Robustness and Evolvability in Transcriptional Regulation

José Aguilar-Rodríguez and Joshua L. Payne

Abstract The relationship between genotype and phenotype is central to our understanding of development, evolution, and disease. This relationship is known as the genotype-phenotype map. Gene regulatory circuits occupy a central position in this map, because they control when, where, and to what extent genes are expressed, and thus drive fundamental physiological, developmental, and behavioral processes in living organisms as different as bacteria and humans. Mutations that affect these gene expression patterns are often implicated in disease, so it is important that gene regulatory circuits are robust to mutation. Such mutations can also bring forth beneficial phenotypic variation that embodies or leads to evolutionary adaptations or innovations. Here, we review recent theoretical and experimental work that sheds light on the robustness and evolvability of gene regulatory circuits.

1 Introduction

Two of the most fundamental properties of living systems are robustness and evolvability (Wagner, 2005; Masel & Trotter, 2010). Robustness is the invariance of a phenotype in the presence of environmental or genetic change. Evolvability is the ability of a living system to generate phenotypic variation that is both heritable and adaptive (Payne & Wagner, 2019). A large number of studies have focused on elucidating the molecular mechanisms of both robustness and evolvability, and

J. Aguilar-Rodríguez
Department of Biology, Stanford University, Stanford, CA, USA

Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, CA, USA
e-mail: jose.aguilar@stanford.edu

J. L. Payne (✉)
Institute of Integrative Biology, ETH, Zurich, Switzerland

Swiss Institute of Bioinformatics, Lausanne, Switzerland
e-mail: joshua.payne@env.ethz.ch

© Springer Nature Switzerland AG 2021

A. Crombach (ed.), *Evolutionary Systems Biology*,
https://doi.org/10.1007/978-3-030-71737-7_9

establishing the relationship between these two properties, at multiples scales of biological organization, ranging from the structural and functional properties of RNA and proteins to the ability of a metabolic network to create biomass from nutrients. The evidence from these studies suggests that robustness can facilitate evolvability. Here, we review a subset of this large body of work, specifically highlighting studies that have focused on the molecular mechanisms of robustness and evolvability in transcriptional regulation. Specifically, we discuss mechanisms of robustness against perturbations caused by genetic mutations (mutational robustness), rather than mechanisms of robustness against nongenetic perturbations (environmental robustness), although mutational robustness is correlated with environmental robustness in many instances (de Visser et al., 2003; Lehner, 2010).

Most of the early studies on the robustness and evolvability of transcriptional regulation were theoretical or computational (Kauffman, 1969; Wagner, 1996; Bergman & Siegal, 2003). The reason for this is the enormous complexity of the gene circuits (and their individual molecular components) that control when, where, and to what extent genes are expressed. Advances in high-throughput technologies are changing this picture, providing mechanistic insight into how transcriptional regulation is robust to mutational change, yet able to bring forth new and beneficial phenotypes. These studies, which are the focus of our chapter, are also validating a long-standing body of theoretical work on the relationship between robustness and evolvability in transcriptional regulation.

While there are multiple mechanisms of gene regulation acting at different stages of information transmission from DNA to protein (Alonso & Wilkins, 2005; Keren et al., 2010; Pauli et al., 2011; Guttman & Rinn, 2012; Pelechano & Steinmetz, 2013; Smith & Meissner, 2013; Tian & Manley, 2016; Zhao et al., 2016), the fundamental spatiotemporal control of gene expression occurs at the level of gene transcription. Transcriptional regulation drives essential physiological processes—e.g., how cells respond to their environment (Ptashne & Gann, 2002)—, behavioral processes—e.g., mating in yeast (Tsong et al., 2006)—, and developmental processes—e.g., embryonic patterning in diptera (Lawrence, 1992). Transcriptional regulation is mediated by sequence-specific DNA-binding proteins known as transcription factors (TFs). They regulate gene transcription by binding short DNA sequences (6–12 base pairs) known as TF binding sites in the promoters or enhancers of genes. The binding of a TF to a gene's regulatory region may activate or repress the transcription of that gene by promoting or blocking the recruitment of the RNA polymerase to the transcription start site. The strength of this regulatory effect is partly determined by the TF's affinity for its site. Genes coding for TFs typically represent 5–10% of the total number of genes in a given genome (Madan Babu et al., 2006; Vaquerizas et al., 2009; Stormo & Zhao, 2010), and their products can regulate the expression of other TFs, forming transcriptional regulatory circuits that control gene expression in space and time. These circuits occupy a central position in the mapping from genotype to phenotype, and drive fundamental physiological, developmental, and behavioral processes in all living organisms from bacteria to humans.

In this chapter, we first review the mechanisms of mutational robustness and evolvability in (*trans*-acting) transcription factors and their (*cis*-acting) DNA-binding sites, that is, in the individual components of gene regulatory circuits. We then review the more global mechanisms of robustness and evolvability that emerge at the level of whole gene regulatory circuits.

2 Robustness and Evolvability of Gene Regulatory Circuit Components

2.1 Robustness and Evolvability of Transcription Factors

Genotype of a TF: Amino acid sequence of the protein
Phenotype of a TF: Ability to bind DNA specifically and regulate gene transcription

One of the most useful and productive distinctions in biology is that between genotype and phenotype, which can be defined at different levels of biological organization, ranging from biological macromolecules to whole organisms (Wagner, 2011). In this section, we define the genotype of a TF as its amino acid sequence. A TF can have multiple phenotypes. For example, its ability to bind a specific short DNA sequence with a given binding affinity or its ability to recruit RNA polymerases at a given rate.

TFs may have several functional domains—conserved protein segments that can function independently, each with a different function (Bornberg-Bauer & Albà, 2013; Stormo, 2013; Toll-Riera & Albà, 2013). TFs typically have just one DNA-binding domain, which can function autonomously (Stormo, 2013). Other TF domains are responsible for dimerization, with many TFs functioning as homodimers or heterodimers. Finally, some TF domains mediate interactions with other proteins to form large molecular complexes that regulate the rate of transcription. For instance, many TFs have an activation domain that interacts with the basal transcriptional machinery and coactivator complexes to initiate transcription (Latchman, 2008). TFs can be classified into families based on the structures and sequence similarity of their DNA-binding domains (Weirauch & Hughes, 2011; Stormo, 2013). TFs from the same family have similar structures, and thus bind DNA with the same overall geometry of interaction (Stormo, 2013). TFs from the same family usually also have a common ancestry, and have diverged through evolutionary processes such as gene duplication and species diversification.

2.1.1 The Robustness of the Protein Structure of Transcription Factors

While most protein mutations tend to be deleterious (Eyre-Walker & Keightley, 2007), the structure and biological activity of proteins are to some extent robust to mutations. However, different proteins, and even different domains within a single

protein, can vary widely in their level of mutational robustness. The robustness of TFs against mutations has important implications for human disease, because numerous Mendelian diseases are caused by mutations in TFs, especially in homeodomain TFs (Veraksa et al., 2000). Similarly, somatic mutations in Cys_2-His_2 zinc finger (C2H2-ZF) domains are commonly mutated in cancer cells, thus likely contributing to the transcriptional dysregulation that is characteristic of this disease (Munro et al., 2018). Even among putatively healthy individuals, there are more than 50,000 polymorphisms segregating in the human population that are found in the DNA-binding domains of sequence-specific TFs (Barrera et al., 2016). Many of these variants are likely to affect binding activity, and thus contribute to phenotypic heterogeneity and disease. However, of 177 nonsynonymous polymorphisms chosen because they are predicted to have an effect on TF activity, 40 were not found to affect DNA binding and/or specificity in vitro (Barrera et al., 2016). This shows how regulatory proteins can be robust to mutational change, including mutations in their DNA-binding domains. This is further exemplified by the human basic helix-loop-helix (bHLH) TF Max, which interfaces DNA with five amino acid residues. Three of these positions can be mutated into any other amino acid without altering binding specificity, although some amino acid substitutions in two of these three positions modulate binding affinity (Maerkl & Quake, 2009). However, the general rule is that mutations in the DNA-binding domains of TFs cause changes in binding specificity (Cook et al., 1994; Mathias et al., 2001; Noyes et al., 2008; Aggarwal et al., 2010; De Masi et al., 2011).

The activation domains of TFs are less conserved across species and less structured than DNA-binding domains (Latchman, 2008). They are intrinsically disordered domains. Activation domains are more robust against amino acid replacements than DNA-binding domains (Majithia et al., 2016; Staller et al., 2018). For example, the function of the human nuclear receptor PPARγ is less affected by amino acid substitutions in its activation domain AF-1 than its DNA-binding domain (Majithia et al., 2016). Strikingly, the activation domain is even less sensitive to mutations than the "hinge" region connecting the DNA-binding domain to the ligand-binding domain of this TF.

Many transcription factors cooperatively bind DNA with other protein factors. Such interactions may alleviate the consequences of mutations in a TF's DNA-binding domain or mutations in its binding sites. Protein–protein interactions with a different TF may stabilize the binding of a TF to particular genomic locations while its protein–DNA interactions evolve gradually. For example, the conserved interaction of Matα1 with Mcm1 may explain the dramatic changes in the binding specificity of this regulator of mating type in ascomycete fungi over relatively short evolutionary time scales (Baker et al., 2011). Similarly, the cooperativity between Mcm1 and Rap1 can stabilize the binding of Mcm1 to weak binding sites (Sorrells et al., 2018). Thus, interactions with protein partners can provide a source of mutational robustness for both TFs and their binding sites.

