomics 6:104-111.

Paigen K, Petkov PM. 2018. PRDM9 and its role in genetic recombination. Trends in Genetics 34:291-300.

Papakostas S, Vøllestad LA, Bruneaux M, Aykanat T, Vanoverbeke J, Ning M, Primmer CR, Leder EH. 2014. Gene pleiotropy constrains gene expression changes in fish adapted to different thermal conditions. Nature Communications 5:4071.

Papp B, Pal C, Hurst LD. 2003. Dosage sensitivity and the evolution of gene families in yeast. Nature 424:194-197.

Park C, Zhang J. 2012. High expression hampers horizontal gene transfer. Genome Biology and Evolution 4:523-532.

Park S, Ruhlman TA, Weng M-L, Hajrah NH, Sabir JS, Jansen RK. 2017. Contrasting patterns of nucleotide substitution rates provide insight into dynamic evolution of plastid and mitochondrial genomes of Geranium. Genome Biology and Evolution 9:1766-1780.

Parker N, Wang Y, Meinke D. 2014. Natural variation in sensitivity to a loss of chloroplast translation in Arabidopsis. Plant Physiology 166:2013-2027.

Patron NJ, Rogers MB, Keeling PJ. 2004. Gene replacement of fructose-1, 6-bisphosphate aldolase supports the hypothesis of a single photosynthetic ancestor of chromalveolates. Eukaryotic Cell 3:1169-1175.

Perkins EL, Sterling JF, Hashem VI, Resnick MA. 1999. Yeast and human genes that affect the Escherichia coli SOS response. Proceedings of the National Academy of Sciences 96:2204-2209.

Pett W, Lavrov DV. 2015. Cytonuclear interactions in the evolution of animal mitochondrial tRNA metabolism. Genome Biology and Evolution 7:2089-2101.

Podlevsky JD, Chen JJ-L. 2016. Evolutionary perspectives of telomerase RNA structure and function. RNA biology 13:720-732.

Ponting CP. 2011. What are the genomic drivers of the rapid evolution of PRDM9? Trends in Genetics 27:165-171.

Popa O, Hazkani-Covo E, Landan G, Martin W, Dagan T. 2011. Directed networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes. Genome Research 21:599-609.

Postel Z, Touzet P. 2020. Cytonuclear genetic incompatibilities in plant speciation. Plants 9:487.

Power JJ, Pinheiro F, Pompei S, Kovacova V, Yüksel M, Rathmann I, Förster M, Lässig M, Maier B. 2021. Adaptive evolution of hybrid bacteria by horizontal gene transfer. Proceedings of the National Academy of Sciences 118:e2007873118.

Presgraves DC. 2018. Evaluating genomic signatures of “the large X-effect” during complex speciation. Molecular ecology 27:3822-3830.

Presgraves DC. 2010. The molecular evolutionary basis of species formation. Nature Reviews Genetics 11:175-180.

Presgraves DC. 2008. Sex chromosomes and speciation in *Drosophila*. Trends in Genetics 24:336-343.

Promislow DEL. 2004. Protein networks, pleiotropy and the evolution of senescence. Proceedings of the Royal Society of London. Series B: Biological Sciences 271:1225-1234.

Rand DM, Haney RA, Fry AJ. 2004. Cytonuclear coevolution: the genomics of cooperation. Trends in Ecology & Evolution 19:645-653.

Rivera MC, Jain R, Moore JE, Lake JA. 1998. Genomic evidence for two functionally distinct gene classes. Proceedings of the National Academy of Sciences 95:6239-6244.

Roth CW, Kobeski F, Walter MF, Biessmann H. 1997. Chromosome end elongation by recombination in the mosquito Anopheles gambiae. Molecular and cellular biology 17:5176-5183.

Salie MJ, Thelen JJ. 2016. Regulation and structure of the heteromeric acetyl-CoA carboxylase. Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids 1861:1207-1213.

Salinas-Giegé T, Giegé R, Giegé P. 2015. tRNA biology in mitochondria. International Journal of Molecular Sciences 16:4518-4559.

Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, Pääbo S, Patterson N, Reich D. 2014. The genomic landscape of Neanderthal ancestry in present-day humans. Nature 507:354-357.

Sato N, Takano H. 2017. Diverse origins of enzymes involved in the biosynthesis of chloroplast peptidoglycan. Journal of plant research 130:635-645.

Schluter D, Rieseberg LH. 2022. Three problems in the genetics of speciation by selection. Proceedings of the National Academy of Sciences 119:e2122153119.

