ANNUAL REVIEWS

Annual Review of Genetics

The 3D-Evo Space: Evolution of Gene Expression and Alternative Splicing Regulation

Federica Mantica[1] and Manuel Irimia[1,2,3]

[1]Center for Genomic Regulation, Barcelona Institute of Science and Technology, Barcelona, Spain; email: mirimia@gmail.com

[2]Universitat Pompeu Fabra, Barcelona, Spain

[3]ICREA, Barcelona, Spain

Annu. Rev. Genet. 2022. 56:315–37

First published as a Review in Advance on September 2, 2022

The *Annual Review of Genetics* is online at genet.annualreviews.org

https://doi.org/10.1146/annurev-genet-071719-020653

Keywords

evo-devo, tissues, transcription, exon, gene regulatory program, bilateria

Abstract

Animal species present relatively high levels of gene conservation, and yet they display a great variety of cell type and tissue phenotypes. These diverse phenotypes are mainly specified through differential gene usage, which relies on several mechanisms. Two of the most relevant mechanisms are regulated gene transcription, usually referred to as gene expression (rGE), and regulated alternative splicing (rAS). Several works have addressed how either rGE or rAS contributes to phenotypic diversity throughout evolution, but a back-to-back comparison between the two molecular mechanisms, specifically highlighting both their common regulatory principles and unique properties, is still missing. In this review, we propose an innovative framework for the unified comparison between rGE and rAS from different perspectives: the three-dimensional (3D)-evo space. We use the 3D-evo space to comprehensively (*a*) review the molecular basis of rGE and rAS (i.e., the molecular axis), (*b*) depict the tissue-specific phenotypes they contribute to (i.e., the tissue axis), and (*c*) describe the determinants that drive the evolution of rGE and rAS programs (i.e., the evolution axis). Finally, we unify the perspectives emerging from the three axes by discussing general trends and specific examples of rGE and rAS tissue program evolution.

1. INTRODUCTION

Bilaterian: species characterized by bilateral symmetry, in at least some life stages, and three embryonic germ layers; their last common ancestor dates to 700 million years ago

The release of the first human genome assembly 21 years ago had major implications for almost all biomedical disciplines (57, 124). The subsequent determination of genomic sequences from additional species spurred several new investigation fields, among which comparative genomics is arguably the most relevant (18). However, genomic sequences alone are not enough to uncover the molecular events underlying most phenotypic differences between species (51). Species as diverse as humans and fruit flies share many genes [e.g., 60% of human disease-associated genes are conserved in the fruit fly (22)], suggesting that biological differences must greatly be due to the way these species make use of the common genomic information and how conserved genes differentially fulfill their functional potential in each species. This differential gene usage is achieved through multiple gene regulatory mechanisms, with two of the most relevant being regulated gene transcription (usually referred to simply as gene expression, rGE) and regulated alternative splicing (rAS).

rGE and rAS are two pretranslational regulatory mechanisms that shape the transcriptome in distinctive ways and have unique evolutionary origins. On one hand, rGE enables quantitative transcriptional changes, determining which genes are activated and to what extent in different contexts. All organisms within the Archaea, Bacteria, and Eukaryota domains present some form of rGE, hinting that this regulatory mechanism dates back to the last universal common ancestor (LUCA). On the other hand, rAS introduces qualitative differences in the pool of generated transcripts. More specifically, rAS mediates the differential combination of a gene's exons and introns to generate precise proportions of multiple transcript isoforms, depending on the biological context. By definition, alternative splicing (regulated or not) can only occur in genes composed of exons and introns, and therefore it likely originated during eukaryogenesis (42, 43), following the appearance of complex exon–intron gene structures (52, 94, 97). Crucially, in several cases, alternative splicing occurs simply as a consequence of spliceosomal noise or errors, without a clear biological significance (121). Thus, in this review, we restrict the focus to rAS, concentrating on those events where the generation of alternative isoforms occurs in a controlled and genetically encoded manner, more likely associated with significant biological outcomes (79, 122, 129).