2.1.2 Robustness in Duplicated Transcription Factors

Gene duplication is one of the main forces shaping eukaryotic genomes (Zhang, 2003). Because gene duplicates (paralogs) are initially redundant, one of the copies can act as a backup, compensating deleterious mutations in the other copy (Keane et al., 2014). Therefore, gene duplication is an important mechanism of mutational robustness (Gu et al., 2003; Conant & Wagner, 2004; Fares, 2015), and plays an important role in the evolution of TFs (Babu & Teichmann, 2003; Teichmann & Babu, 2004). Many paralogous TFs recognize the same or very similar sets of binding sites in vitro (Weirauch et al., 2014), and they also bind many of the same genomic regions in vivo (Hollenhorst et al., 2007), indicating that they may be fully or partially redundant. Even distant TF paralogs can partially compensate one another against loss-of-function mutations (Kafri et al., 2005; He & Zhang, 2006; Tischler et al., 2006). Additionally, TFs with similar binding specificities tend to regulate sets of genes with similar biological functions. This may minimize the negative consequences of the "cross-talk" that occurs when different TFs bind similar sets of sites (Itzkovitz et al., 2006), or the negative consequence of specificity-changing mutations in a TFs DNA-binding domain.

2.1.3 Many Transcription Factors Are Clients of the Molecular Chaperone HSP90

A protein is classified as a client of HSP90 if it interacts physically with the chaperone and if the inhibition of HSP90 function reduces protein client activity. The most common protein clients of the eukaryotic chaperone HSP90 are TFs, including nuclear steroid receptors, but also PAS family TFs, p53, STAT3, and chromatin proteins such as trithorax (Taipale et al., 2010). BES1, a TF in the steroid hormone pathway in *Arabidopsis thaliana*, is a client of HSP90, but its closest paralog, BZR1, is not an HSP90 client. This difference in the client status of two highly similar proteins facilitates a test of whether HSP90 can enhance the mutational robustness of a TF. *BES1* shows relaxed selection compared to *BZR1* as expected if HSP90 allows BES1 to explore a greater fraction of genotype space without losing function (Lachowiec et al., 2013). Similarly, HSP90 clients in yeast, including many TFs, evolve faster than their nonclient paralogs, suggesting that HSP90 can increase the mutational tolerance of its client TFs (Lachowiec et al., 2013; Alvarez-Ponce et al., 2019).

2.1.4 The Evolvability of Transcription Factors

New gene expression patterns can evolve by changes in TFs and their binding sites that lead to the rewiring of a gene regulatory circuit. However, the adaptive evolution of a TF can be heavily constrained by both epistasis and pleiotropy. Epistasis between different residues can severely restrict the evolutionary trajectories of any

evolving protein (Starr & Thornton, 2016), including TFs, and mutations in a TF can have strong pleiotropic effects on a regulatory circuit because all its gene targets can be affected by changes in its binding specificity (Britten & Davidson, 1969; Stern, 2000; Carroll, 2005; Wray, 2007).

For these reasons, while there are clear cases of mutations in TFs contributing to the adaptive evolution of gene regulation (Galant & Carroll, 2002; Ronshaugen et al., 2002; Lynch et al., 2008, 2011), most studies of regulatory evolution have focused on the evolution of *cis*-regulatory elements, such as promoters and enhancers (Prud'homme et al., 2007; Carroll, 2008; Stern & Orgogozo, 2008). However, some properties of a TF can promote its evolvability. For example, the organization of TFs in functionally autonomous protein domains that can evolve independently allows for the evolutionary emergence of new TFs by domain rearrangement, including the gain or loss of protein domains, the shuffling of already existing domains, and expansions or contractions in the number of a given protein domain (Bornberg-Bauer & Albà, 2013). Such domain rearrangements in TF families can lead to a major functional shift and the subsequent expansion of new sub-families of TFs (Schmitz et al., 2016). Changes in TF function can be mediated by either a change in DNA-binding specificity or changes in protein–protein interactions with other TFs or signaling proteins. Together with gene duplication, domain rearrangement was a potent force in the evolution of major TF families (Schmitz et al., 2016), including bHLH TFs (Amoutzias et al., 2004; Morgenstern & Atchley, 2018).

The robustness of a protein can facilitate the acquisition of novel functions (Bloom et al., 2006). In other words, genetic robustness and evolvability can be synergistic (Wagner, 2008). For example, the robustness of TFs due to the existence of paralogous TFs can promote the evolution of novel adaptive regulatory roles. Gene duplication can facilitate the evolution of TFs with divergent binding specificities that control different sets of genes and facilitate adaptation to new niches (Perez et al., 2014). A duplication of a Hox3 TF in two paralogs deep in the lineage of Cyclorrhaphan flies allowed one of the two paralogs (Bicoid, Bcd) to gain the important developmental role of controlling anterior-posterior patterning in fly embryos (Stauber et al., 1999). After the emergence by duplication of Bcd, this TF acquired at least two large-effect mutations that changed its DNA specificity and played a major role in the evolution of this TFs controlling role during early fly development (Liu et al., 2018). In combination with protein domain rearrangements, and given enough evolutionary time, the diversification and expansion of TF families can have deep evolutionary consequences, and it tends to be associated with increases on organismic complexity, both morphological and in terms of the number of cell types (Carroll et al., 2001; De Bodt et al., 2003; Irish, 2003; Levine & Tjian, 2003; Degnan et al., 2009; Vaquerizas et al., 2009; Ruiz-Trillo et al., 2013; Albertin et al., 2015; Schmitz et al., 2016).

C2H2-ZF TF are the most common class of TFs in metazoans (Vaquerizas et al., 2009). The binding specificity of a C2H2-ZF domain is mainly conferred by four DNA-contacting residues within the domain's α-helix (Pabo et al., 2001). A C2H2-ZF domain can bind a wide range of three or four base pairs. C2H2-ZF

TFs typically contain tandem arrays of these domains that bind contiguous DNA sites, which allows this type of TFs the ability to recognize an incredibly large diversity of DNA sequences of variable length (Basciotta et al., 2013). Domain rearrangement and gene duplication have played an important role in the expansion and diversification of C2H2-ZF TFs in animals (Schmitz et al., 2016). Even without gene duplication, the binding specificities of C2H2-ZFs can change over short evolutionary timescales. For example, one-to-one orthologous C2H2-ZF TFs typically show divergence in their DNA-contacting residues across closely related *Drosophila* species (Nadimpalli et al., 2015). The predicted DNA-binding specificities of these domains gradually change as a function of phylogenetic distance, suggesting that single-copy TFs can diverge in their DNA-binding specificities via small evolutionarily viable steps. Robustness may be behind this process of binding specificity modifications in single-copy TFs. While the binding energy of other eukaryotic C2H2-ZF TFs largely depends on base-contacting amino acids, C2H2-ZF TFs from metazoans use non-base contacting amino acids to establish hydrogen bonds with the phosphate backbone of DNA that increase their overall binding energy (Najafabadi et al., 2017). These non-base contacting amino acids provide robustness to mutations in base-contacting amino acids, which may have led to the ability of single-copy C2H2-ZF TFs to rapidly diverge in binding preference. This robustness of the C2H2-ZF domain may have played an important role in the regulatory evolution not only of *Drosophila* but also of other metazoans, including humans, where there is evidence of adaptive evolution in C2H2-ZF domains (Emerson & Thomas, 2009). Similarly, robustness played a role in the evolution of binding specificity in steroid receptors (Starr et al., 2017; Payne & Wagner, 2019).

Besides specificity-altering mutations in DNA-binding domains, gene regulation can also evolve via mutations that change how TFs respond to upstream signaling pathways. For example, a TF that plays an essential role during pregnancy, CEBPB, changed its response to cAMP/PKA signaling from repression to activation due to three amino acid replacements that affected phosphorylation sites in an internal regulatory domain (Lynch et al., 2011). This novel function evolved coincident with the evolution of pregnancy in placental mammals. The alteration of post-translational modification sites is therefore an additional mechanism by which TFs evolve. Such changes facilitate gene regulatory innovations in signaling-dependent transcriptional circuits by altering the function of TFs in specific cell types, while avoiding or minimizing deleterious pleiotropic effects on other cellular functions.

In sum, the regulatory proteins involved in transcriptional regulation often exhibit robustness to mutation, which can facilitate evolvability.