Schmitz-Linneweber C, Kushnir S, Babiychuk E, Poltnigg P, Herrmann RG, Maier RM. 2005. Pigment deficiency in nightshade/tobacco cybrids is caused by the failure to edit the plastid ATPase alpha-subunit mRNA. Plant Cell 17:1815-1828.

Schumer M, Xu C, Powell DL, Durvasula A, Skov L, Holland C, Blazier JC, Sankararaman S, Andolfatto P, Rosenthal GG. 2018. Natural selection interacts with recombination to shape the evolution of hybrid genomes. Science 360:656-660.

Sehnal D, Bittrich S, Deshpande M, Svobodová R, Berka K, Bazgier V, Velankar S, Burley SK, Koča J, Rose AS. 2021. Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. Nucleic Acids Research 49:W431-W437.

Serrato-Capuchina A, Matute DR. 2018. The role of transposable elements in speciation. Genes 9:254.

Shaul S, Nussinov R, Pupko T. 2006. Paths of lateral gene transfer of lysyl-aminoacyl-tRNA synthetases with a unique evolutionary transition stage of prokaryotes coding for class I and II varieties by the same organisms. BMC Evolutionary Biology 6:22.

Shutt TE, Gray MW. 2006. Bacteriophage origins of mitochondrial replication and transcription proteins. Trends in Genetics 22:90-95.

Skippington E, Ragan MA. 2012. Phylogeny rather than ecology or lifestyle biases the construction of Escherichia coli–Shigella genetic exchange communities. Open Biology 2:120112.

Sloan DB, Havird JC, Sharbrough J. 2017. The on-again, off-again relationship between mitochondrial genomes and species boundaries. Molecular ecology 26:2212-2236.

Sloan DB, Warren JM, Williams AM, Wu Z, Abdel-Ghany SE, Chicco AJ, Havird JC. 2018. Cytonuclear integration and co-evolution. Nature Reviews Genetics 19:635-648.

Slomka S, Françoise I, Hornung G, Asraf O, Biniashvili T, Pilpel Y, Dahan O. 2020. Experimental evolution of Bacillus subtilis reveals the evolutionary dynamics of horizontal gene transfer and suggests adaptive and neutral effects. Genetics 216:543-558.

Small I, Akashi K, Chapron A, Dietrich A, Duchene AM, Lancelin D, Maréchal-Drouard L, Menand B, Mireau H, Moudden Y. 1999. The strange evolutionary history of plant mitochondrial tRNAs and their aminoacyl-tRNA synthetases. JOURNAL OF HEREDITY 90:333-337.

Smith DR, Keeling PJ. 2015. Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. Proceedings of the National Academy of Sciences 112:10177-10184.

927 Smith SD, Wang S, Rausher MD. 2013. Functional evolution of an anthocyanin pathway enzyme during a
928 flower color transition. Molecular Biology and Evolution 30:602-612.

929 Sorek R, Zhu Y, Creevey CJ, Francino MP, Bork P, Rubin EM. 2007. Genome-wide experimental
930 determination of barriers to horizontal gene transfer. Science 318:1449-1452.

931 Sutherland KM, Ward LM, Colombero CR, Johnston DT. 2021. Inter-domain horizontal gene transfer of
932 nickel-binding superoxide dismutase. Geobiology 19:450-459.

933 Swamy KB, Schuyler SC, Leu J-Y. 2021. Protein complexes form a basis for complex hybrid
934 incompatibility. Frontiers in Genetics 12:609766.

935 Sweasy JB, Loeb LA. 1993. Detection and characterization of mammalian DNA polymerase beta mutants
936 by functional complementation in Escherichia coli. Proceedings of the National Academy of Sciences
937 90:4626-4630.

938 Tagwerker C, Dupont CL, Karas BJ, Ma L, Chuang R-Y, Benders GA, Ramon A, Novotny M, Montague
939 MG, Venepally P. 2012. Sequence analysis of a complete 1.66 Mb Prochlorococcus marinus MED4
940 genome cloned in yeast. Nucleic Acids Research 40:10375-10383.

941 Timmis JN, Ayliffe MA, Huang CY, Martin W. 2004. Endosymbiotic gene transfer: Organelle genomes
942 forge eukaryotic chromosomes. Nature Review Genetics 5:123-135.

943 Tovar J, León-Avila G, Sánchez LB, Sutak R, Tachezy J, Van Der Giezen M, Hernández M, Müller M,
944 Lucocq JM. 2003. Mitochondrial remnant organelles of Giardia function in iron-sulphur protein maturation.
945 Nature 426:172-176.

946 van Mourik PM, de Jong J, Agpalo D, Claussin C, Rothstein R, Chang M. 2016. Recombination-mediated
947 telomere maintenance in Saccharomyces cerevisiae is not dependent on the Shu complex. PloS one
948 11:e0151314.