Previous reviews have addressed how either rGE or rAS contributes to phenotypic diversity throughout animal evolution (39, 96, 104, 125). However, these separate studies fail to highlight how the two molecular mechanisms, despite shaping the transcriptome with unique modalities, share a deep-rooted regulatory logic at numerous levels. To fill this gap, we propose a back-to-back comparison between rGE and rAS in the context of tissue evolution across bilaterian animals, introducing an innovative framework that allows the unified discussion of different perspectives: the three-dimensional (3D)-evo space. The 3D-evo space is defined by three axes, each associated with a distinct theme (**Figure 1**). The molecular axis describes the molecular basis of rGE and rAS in the light of their common regulatory principles and the distinct modalities through which they shape the transcriptome. The tissue axis exemplifies how rGE and rAS programs, orchestrated respectively by transcription and splicing master regulators, define tissue phenotypes across bilaterian animals. The evolution axis analyzes four common features of rGE and rAS programs, related either to the master regulators [transcription factors (TFs)/splicing factors (SFs)] or to their respective targets (genes/exons), which determine the evolutionary fate of the ancestral regulatory programs. Finally, we conceptually step into the 3D-evo volume and we unify the perspectives emerging from the three axes by discussing general trends and specific examples of the evolution of rGE and rAS tissue programs across bilaterian animals.