2.2 Robustness and Evolvability of Transcription Factor Binding Sites

Genotype of a *cis*-regulatory element: short DNA sequence
Phenotype of a *cis* element: molecular ability to bind a regulatory protein

2.2.1 The Robustness of Regulatory Sequences

In the previous section, we reviewed how regulatory proteins can be robust to mutational change, and how such robustness can enhance evolvability. The regulatory regions that these proteins bind are also robust to genetic change, and similarly their robustness can synergize with their evolvability. The very short length of TF binding sites confers an additional source of robustness against mutations. Shorter binding sites are less easily disrupted by mutations because they offer a smaller mutational target, a benefit that comes at the expense of reduced specificity (Stewart et al., 2012). However, TF binding sites can also be intrinsically robust to mutations. TFs can specifically bind to dozens or hundreds of different short DNA sequences (Sengupta et al., 2002; Berger et al., 2006; Badis et al., 2009; Wong et al., 2013; Weirauch et al., 2014), and these sequences tend to be mutationally interconnected to one another forming genotype networks of TF binding sites (Payne & Wagner, 2014; Khalid et al., 2016; Aguilar-Rodríguez et al., 2018). In such a genotype network, vertices represent DNA sequences with the ability to specifically bind a particular TF, and two vertices are connected by a link if their associated sequences differ by just a single small mutation, such as a point mutation. The existence of genotype networks of TF binding sites implies that mutations to a binding site will often create mutant sequences that are still able to bind the same TF, thus conferring mutational robustness. Additionally, mutational neighbors in a genotype network tend to have similar binding affinity for a given TF, indicating that binding affinity is also robust to mutation (Payne & Wagner, 2014; Aguilar-Rodríguez et al., 2017). This is important because changes in binding affinity can lead to changes in gene expression (Kasowski et al., 2010; Shultzaberger et al., 2010; Sharon et al., 2012), and both large (Giaever et al., 2002; Gerdes et al., 2003; Dietzl et al., 2007; Hillenmeyer et al., 2008; Ramani et al., 2012; Hart et al., 2015) and small (Dykhuizen et al., 1987; Dekel & Alon, 2005; Rest et al., 2013; Keren et al., 2016) deviations from an optimal mean level of expression can be detrimental to organismal fitness. However, some genes have a nonlinear fitness-expression function with a plateau of maximal fitness for a wide range of expression levels (Rest et al., 2013; Bergen et al., 2016; Keren et al., 2016; Duveau et al., 2017). Therefore, the promoters of these genes are robust to many *cis*-regulatory mutations. For example, nearly all mutations and polymorphisms in the promoter of the yeast gene *TDH3* have no significant effect on fitness in a rich medium containing glucose (Duveau et al., 2017).

The robustness of TF binding sites allows the accumulation of genetic diversity in binding sites, both within species (Aguilar-Rodríguez et al., 2017, 2018), and between species (Weirauch & Hughes, 2010). Intra-specific variation in TF sites is pervasive (Zheng et al., 2011; Garfield et al., 2012; Spivakov et al., 2012; Khurana et al., 2013; Arbiza et al., 2013), and such differences often do not impact the expression level of target genes (Kasowski et al., 2010; Zheng et al., 2010). Similarly, over longer evolutionary time scales, regulatory sequences can diverge considerably at the sequence level without a corresponding divergence at

the phenotypic level of the gene expression patterns they control (Ludwig et al., 2000; Odom et al., 2007).

Another important source for robustness of TF binding sites is the presence of multiple binding sites for the same TF in close proximity to one another (Johnson et al., 1979; Giniger & Ptashne, 1988; Carey et al., 1990; Thanos & Maniatis, 1995; Wasserman & Fickett, 1998; Krivan & Wasserman, 2001; Pfeiffer et al., 2002; Ezer et al., 2015). These homotypic clusters of binding sites are common across all domains of life (Lifanov et al., 2003; Gotea et al., 2010; Gama-Castro et al., 2011). Mutations in a TF binding site can be compensated by the presence of nearby non-mutated TF binding sites (Somma et al., 1991; Spivakov et al., 2012; Kilpinen et al., 2013). This mechanism of mutational robustness has already been reviewed more extensively elsewhere (Payne & Wagner, 2015).

However, TF binding sites are not robust to all mutations, and mutations in these regulatory sequences can often be deleterious and cause disease (Musunuru et al., 2010; Harismendy et al., 2011). Indeed, the majority of single-nucleotide variants in DNA regulatory regions associated with different human diseases tend to alter TF binding sites (Maurano et al., 2012). For example, de novo mutations in DNA regulatory elements active in the human brain are associated with different neurodevelopmental disorders and predicted to increase the binding affinity of the binding sites in which they fall (Short et al., 2018). Mutations that fall in TF binding sites can increase cancer risk (Pomerantz et al., 2009; Khurana et al., 2013; Weinhold et al., 2014; Katainen et al., 2015; Melton et al., 2015). For example, noncoding single-nucleotide variants associated to breast cancer can modulate TF binding affinity resulting in transcriptional misregulation (Liu et al., 2017), which is a hallmark of many cancer types (Lee & Young, 2013; Bhagwat & Vakoc, 2015).

2.2.2 The Evolvability of Regulatory Sequences

Substitutions in *cis*-regulatory sequences may produce novel gene expression patterns associated to evolutionary innovations and adaptations (Wray, 2007; Prud'homme et al., 2007). Single-base pair substitutions in a TF binding site can change the regulatory control of a target gene from one TF to another (Payne & Wagner, 2014; Aguilar-Rodríguez et al., 2018), and this may lead to profound changes in development, physiology or behavior. For example, in rice, a single mutation in the promotor of a C2H2-ZF TF gene reduces its expression by creating a binding site for the transcriptional repressor MYB, and this change in expression increases resistance against rice blast—a fungal disease that can cause significant crop loss (Li et al., 2017). Standing genetic variation within TF binding sites can also contribute to evolutionary adaptation. For example, a recent high-throughput precise genome editing screen found that among 16,006 natural genetic variants in yeast, 572 variants with a significant fitness effect in glucose media were highly enriched in promoters, particularly in TF binding sites (Sharon et al., 2018). The genetic diversity accumulated within binding sites as a consequence of their robustness to mutational change provides an ideal "testing ground" for new mutations by

allowing the exploration of many different genetic backgrounds (Aguilar-Rodríguez et al., 2017, 2018). Robustness is a way to explore new mutational neighborhoods while preserving a phenotype, because it is almost always possible to transform one site into another via a series of mutations that preserve TF binding. Some of these new mutations may create binding sites for a different TF, which may lead to adaptive changes in gene expression (Payne & Wagner, 2014). Therefore, the genetic robustness of TF binding sites can synergize with their evolvability. For example, a comparative analysis of two well-studied transcriptional repressors from phages, showed that it was easier to evolve a cognate site from a non-cognate site for the repressor whose cognate sites are more robust to mutations (Igler et al., 2018). The structure of the genotype networks of TF binding sites furthers our understanding of the "robust-yet-evolvable" nature of these DNA sequences. They tend to be "small-world" (Watts & Strogatz, 1998; Aguilar-Rodríguez et al., 2018), which indicates that binding sites tend to be highly clustered in genotype space (robustness), but also that it is possible to traverse the network with just a few mutations, thus providing efficient access to adjacent genotype networks of other TFs (evolvability).

TF binding sites are short enough that it is possible to study their evolvability comprehensively (Rowe et al., 2010; Jimenez et al., 2013). For example, one can easily measure how strongly a TF binds tens of thousands of different DNA sequences (Berger et al., 2006), and this information is available for thousands of TFs from hundreds of species comprising multiple TF families (Weirauch et al., 2014). Binding affinity is an important molecular phenotype because it is an important contributor to a TFs ability to activate or repress a target gene, and the gene expression patterns that emerge from such TF-DNA interactions embody fundamental biological processes. The regulatory effect on gene expression of a TF can be either fine-tuned or even radically transformed by affinity-altering mutations in TF binding sites (Shultzaberger et al., 2010; Sharon et al., 2012). The mapping of a DNA sequence to binding affinity can be described as an adaptive landscape where one can study how mutation and natural selection can change the capacity of a DNA sequence to bind a particular TF (Berg et al., 2004). A recent study of more than a thousand such landscapes characterized their ruggedness using a variety of measures and found that they are highly navigable via a Darwinian process of mutation and selection, indicating that binding affinity—and thereby gene expression—is readily fine-tuned via mutations in TF binding sites (Aguilar-Rodríguez et al., 2017). These landscapes typically have just a single peak, and these peaks tend to be accessible from any location in the landscape via mutational pathways that increase monotonically in binding affinity. This type of smooth landscape promotes the evolvability of TF binding sites because mutation can bring forth beneficial phenotypic variation from any location on the landscape (Payne & Wagner, 2019). Therefore, the navigability of these TF binding affinity landscapes may have contributed to the enormous success of altering transcriptional regulation as a way to generate variation and innovation throughout evolution.

3 Robustness and Evolvability of Whole Gene Regulatory Circuits

Genotype of a gene circuit: genes for TFs and *cis*-regulatory sequences
Phenotype of a gene circuit: spatiotemporal gene expression pattern

The genotype of a gene regulatory circuit comprises the DNA sequences that encode the circuit's constituent transcription factors, as well as the binding sites for these factors in the promoters and enhancers of the circuit's genes. The phenotype of a gene regulatory circuit is its spatiotemporal gene expression pattern. Most of what we know about the robustness and evolvability of regulatory circuits comes from abstract computational models, such as Boolean circuits (Kauffman, 1969; Wagner, 1996). These studies have shown that there are many genotypes that have the same phenotype, meaning that a large number of circuit configurations are capable of driving the same gene expression pattern. Additionally, these genotypes tend to be arranged as a genotype network, such that it is usually possible to mutate any one regulatory circuit with a given phenotype into any other via a series of intermediates that also produce the phenotype (Ciliberti et al., 2007a, b). Such genotype networks confer robustness and evolvability to regulatory circuits, as they do at the level of the circuit components.