949 Wagner GP, Zhang J. 2011. The pleiotropic structure of the genotype–phenotype map: the evolvability of
950 complex organisms. Nature Reviews Genetics 12:204-213.

951 Waller RF, Slamovits CH, Keeling PJ. 2006. Lateral gene transfer of a multigene region from
952 cyanobacteria to dinoflagellates resulting in a novel plastid-targeted fusion protein. Molecular Biology and
953 Evolution 23:1437-1443.

954 Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, Sabatini DM. 2015.
955 Identification and characterization of essential genes in the human genome. Science 350:1096-1101.

Wang Z, Liao B-Y, Zhang J. 2010. Genomic patterns of pleiotropy and the evolution of complexity. Proceedings of the National Academy of Sciences 107:18034-18039.

Warren JM, Salinas-Giegé T, Triant DA, Taylor DR, Drouard L, Sloan DB. 2021. Rapid shifts in mitochondrial tRNA import in a plant lineage with extensive mitochondrial tRNA gene loss. Molecular Biology and Evolution 38:5735-5751.

Warren JM, Sloan DB. 2022. Extensive retargeting of plant aminoacyl tRNA synthetases correlates with mitochondrial tRNA gene loss. bioRxiv.

Warren JM, Sloan DB. 2021. Hopeful monsters: unintended sequencing of famously malformed mite mitochondrial tRNAs reveals widespread expression and processing of sense–antisense pairs. NAR genomics and bioinformatics 3:lqaa111.

Warren JM, Sloan DB. 2020. Interchangeable parts: The evolutionarily dynamic tRNA population in plant mitochondria. Mitochondrion 52:144-156.

Watanabe K. 2010. Unique features of animal mitochondrial translation systems–The non-universal genetic code, unusual features of the translational apparatus and their relevance to human mitochondrial diseases. Proceedings of the Japan Academy, Series B 86:11-39.

Wellner A, Gophna U. 2008. Neutrality of foreign complex subunits in an experimental model of lateral gene transfer. Molecular Biology and Evolution 25:1835-1840.

Wellner A, Lurie MN, Gophna U. 2007. Complexity, connectivity, and duplicability as barriers to lateral gene transfer. Genome biology 8:R156.

Weng ML, Ruhlman TA, Jansen RK. 2016. Plastid-nuclear interaction and accelerated coevolution in plastid ribosomal genes in Geraniaceae. Genome Biology and Evolution 8:1824-1838.

Williams AM, Carter OG, Forsythe ES, Mendoza HK, Sloan DB. 2022. Gene duplication and rate variation in the evolution of plastid ACCase and Clp genes in angiosperms. Molecular phylogenetics and evolution 168:107395.

Williams D, Gogarten JP, Papke RT. 2012. Quantifying homologous replacement of loci between haloarchaeal species. Genome Biology and Evolution 4:1223-1244.

Woese CR, Olsen GJ, Ibba M, Söll D. 2000. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. Microbiology and Molecular Biology Reviews 64:202-236.

Wolfe KH, Li WH, Sharp PM. 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proceedings of the National Academy of Sciences 84:9054-9058.

Woods LC, Gorrell RJ, Taylor F, Connallon T, Kwok T, McDonald MJ. 2020. Horizontal gene transfer potentiates adaptation by reducing selective constraints on the spread of genetic variation. Proceedings of the National Academy of Sciences 117:26868-26875.

Yan Z, Ye G, Werren J. 2019. Evolutionary rate correlation between mitochondrial-encoded and mitochondria-associated nuclear-encoded proteins in insects. Molecular Biology and Evolution 36:1022-1036.

Yaremchuk A, Kriklivyi I, Tukalo M, Cusack S. 2002. Class I tyrosyl-tRNA synthetase has a class II mode of cognate tRNA recognition. EMBO Journal 21:3829-3840.

Zhang J, Ruhlman TA, Sabir J, Blazier JC, Jansen RK. 2015. Coordinated rates of evolution between interacting plastid and nuclear genes in Geraniaceae. The Plant Cell 27:563-573.

Zhang J, Yang J-R. 2015. Determinants of the rate of protein sequence evolution. Nature Reviews Genetics 16:409-420.

Zhang J-M, Zou L. 2020. Alternative lengthening of telomeres: from molecular mechanisms to therapeutic outlooks. Cell & Bioscience 10:30.

Zotenko E, Mestre J, O'Leary DP, Przytycka TM. 2008. Why do hubs in the yeast protein interaction network tend to be essential: reexamining the connection between the network topology and essentiality. PLoS Computational Biology 4:e1000140.