Our understanding of the robustness and evolvability of regulatory circuits continues to advance, with recent studies uncovering how circuit architecture influences the robustness of gene expression patterns to variation in morphogen production rates (Raspopovic et al., 2014; Li et al., 2018), how the robustness of regulatory circuits makes their evolution contingent upon chance mutational events (Starr et al., 2017), and how the evolvability of a regulatory circuit can transcend that of its constituent components (Lagator et al., 2017). In addition, a series of recent studies have shown how robustness and evolvability depend upon the dynamical mechanism a circuit uses to generate its phenotype (Jiménez et al., 2015; Schaerli et al., 2018), and how genotype networks facilitate "system drift," which enhances evolvability (Nocedal et al., 2017; Jaeger, 2018). These recent insights, realized using a combination of increasingly sophisticated computational models and experiments, are the focus of this section.

Gene regulatory circuits can produce the same gene expression pattern using distinct dynamical mechanisms. For example, Cotterell and Sharpe (2010) produced an atlas of regulatory circuits that interpret a morphogen gradient to produce a single stripe of gene expression, using a model based on the gap gene circuit, which drives segmental patterning in dipteran insects. This atlas includes circuits that produce stripes using six distinct dynamical mechanisms—unique spatiotemporal patterns of expression that all converge on a stripe. Analysis of the circuits employing each dynamical mechanism revealed variation in mutational robustness, measured either as the number of distinct circuit topologies, or as the volume of parameter space for a specific topology, that produce a stripe. Thus, identical gene expression phenotypes

can vary in their mutational robustness, depending upon the dynamical mechanism used to generate the phenotype.

The underlying explanation for this phenomenon is that distinct genotype networks are formed by the different sets of stripe-forming circuits with each dynamical mechanism. This means that it is generally not possible to smoothly transition via mutation from one dynamical mechanism to another. Jiménez et al. (2015) explored the consequences of this genotype network fragmentation for evolvability, defined as the ability of mutation to bring forth novel expression phenotypes, such as spatial gradients, inverse stripes, or multiple stripes. They found that evolvability is mechanism-dependent. Circuits using distinct mechanisms differ not only in which phenotypes mutation can bring forth, but also in the relative likelihood of mutations bringing forth such phenotypes. These findings were recently validated with experiments using synthetic stripe-forming regulatory circuits, in which random mutations to two circuits with distinct dynamical mechanisms produced different distributions of novel phenotypes (Schaerli et al., 2018). Such mechanism-dependent evolvability is an example of evolutionary constraint, in which circuits with the same phenotype, but different dynamical mechanisms, differ in their ability to generate phenotypic variation via mutation.

A population of circuits with any one phenotype can accrue mutations that alter the quantitative and qualitative features of the circuit, without affecting the circuit's gene expression phenotype. This phenomenon is referred to as system drift (Weiss & Fullerton, 2000). For example, the final patterning output of the gap gene circuit is conserved among *Drosophila melanogaster* and the scuttle fly *Megaselia abdita*, two species that last shared a common ancestor approximately 180 million years ago. In contrast, the dynamical mechanisms used to generate this phenotype show significant quantitative differences (Wotton et al., 2015). Specifically, gap domains appear more posteriorly and retract from the pole later in *M. adbita* than in *D. melanogaster*. Analysis of a data driven mathematical model of gap gene expression demonstrates that such system drift occurs because the set of gap gene circuits that drive segmental patterning—even those employing distinct dynamical mechanisms—are arranged as a genotype network (Crombach et al., 2016). This permits the accumulation of mutations that affect the strength and identity of a circuit's regulatory interactions, without affecting the final patterning phenotype. This also influences evolvability, because the potential for mutation to cause phenotypic variation varies across a genotype network (Wagner, 2011).

Phenotype-preserving mutations can thus serve as stepping-stones for evolutionary innovations. Another example of this phenomenon is the fungal transcriptional regulator Ndt80, a DNA-binding protein that is conserved across a large group of fungal species that last shared a common ancestor approximately 300 million years ago (Nocedal et al., 2017). Ndt80 is part of a regulatory circuit controlling meiosis and sporulation in most of these species, but it controls the formation of biofilms in *Candida albicans*. This shift in function resulted from system drift. This is evidenced by a comparative analysis of six fungal lineages, which uncovered extensive rewiring in the regulon of Ndt80, even in those lineages where the function of Ndt80 had not changed (Nocedal et al., 2017). This rewiring facilitated

the exploration of new regulatory circuits, potentiating the evolution of the novel phenotype of biofilm formation.

Such rewiring can be extreme. A comparative analysis of the regulatory circuits controlling the conversion of galactose to glucose-1-phosphate in *S. cerevisiae* and *C. albicans* provides an illustrative example (Dalal et al., 2016). In both species, three *GAL* genes are needed for this conversion. In *S. cerevisiae*, these genes are activated by the TF Gal4, whereas in *C. albicans*, they are activated by the TFs Rtg1 and Rtg3. This rewiring primarily occurred via changes in the *cis*-regulatory sequences of the *GAL* genes, which not only resulted in a qualitative change to the structure of the circuit, but also to quantitative changes in the induction ratios of the *GAL* genes and in their response to non-galactose signals. Some of these quantitative changes may have been adaptive for *S. cerevisiae*, because they contribute to the rapid fermentation of different sugars.

In sum, the robustness of gene regulatory circuits facilitates system drift, which enhances evolvability. The extent to which system drift occurs can depend upon the dynamical mechanism a circuit uses to generate its phenotype, which may constrain evolvability. However, system drift need not preserve dynamical mechanism, and some quantitative changes to dynamical mechanism may themselves be adaptive.

4 Concluding Remarks

While the idea behind genotype-phenotype maps can be traced back to the work of Sewall Wright (1932) and John Maynard Smith (1970), the term genotype-phenotype map ("genotype-phenotype mapping") itself was only coined in 1970 by Jim Burns (1970), who outlined the research goals of evolutionary systems biology before the development of systems biology made it feasible. He recognized early on the importance of integrating the mechanistic perspective of biochemistry, cell and molecular biology within the unifying framework of evolutionary biology:

> *It is the quantitative phenotype, arising from the genotypic prescriptions and the environment, which is of critical importance for the cell's survival and which therefore features in population genetic theory. A study of this synthetic problem would thus, by providing genotype-phenotype mappings for simple synthetic systems, help to connect two major areas of biological theory: the biochemical and the population genetic.*

The term "genotype-phenotype map" was re-introduced in 1991 by the developmental biologist Pere Alberch as a useful concept for the integration of genetics into the study of the complex developmental processes that generate morphological phenotypes such as the vertebrate limb (Alberch, 1991). Gene regulatory circuits occupy a central position in the map that goes from a genome to the high-level morphological phenotypes that interested Alberch.

In this chapter, we have reviewed the robustness of such circuits against genetic change, and how such robustness may have contributed to the enormous success of transcriptional regulation as a source of evolutionary novelty. We have

explored some of the mechanistic causes for the robustness and evolvability of transcription factors, the regulatory DNA sequences that they bind, and the gene circuits that emerge from the complex interactions between transcription factors and their binding sites. Space constraints do not allow us to review other molecular mechanism of evolvability in transcriptional regulation, such as stochastic gene expression that can promote evolvability by generating phenotypic heterogeneity in isogenic populations (Payne & Wagner, 2019). Also, we have focused on transcriptional regulation, although there are important levels of gene expression regulation, such as regulatory noncoding RNA, alternative splicing, epigenetic gene regulation, and protein posttranslational modifications. However, less is known about the mechanisms of robustness and evolvability at these levels, or about how they interact with the mechanisms reviewed here, although there is progress being made in this area (Payne et al., 2018). In the foreseeable future, the concepts and tools of evolutionary systems biology, aided by new technological advances, will further our understanding of the mechanisms by which gene regulation is robust to genetic change, yet capable of bringing forth evolutionary adaptations and innovations.

Acknowledgments We apologize to our peers whose important works in this rapidly advancing field were not cited due to space constraints. J.A.-R. acknowledges support by Swiss National Science Foundation grant P2ZHP3_174735 and European Molecular Biology Organization grant ALTF 724-2018. J.L.P. acknowledges support by Swiss National Science Foundation grant PP00P3_170604.

References

Aggarwal, P., Das Gupta, M., Joseph, A. P., et al. (2010). Identification of specific DNA binding residues in the TCP family of transcription factors in *Arabidopsis*. *Plant Cell, 22*, 1174–1189. https://doi.org/10.1105/tpc.109.066647

Aguilar-Rodríguez, J., Payne, J. L., & Wagner, A. (2017). A thousand empirical adaptive landscapes and their navigability. *Nature Ecology and Evolution, 1*, 45. https://doi.org/10.1038/s41559-016-0045

Aguilar-Rodríguez, J., Peel, L., Stella, M., et al. (2018). The architecture of an empirical genotype-phenotype map. *Evolution, 72*, 1242. https://doi.org/10.1111/evo.13487

Alberch, P. (1991). From genes to phenotype: Dynamical systems and evolvability. *Genetica, 84*, 5–11.

Albertin, C. B., Simakov, O., Mitros, T., et al. (2015). The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature, 524*, 220–224. https://doi.org/10.1038/nature14668

Alonso, C. R., & Wilkins, A. S. (2005). The molecular elements that underlie developmental evolution. *Nature Reviews. Genetics, 6*, 709–715. https://doi.org/10.1038/nrg1676

Alvarez-Ponce, D., Aguilar-Rodríguez, J., & Fares, M. A. (2019). Molecular chaperones accelerate the evolution of their protein clients in yeast. *Genome Biology and Evolution, 11*(8), 2360–2375.

Amoutzias, G. D., Robertson, D. L., Oliver, S. G., & Bornberg-Bauer, E. (2004). Convergent evolution of gene networks by single-gene duplications in higher eukaryotes. *EMBO Reports, 5*, 274–279. https://doi.org/10.1038/sj.embor.7400096

Arbiza, L., Gronau, I., Aksoy, B. A., et al. (2013). Genome-wide inference of natural selection on human transcription factor binding sites. *Nature Genetics, 45*, 723–729. https://doi.org/10.1038/ng.2658

Babu, M. M., & Teichmann, S. A. (2003). Evolution of transcription factors and the gene regulatory network in Escherichia coli. *Nucleic Acids Research, 31*, 1234–1244. https://doi.org/10.1093/nar/gkg210

Badis, G., Berger, M. F., Philippakis, A. A., et al. (2009). Diversity and complexity in DNA recognition by transcription factors. *Science, 324*, 1720–1723. https://doi.org/10.1126/science.1162327

Baker, C. R., Tuch, B. B., & Johnson, A. D. (2011). Extensive DNA-binding specificity divergence of a conserved transcription regulator. *Proceedings of the National Academy of Sciences, 108*, 7493–7498. https://doi.org/10.1073/pnas.1019177108

Barrera, L. A., Vedenko, A., Kurland, J. V., et al. (2016). Survey of variation in human transcription factors reveals prevalent DNA binding changes. *Science, 351*, 1450–1454. https://doi.org/10.1126/science.aad2257

Basciotta, M. D., Christensen, R. G., Celniker, S. E., et al. (2013). Global analysis of *Drosophila* Cys_2-His_2 zinc finger proteins reveals a multitude of novel recognition motifs and binding determinants. *Genome Research, 23*, 928–940. https://doi.org/10.1101/gr.151472.112

Berg, J., Willmann, S., & Lässig, M. (2004). Adaptive evolution of transcription factor binding sites. *BMC Evolutionary Biology, 4*, 42. https://doi.org/10.1186/1471-2148-4-42

Bergen, A. C., Olsen, G. M., & Fay, J. C. (2016). Divergent MLS1 promoters lie on a fitness plateau for gene expression. *Molecular Biology and Evolution, 33*, 1270–1279. https://doi.org/10.1093/molbev/msw010

Berger, M. F., Philippakis, A. A., Qureshi, A. M., et al. (2006). Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nature Biotechnology, 24*, 1429–1435. https://doi.org/10.1038/nbt1246

Bergman, A., & Siegal, M. L. (2003). Evolutionary capacitance as a general feature of complex gene networks. *Nature, 424*, 549–552. https://doi.org/10.1038/nature01765

Bhagwat, A. S., & Vakoc, C. R. (2015). Targeting transcription factors in cancer. *Trends in Cancer, 1*, 53–65. https://doi.org/10.1016/j.trecan.2015.07.001

Bloom, J. D., Labthavikul, S. T., Otey, C. R., & Arnold, F. H. (2006). Protein stability promotes evolvability. *Proceedings of the National Academy of Sciences, 103*, 5869–5874. https://doi.org/10.1073/pnas.0510098103

Bornberg-Bauer, E., & Albà, M. M. (2013). Dynamics and adaptive benefits of modular protein evolution. *Current Opinion in Structural Biology, 23*, 459–466.

Britten, R. J., & Davidson, E. H. (1969). Gene regulation for higher cells: A theory new facts regarding the organization of the genome. *Science, 25*, 349–357. https://doi.org/10.1016/j.ympev.2006.05.020

Burns, J. (1970). The synthetic problem and the genotype-phenotype relation in cellular metabolism. In C. H. Waddington (Ed.), *Towards a theoretical biology. Volume 3: Drafts. An IUBS Symposium* (pp. 47–51). Aldine Publishing Company.

Carey, M., Lin, Y. S., Green, M. R., & Ptashne, M. (1990). A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. *Nature, 345*, 361–364. https://doi.org/10.1038/345361a0

Carroll, S. B. (2005). Evolution at two levels: On genes and form. *PLoS Biology, 3*, e245. https://doi.org/10.1371/journal.pbio.0030245

Carroll, S. B. (2008). Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell, 134*, 25–36. https://doi.org/10.1016/j.cell.2008.06.030

Carroll, S. B., Grenier, J. K., & Weatherbee, S. D. (2001). *From DNA to diversity. Molecular genetics and the evolution of animal design*. Blackwell.

Ciliberti, S., Martin, O. C., & Wagner, A. (2007a). Innovation and robustness in complex regulatory gene networks. *Proceedings of the National Academy of Sciences, 104*, 13591–13596. https://doi.org/10.1073/pnas.0705396104

Ciliberti, S., Martin, O. C., & Wagner, A. (2007b). Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Computational Biology, 3*, 0164–0173. https://doi.org/10.1371/journal.pcbi.0030015

Conant, G. C., & Wagner, A. (2004). Duplicate genes and robustness to transient gene knock-downs in Caenorhabditis elegans. *Proceedings of the Royal Society B: Biological Sciences, 271*, 89–96. https://doi.org/10.1098/rspb.2003.2560

Cook, W. J., Mosley, S. P., Audino, D. C., et al. (1994). Mutations in the zinc-finger region of the yeast regulatory protein ADR1 affect both DNA binding and transcriptional activation. *The Journal of Biological Chemistry, 269*, 9374–9379.

Cotterell, J., & Sharpe, J. (2010). An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. *Molecular Systems Biology, 6*, 425. https://doi.org/10.1038/msb.2010.74

Crombach, A., Wotton, K. R., Jiménez-Guri, E., & Jaeger, J. (2016). Gap gene regulatory dynamics evolve along a genotype network. *Molecular Biology and Evolution, 33*, 1293–1307. https://doi.org/10.1093/molbev/msw013

Dalal, C. K., Zuleta, I. A., Mitchell, K. F., et al. (2016). Transcriptional rewiring over evolutionary timescales changes quantitative and qualitative properties of gene expression. *eLife, 5*, 1–24. https://doi.org/10.7554/eLife.18981

De Bodt, S., Raes, J., Van De Peer, Y., & Theißen, G. (2003). And then there were many: MADS goes genomic. *Trends in Plant Science, 8*, 475–483. https://doi.org/10.1016/j.tplants.2003.09.006

De Masi, F., Grove, C. A., Vedenko, A., et al. (2011). Using a structural and logics systems approach to infer bHLH-DNA binding specificity determinants. *Nucleic Acids Research, 39*, 4553–4563. https://doi.org/10.1093/nar/gkr070

de Visser, J. A. G. M., Hermisson, J., Wagner, G. P., et al. (2003). Perspective: Evolution and detection of genetic robustness. *Evolution (N Y), 57*, 1959–1972. https://doi.org/10.1554/02-750R

Degnan, B. M., Vervoort, M., Larroux, C., & Richards, G. S. (2009). Early evolution of metazoan transcription factors. *Current Opinion in Genetics & Development, 19*, 591–599. https://doi.org/10.1016/j.gde.2009.09.008

Dekel, E., & Alon, U. (2005). Optimality and evolutionary tuning of the expression level of a protein. *Nature, 436*, 588–592. https://doi.org/10.1038/nature03842

Dietzl, G., Chen, D., Schnorrer, F., et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature, 448*, 151–156. https://doi.org/10.1038/nature05954

Duveau, F., Toubiana, W., & Wittkopp, P. J. (2017). Fitness effects of cis-regulatory variants in the *Saccharomyces cerevisiae* TDH3 promoter. *Molecular Biology and Evolution, 34*, 2908–2912. https://doi.org/10.1093/molbev/msx224

Dykhuizen, D. E., Dean, A. M., & Hartl, D. L. (1987). Metabolic flux and fitness. *Genetics, 115*, 25–31.

Emerson, R. O., & Thomas, J. H. (2009). Adaptive evolution in zinc finger transcription factors. *PLoS Genetics, 5*, e1000325. https://doi.org/10.1371/journal.pgen.1000325

Eyre-Walker, A., & Keightley, P. D. (2007). The distribution of fitness effects of new mutations. *Nature Reviews. Genetics, 8*, 610–618. https://doi.org/10.1038/nrg2146

Ezer, D., Zabet, N. R., & Adryan, B. (2015). Homotypic clusters of transcription factor binding sites: A model system for understanding the physical mechanics of gene expression. *Computational and Structural Biotechnology Journal, 10*, 63–69. https://doi.org/10.1016/j.csbj.2014.07.005

Fares, M. A. (2015). The origins of mutational robustness. *Trends in Genetics, 31*, 373–381. https://doi.org/10.1016/j.tig.2015.04.008

Galant, R., & Carroll, S. B. (2002). Evolution of a transcriptional repression domain in an insect Hox protein. *Nature, 415*, 910–913. https://doi.org/10.1038/nature717

Gama-Castro, S., Salgado, H., Peralta-Gil, M., et al. (2011). RegulonDB version 7.0: Transcriptional regulation of *Escherichia coli* K-12 integrated within genetic sensory response units (Gensor Units). *Nucleic Acids Research, 39*, D98–D105. https://doi.org/10.1093/nar/gkq1110

Garfield, D., Haygood, R., Nielsen, W. J., & Wray, G. A. (2012). Population genetics of cis-regulatory sequences that operate during embryonic development in the sea urchin *Strongylocentrotus purpuratus*. *Evolution & Development, 14*, 152–167. https://doi.org/10.1111/j.1525-142X.2012.00532.x

Gerdes, S. Y., Scholle, M. D., Campbell, J. W., et al. (2003). Experimental determination and system level analysis of essential genes in *Escherichia coli* MG1655. *Journal of Bacteriology, 185*, 5673–5684. https://doi.org/10.1128/JB.185.19.5673

Giaever, G., Chu, A. M., Ni, L., et al. (2002). Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature, 418*, 387–391. https://doi.org/10.1038/nature00935

Giniger, E., & Ptashne, M. (1988). Cooperative DNA binding of the yeast transcriptional activator GAL4. *Proceedings of the National Academy of Sciences of the United States of America, 85*, 382–386.

Gotea, V., Visel, A., Westlund, J. M., et al. (2010). Homotypic clusters of transcription factor binding sites are a key component of human promoters and enhancers. *Genome Research, 20*, 565–577. https://doi.org/10.1101/gr.104471.109

Gu, Z., Steinmetz, L. M., Gu, X., et al. (2003). Role of duplicate genes in genetic robustness against null mutations. *Nature, 421*, 63–66. https://doi.org/10.1038/nature01198

Guttman, M., & Rinn, J. L. (2012). Modular regulatory principles of large non-coding RNAs. *Nature, 482*, 339–346. https://doi.org/10.1038/nature10887

Harismendy, O., Notani, D., Song, X., et al. (2011). 9p21 DNA variants associated with coronary artery disease impair interferon-γ 3 signalling response. *Nature, 470*, 264–270. https://doi.org/10.1038/nature09753

Hart, T., Chandrashekhar, M., Aregger, M., et al. (2015). High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. *Cell, 163*, 1515–1526. https://doi.org/10.1016/j.cell.2015.11.015

He, X., & Zhang, J. (2006). Transcriptional reprogramming and backup between duplicate genes: Is it a genomewide phenomenon? *Genetics, 172*, 1363–1367. https://doi.org/10.1534/genetics.105.049890

Hillenmeyer, M. E., Fung, E., Wildenhain, J., et al. (2008). The chemical genomic portrait of yeast: Uncovering a phenotype for all genes. *Science, 320*, 362–365. https://doi.org/10.1126/science.1150021

Hollenhorst, P. C., Shah, A. A., Hopkins, C., & Graves, B. J. (2007). Genome-wide analyses reveal properties of redundant and specific promoter occupancy within the ETS gene family. *Genes & Development, 21*, 1882–1894. https://doi.org/10.1101/gad.1561707

Igler, C., Lagator, M., Tkačik, G., et al. (2018). Evolutionary potential of transcription factors for gene regulatory rewiring. *Nature Ecology and Evolution, 2*, 1633–1643. https://doi.org/10.1038/s41559-018-0651-y

Irish, V. F. (2003). The evolution of floral homeotic gene function. *BioEssays, 25*, 637–646. https://doi.org/10.1002/bies.10292

Itzkovitz, S., Tlusty, T., & Alon, U. (2006). Coding limits on the number of transcription factors. *BMC Genomics, 7*, 1–15. https://doi.org/10.1186/1471-2164-7-239

Jaeger, J. (2018). Shift happens: The developmental and evolutionary dynamics of the gap gene system. *Current Opinion in Systems Biology, 11*, 65–73. https://doi.org/10.1016/j.coisb.2018.08.004

Jiménez, A., Munteanu, A., Sharpe, J., et al. (2015). Dynamics of gene circuits shapes evolvability. *Proceedings of the National Academy of Sciences of the United States of America, 112*, 2103–2108. https://doi.org/10.1073/pnas.1411065112

Jimenez, J. I., Xulvi-Brunet, R., Campbell, G. W., et al. (2013). Comprehensive experimental fitness landscape and evolutionary network for small RNA. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 14984–14989. https://doi.org/10.1073/pnas.1307604110

Johnson, A. D., Meyer, B. J., & Ptashne, M. (1979). Interactions between DNA-bound repressors govern regulation by the lambda phage repressor. *Proceedings of the National Academy of Sciences of the United States of America, 76*, 5061–5065.

Kafri, R., Bar-Even, A., & Pilpel, Y. (2005). Transcription control reprogramming in genetic backup circuits. *Nature Genetics, 37*, 295–299. https://doi.org/10.1038/ng1523

Kasowski, M., Grubert, F., Heffelfinger, C., et al. (2010). Variation in transcription factor binding among humans. *Science, 328*, 232–235. https://doi.org/10.1126/science.1183621

Katainen, R., Dave, K., Pitkänen, E., et al. (2015). CTCF/cohesin-binding sites are frequently mutated in cancer. *Nature Genetics, 47*, 818–821. https://doi.org/10.1038/ng.3335

Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology, 22*, 437–467. https://doi.org/10.1354/vp.45-1-111

Keane, O. M., Toft, C., Carretero-Paulet, L., et al. (2014). Preservation of genetic and regulatory robustness in ancient gene duplicates of *Saccharomyces cerevisiae*. *Genome Research, 24*, 1830–1841. https://doi.org/10.1101/gr.176792.114

Keren, H., Lev-Maor, G., & Ast, G. (2010). Alternative splicing and evolution: Diversification, exon definition and function. *Nature Reviews. Genetics, 11*, 345–355. https://doi.org/10.1038/nrg2776

Keren, L., Hausser, J., Lotan-Pompan, M., et al. (2016). Massively parallel interrogation of the effects of gene expression levels on fitness. *Cell, 166*, 1282–1294.e18. https://doi.org/10.1016/j.cell.2016.07.024

Khalid, F., Aguilar-Rodríguez, J., Wagner, A., & Payne, J. L. (2016). Genonets server–a web server for the construction, analysis and visualization of genotype networks. *Nucleic Acids Research, 44*, W70–W76.

Khurana, E., Fu, Y., Colonna, V., et al. (2013). Integrative annotation of variants from 1092 humans: Application to cancer genomics. *Science, 342*, 1235587. https://doi.org/10.1126/science.1235587

Kilpinen, H., Waszak, S. M., Gschwind, A. R., et al. (2013). Coordinated effects of sequence variation on DNA binding, chromatin structure, and transcription. *Science, 342*, 744–747. https://doi.org/10.1126/science.1242463

Krivan, W., & Wasserman, W. W. (2001). A predictive model for regulatory sequences directing liver-specific transcription. *Genome Research, 11*, 1559–1566. https://doi.org/10.1101/gr.180601

Lachowiec, J., Lemus, T., Thomas, J. H., et al. (2013). The protein chaperone HSP90 can facilitate the divergence of gene duplicates. *Genetics, 193*, 1269–1277. https://doi.org/10.1534/genetics.112.148098

Lagator, M., Sarikas, S., Acar, H., et al. (2017). Regulatory network structure determines patterns of intermolecular epistasis. *eLife, 6*, 1–22. https://doi.org/10.7554/eLife.28921

Latchman, D. S. (2008). *Eukaryotic transcription factors*. Elsevier Science.

Lawrence, P. A. (1992). *The making of a fly: The genetics of animal design*. Wiley-Blackwell.

Lee, T. I., & Young, R. A. (2013). Transcriptional regulation and its misregulation in disease. *Cell, 152*, 1237–1251. https://doi.org/10.1016/j.cell.2013.02.014

Lehner, B. (2010). Genes confer similar robustness to environmental, stochastic, and genetic perturbations in yeast. *PLoS One, 5*, e9035. https://doi.org/10.1371/journal.pone.0009035

Levine, M., & Tjian, R. (2003). Transcription regulation and animal diversity. *Nature, 424*, 147–151. https://doi.org/10.1038/nature01763

Li, P., Markson, J. S., Wang, S., et al. (2018). Morphogen gradient reconstitution reveals Hedgehog pathway design principles. *Science, 360*, 543–548. https://doi.org/10.1126/science.aao0645

Li, W., Zhu, Z., Chern, M., et al. (2017). A natural allele of a transcription factor in rice confers broad-spectrum blast resistance. *Cell, 170*, 114–126.e15. https://doi.org/10.1016/j.cell.2017.06.008

Lifanov, A. P., Makeev, V. J., Nazina, A. G., & Papatsenko, D. A. (2003). Homotypic regulatory clusters in Drosophila. *Genome Research, 13*, 579–588. https://doi.org/10.1101/gr.668403

Liu, Q., Onal, P., Datta, R. R., et al. (2018). Ancient mechanisms for the evolution of the bicoid homeodomain's function in fly development. *eLife, 7*, 1–28. https://doi.org/10.7554/eLife.34594

Liu, Y., Walavalkar, N. M., Dozmorov, M. G., et al. (2017). Identification of breast cancer associated variants that modulate transcription factor binding. *PLoS Genetics, 13*, 1–21. https://doi.org/10.1371/journal.pgen.1006761

Ludwig, M. Z., Bergman, C., Patel, N. H., & Kreitman, M. (2000). Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature, 403*, 564–567. https://doi.org/10.1038/35000615

Lynch, V. J., May, G., & Wagner, G. P. (2011). Regulatory evolution through divergence of a phosphoswitch in the transcription factor CEBPB. *Nature, 480*, 383–386. https://doi.org/10.1038/nature10595

Lynch, V. J., Tanzer, A., Wang, Y., et al. (2008). Adaptive changes in the transcription factor HoxA-11 are essential for the evolution of pregnancy in mammals. *Proceedings of the National Academy of Sciences, 105*, 14928–14933. https://doi.org/10.1073/pnas.0802355105

Madan Babu, M., Teichmann, S. A., & Aravind, L. (2006). Evolutionary dynamics of prokaryotic transcriptional regulatory networks. *Journal of Molecular Biology, 358*, 614–633. https://doi.org/10.1016/j.jmb.2006.02.019

Maerkl, S. J., & Quake, S. R. (2009). Experimental determination of the evolvability of a transcription factor. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 18650–18655. https://doi.org/10.1073/pnas.0907688106

Majithia, A. R., Tsuda, B., Agostini, M., et al. (2016). Prospective functional classification of all possible missense variants in PPARG. *Nature Genetics, 48*, 1570–1575. https://doi.org/10.1038/ng.3700

Masel, J., & Trotter, M. V. (2010). Robustness and evolvability. *Trends in Genetics, 26*, 406–414. https://doi.org/10.1016/j.tig.2010.06.002

Mathias, J. R., Zhong, H., Jin, Y., & Vershon, A. K. (2001). Altering the DNA-binding specificity of the yeast Matalpha 2 homeodomain protein. *The Journal of Biological Chemistry, 276*, 32696–32703. https://doi.org/10.1074/jbc.M103097200

Maurano, M. T., Humbert, R., Rynes, E., et al. (2012). Systematic localization of common disease-associated variation in regulatory DNA. *Science, 337*, 1190–1195. https://doi.org/10.1126/science.1222794

Maynard Smith, J. (1970). Natural selection and the concept of a protein space. *Nature, 225*, 563–564. https://doi.org/10.1038/225563a0

Melton, C., Reuter, J. A., Spacek, D. V., & Snyder, M. (2015). Recurrent somatic mutations in regulatory regions of human cancer genomes. *Nature Genetics, 47*, 710–716. https://doi.org/10.1038/ng.3332

Morgenstern, B., & Atchley, W. R. (2018). Evolution of bHLH transcription factors: Modular evolution by domain shuffling? *Molecular Biology and Evolution, 16*, 1654–1663.

Munro, D., Ghersi, D., & Singh, M. (2018). Two critical positions in zinc finger domains are heavily mutated in three human cancer types. *PLoS Computational Biology, 14*, 1–17. https://doi.org/10.1371/journal.pcbi.1006290

Musunuru, K., Strong, A., Frank-Kamenetsky, M., et al. (2010). From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. *Nature, 466*, 714–719. https://doi.org/10.1038/nature09266

Nadimpalli, S., Persikov, A. V., & Singh, M. (2015). Pervasive variation of transcription factor orthologs contributes to regulatory network evolution. *PLoS Genetics, 11*, e1005011. https://doi.org/10.1371/journal.pgen.1005011

Najafabadi, H. S., Garton, M., Weirauch, M. T., et al. (2017). Non-base-contacting residues enable kaleidoscopic evolution of metazoan C2H2 zinc finger DNA binding. *Genome Biology, 18*, 1–15. https://doi.org/10.1186/s13059-017-1287-y

Nocedal, I., Mancera, E., & Johnson, A. D. (2017). Gene regulatory network plasticity predates a switch in function of a conserved transcription regulator. *eLife, 6*, 1–20. https://doi.org/10.7554/elife.23250

Noyes, M. B., Christensen, R. G., Wakabayashi, A., et al. (2008). Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell, 133*, 1277–1289. https://doi.org/10.1016/j.cell.2008.05.023

Odom, D. T., Dowell, R. D., Jacobsen, E. S., et al. (2007). Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nature Genetics, 39*, 730–732. https://doi.org/10.1038/ng2047

Pabo, C. O., Peisach, E., & Grant, R. A. (2001). Design and selection of novel Cys2His2 zinc finger proteins. *Annual Review of Biochemistry, 70*, 313–340. https://doi.org/10.1146/annurev.biochem.70.1.313

Pauli, A., Rinn, J. L., & Schier, A. F. (2011). Non-coding RNAs as regulators of embryogenesis. *Nature Reviews. Genetics, 12*, 136–149. https://doi.org/10.1038/nrg2904

Payne, J. L., Khalid, F., & Wagner, A. (2018). RNA-mediated gene regulation is less evolvable than transcriptional regulation. *Proceedings of the National Academy of Sciences of the United States of America, 115*, E3481–E3490. https://doi.org/10.1073/pnas.1719138115

Payne, J. L., & Wagner, A. (2014). The robustness and evolvability of transcription factor binding sites. *Science, 343*, 875–877. https://doi.org/10.1126/science.1249046

Payne, J. L., & Wagner, A. (2015). Mechanisms of mutational robustness in transcriptional regulation. *Frontiers in Genetics, 6*, 322. https://doi.org/10.3389/fgene.2015.00322

Payne, J. L., & Wagner, A. (2019). The causes of evolvability and their evolution. *Nature Reviews. Genetics, 20*, 24–38. https://doi.org/10.1038/s41576-018-0069-z

Pelechano, V., & Steinmetz, L. M. (2013). Gene regulation by antisense transcription. *Nature Reviews. Genetics, 14*, 880–893. https://doi.org/10.1038/nrg3594

Perez, J. C., Fordyce, P. M., Lohse, M. B., et al. (2014). How duplicated transcription regulators can diversify to govern the expression of nonoverlapping sets of genes. *Genes & Development, 28*, 1272–1277. https://doi.org/10.1101/gad.242271.114

Pfeiffer, B. D., Berman, B. P., Levine, M., et al. (2002). Exploiting transcription factor binding site clustering to identify cis-regulatory modules involved in pattern formation in the *Drosophila* genome. *Proceedings of the National Academy of Sciences, 99*, 757–762. https://doi.org/10.1073/pnas.231608898

Pomerantz, M. M., Ahmadiyeh, N., Jia, L., et al. (2009). The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nature Genetics, 41*, 882–884. https://doi.org/10.1038/ng.403

Prud'homme, B., Gompel, N., & Carroll, S. B. (2007). Emerging principles of regulatory evolution. *Proceedings of the National Academy of Sciences, 104*, 8605–8612. https://doi.org/10.1073/pnas.0700488104

Ptashne, M., & Gann, A. (2002). *Genes & signals*. Cold Spring Harbor Laboratory Press.

Ramani, A. K., Chuluunbaatar, T., Verster, A. J., et al. (2012). The majority of animal genes are required for wild-type fitness. *Cell, 148*, 792–802. https://doi.org/10.1016/j.cell.2012.01.019

Raspopovic, J., Marcon, L., Russo, L., & Sharpe, J. (2014). Modeling digits. Digit patterning is controlled by a Bmp-Sox9-Wnt Turing network modulated by morphogen gradients. *Science, 345*, 566–570. https://doi.org/10.1126/science.1252960

Rest, J. S., Morales, C. M., Waldron, J. B., et al. (2013). Nonlinear fitness consequences of variation in expression level of a eukaryotic gene. *Molecular Biology and Evolution, 30*, 448–456. https://doi.org/10.1093/molbev/mss248

Ronshaugen, M., McGinnis, N., & McGinnis, W. (2002). Hox protein mutation and macroevolution of the insect body plan. *Nature, 415*, 914–917. https://doi.org/10.1038/nature716

Rowe, W., Platt, M., Wedge, D. C., et al. (2010). Analysis of a complete DNA-protein affinity landscape. *Journal of the Royal Society Interface, 7*, 397–408. https://doi.org/10.1098/rsif.2009.0193

Ruiz-Trillo, I., de Mendoza, A., Torruella, G., et al. (2013). Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *Proceedings of the National Academy of Sciences, 110*, E4858–E4866. https://doi.org/10.1073/pnas.1311818110

Schaerli, Y., Jiménez, A., Duarte, J. M., et al. (2018). Synthetic circuits reveal how mechanisms of gene regulatory networks constrain evolution. *Molecular Systems Biology, 14*, e8102. https://doi.org/10.15252/msb.20178102

Schmitz, J. F., Zimmer, F., & Bornberg-Bauer, E. (2016). Mechanisms of transcription factor evolution in Metazoa. *Nucleic Acids Research, 44*, 6287–6297. https://doi.org/10.1093/nar/gkw492

Sengupta, A. M., Djordjevic, M., & Shraiman, B. I. (2002). Specificity and robustness in transcription control networks. *Proceedings of the National Academy of Sciences, 99*, 2072–2077. https://doi.org/10.1073/pnas.022388499

Sharon, E., Chen, S. A., Khosla, N. M., et al. (2018). Functional genetic variants revealed by massively parallel precise genome editing. *Cell, 175*, 544–557.e16. https://doi.org/10.1016/j.cell.2018.08.057

Sharon, E., Kalma, Y., Sharp, A., et al. (2012). Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nature Biotechnology, 30*, 521–530. https://doi.org/10.1038/nbt.2205

Short, P. J., McRae, J. F., Gallone, G., et al. (2018). De novo mutations in regulatory elements in neurodevelopmental disorders. *Nature, 555*, 611–616. https://doi.org/10.1038/nature25983

Shultzaberger, R. K., Malashock, D. S., Kirsch, J. F., & Eisen, M. B. (2010). The fitness landscapes of cis-acting binding sites in different promoter and environmental contexts. *PLoS Genetics, 6*, e1001042. https://doi.org/10.1371/journal.pgen.1001042

Smith, Z. D., & Meissner, A. (2013). DNA methylation: Roles in mammalian development. *Nature Reviews. Genetics, 14*, 204–220. https://doi.org/10.1038/nrg3354

Somma, M. P., Pisano, C., & Lavia, P. (1991). The housekeeping promoter from the mouse CpG island HTF9 contains multiple protein-binding elements that are functionally redundant. *Nucleic Acids Research, 19*, 2817–2824. https://doi.org/10.1093/nar/19.11.2817

Sorrells, T. R., Johnson, A. N., Howard, C. J., et al. (2018). Intrinsic cooperativity potentiates parallel cis-regulatory evolution. *eLife, 7*, 1–29. https://doi.org/10.7554/eLife.37563

Spivakov, M., Akhtar, J., Kheradpour, P., et al. (2012). Analysis of variation at transcription factor binding sites in *Drosophila* and humans. *Genome Biology, 13*, R49. https://doi.org/10.1186/gb-2012-13-9-r49

Staller, M. V., Holehouse, A. S., Swain-Lenz, D., et al. (2018). A high-throughput mutational scan of an intrinsically disordered acidic transcriptional activation domain. *Cell Systems, 6*, 444–455. https://doi.org/10.1016/j.cels.2018.01.015

Starr, T. N., Picton, L. K., & Thornton, J. W. (2017). Alternative evolutionary histories in the sequence space of an ancient protein. *Nature, 549*, 409–413. https://doi.org/10.1038/nature23902

Starr, T. N., & Thornton, J. W. (2016). Epistasis in protein evolution. *Protein Science, 25*, 1204–1218. https://doi.org/10.1002/pro.2897

Stauber, M., Jäckle, H., & Schmidt-Ott, U. (1999). The anterior determinant bicoid of *Drosophila* is a derived Hox class 3 gene. *Proceedings of the National Academy of Sciences of the United States of America, 96*, 3786–3789. https://doi.org/10.1073/pnas.96.7.3786

Stern, D. L. (2000). Evolutionary developmental biology and the problem of variation. *Evolution, 54*, 1079–1091. https://doi.org/10.1111/j.0014-3820.2000.tb00544.x

Stern, D. L., & Orgogozo, V. (2008). The loci of evolution: How predictable is genetic evolution? *Evolution (N Y), 62*, 2155–2177. https://doi.org/10.1111/j.1558-5646.2008.00450.x

Stewart, A. J., Hannenhalli, S., & Plotkin, J. B. (2012). Why transcription factor binding sites are ten nucleotides long. *Genetics, 192*, 973–985. https://doi.org/10.1534/genetics.112.143370

Stormo, G. D. (2013). *Introduction to protein-DNA interactions: Structure, thermodynamics, and bioinformatics*. Cold Spring Harbor Laboratory Press.

Stormo, G. D., & Zhao, Y. (2010). Determining the specificity of protein-DNA interactions. *Nature Reviews. Genetics, 11*, 751–760. https://doi.org/10.1038/nrg2845

Taipale, M., Jarosz, D. F., & Lindquist, S. (2010). HSP90 at the hub of protein homeostasis: Emerging mechanistic insights. *Nature Reviews. Molecular Cell Biology, 11*, 515–528. https://doi.org/10.1038/nrm2918

Teichmann, S. A., & Babu, M. M. (2004). Gene regulatory network growth by duplication. *Nature Genetics, 36*, 492–496. https://doi.org/10.1038/ng1340

Thanos, D., & Maniatis, T. (1995). Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell, 83*, 1091–1100. https://doi.org/10.1016/0092-8674(95)90136-1

Tian, B., & Manley, J. L. (2016). Alternative polyadenylation of mRNA precursors. *Nature Reviews. Molecular Cell Biology, 18*, 18–30. https://doi.org/10.1038/nrm.2016.116

Tischler, J., Lehner, B., Chen, N., & Fraser, A. G. (2006). Combinatorial RNA interference in *Caenorhabditis elegans* reveals that redundancy between gene duplicates can be maintained for more than 80 million years of evolution. *Genome Biology, 7*, 1–13. https://doi.org/10.1186/gb-2006-7-8-r69

Toll-Riera, M., & Albà, M. M. (2013). Emergence of novel domains in proteins. *BMC Evolutionary Biology, 13*, 47. https://doi.org/10.1186/1471-2148-13-47

Tsong, A. E., Tuch, B. B., Li, H., & Johnson, A. D. (2006). Evolution of alternative transcriptional circuits with identical logic. *Nature, 443*, 415–420. https://doi.org/10.1038/nature05099

Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., & Luscombe, N. M. (2009). A census of human transcription factors: Function, expression and evolution. *Nature Reviews. Genetics, 10*, 252–263. https://doi.org/10.1038/nrg2538

Veraksa, A., Del Campo, M., & McGinnis, W. (2000). Developmental patterning genes and their conserved functions: From model organisms to humans. *Molecular Genetics and Metabolism, 69*, 85–100. https://doi.org/10.1006/mgme.2000.2963

Wagner, A. (1996). Does evolutionary plasticity evolve? *Evolution (N Y), 50*, 1008. https://doi.org/10.2307/2410642

Wagner, A. (2005). *Robustness and evolvability in living systems*. Princeton University Press.

Wagner, A. (2008). Robustness and evolvability: A paradox resolved. *Proceedings of the Royal Society B, 275*, 91–100. https://doi.org/10.1098/rspb.2007.1137

Wagner, A. (2011). *The origins of evolutionary innovations: A theory of transformative change in living systems*. Oxford University Press.

Wasserman, W. W., & Fickett, J. W. (1998). Identification of regulatory regions which confer muscle-specific gene expression. *Journal of Molecular Biology, 278*, 167–181. https://doi.org/10.1006/jmbi.1998.1700

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*, 440–442. https://doi.org/10.1038/30918

Weinhold, N., Jacobsen, A., Schultz, N., et al. (2014). Genome-wide analysis of noncoding regulatory mutations in cancer. *Nature Genetics, 46*, 1160–1165. https://doi.org/10.1038/ng.3101

Weirauch, M. T., & Hughes, T. R. (2010). Conserved expression without conserved regulatory sequence: The more things change, the more they stay the same. *Trends in Genetics, 26*, 66–74. https://doi.org/10.1016/j.tig.2009.12.002

Weirauch, M. T., & Hughes, T. R. (2011). A catalogue of eukaryotic transcription factor types, their evolutionary origin, and species distribution. In T. R. Hughes (Ed.), *A Handbook of transcription factors* (pp. 25–73). Springer.

Weirauch, M. T., Yang, A., Albu, M., et al. (2014). Determination and inference of eukaryotic transcription factor sequence specificity. *Cell, 158*, 1431–1443. https://doi.org/10.1016/j.cell.2014.08.009

Weiss, K. M., & Fullerton, S. M. (2000). Minireview: Phenogenetic drift and the evolution of the genotype–phenotype relationships. *Theoretical Population Biology, 57*, 187–195.

Wong, K. C., Chan, T. M., Peng, C., et al. (2013). DNA motif elucidation using belief propagation. *Nucleic Acids Research, 41*, 1–12. https://doi.org/10.1093/nar/gkt574

Wotton, K. R., Jiménez-Guri, E., Crombach, A., et al. (2015). Quantitative system drift compensates for altered maternal inputs to the gap gene network of the scuttle fly *Megaselia abdita*. *eLife, 4*, 1–28. https://doi.org/10.7554/eLife.04785

Wray, G. A. (2007). The evolutionary significance of cis-regulatory mutations. *Nature Reviews. Genetics, 8*, 206–216. https://doi.org/10.1038/nrg2063

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. In D. F. Jones (Ed.), *Proceedings of the Sixth International Congress on Genetics* (pp. 356–366). The Genetics Society of America.

Zhang, J. (2003). Evolution by gene duplication: An update. *Trends in Ecology & Evolution, 18*, 292–298. https://doi.org/10.1016/S0169-5347(03)00033-8

Zhao, B. S., Roundtree, I. A., & He, C. (2016). Post-transcriptional gene regulation by mRNA modifications. *Nature Reviews. Molecular Cell Biology, 18*, 31–42. https://doi.org/10.1038/nrm.2016.132

Zheng, W., Gianoulis, T. A., Karczewski, K. J., et al. (2011). Regulatory variation within and between species. *Annual Review of Genomics and Human Genetics, 12*, 327–346. https://doi.org/10.1146/annurev-genom-082908-150139

Zheng, W., Zhao, H., Mancera, E., et al. (2010). Genetic analysis of variation in transcription factor binding in yeast. *Nature, 464*, 1187–1191. https://doi.org/10.1038/nature08934