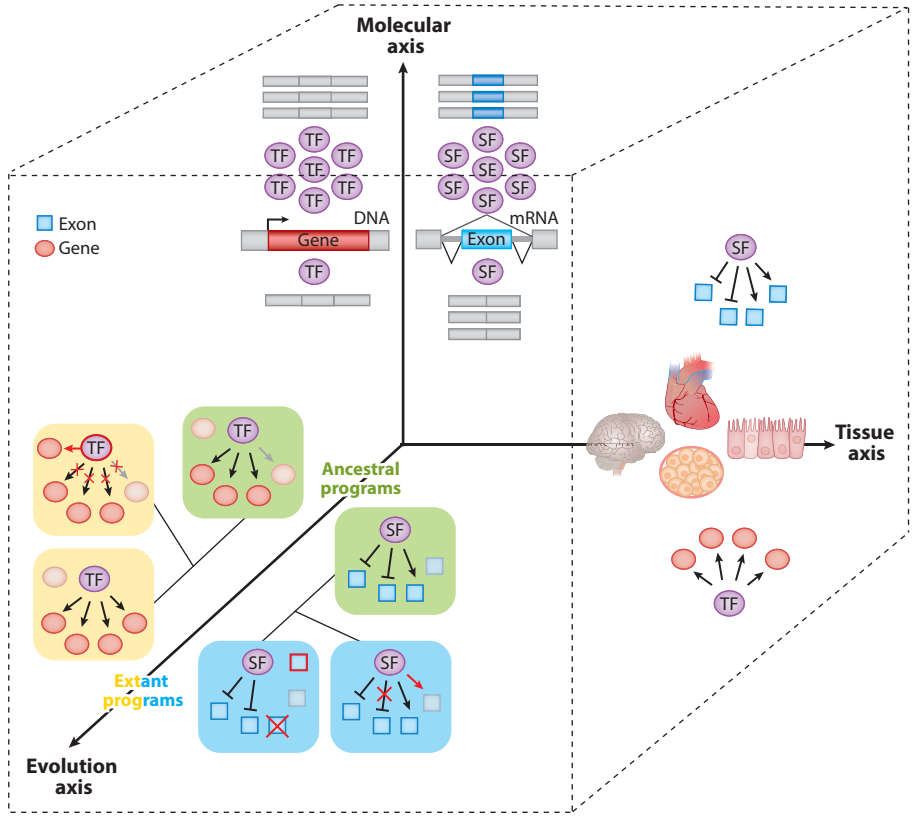

Figure 1

Schematic representation of the three-dimensional (3D)-evo space, which is defined by three conceptual axes. The molecular axis describes the molecular basis of regulated gene expression (rGE) and regulated alternative splicing (rAS), focusing on their common regulatory principles and different regulatory outcomes (quantitative versus qualitative). The tissue axis exemplifies how transcription factors (TFs) and splicing factors (SFs) respectively orchestrate rGE and rAS programs that shape tissue phenotypes in bilaterian animals. The evolution axis analyzes the determinants that control the evolutionary fate (*yellow* and *blue*) of ancestral regulatory programs (*green*). Each evolutionary scenario is described in detail in **Figure 4**.

2. THE MOLECULAR AXIS

rGE and rAS are both key molecular mechanisms for shaping the characteristic transcriptome of each cell and tissue type in multicellular organisms, even if they do it with distinctive modalities: While rGE determines the quantity of transcript generated by a given gene (i.e., how much a gene is expressed), rAS primarily shapes qualitative features (i.e., what versions, or isoforms, of a gene are expressed) (**Figure 2*a*,*b***). However, rGE and rAS largely follow the same fundamental regulatory principles: They both rely on the specific interpretation of *cis*-regulatory elements by a set of *trans*-acting factors (**Figure 2*c*,*d***). In the context of the molecular axis, we first discuss the molecular basis of rGE and rAS in the light of their common regulatory logic; then, we describe how master regulators of rGE and rAS can elegantly coordinate multiple genes and exons into functional and biologically relevant regulatory programs and explore the interconnections between the two regulatory layers.

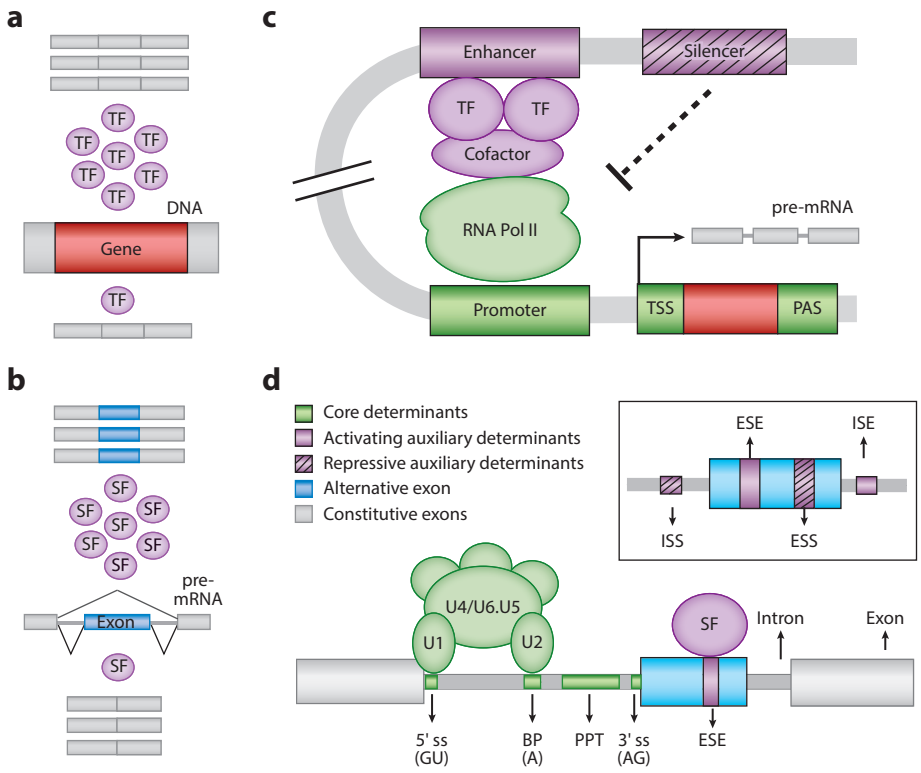

Figure 2

The molecular basis of rGE and rAS. (*a*,*b*) rGE and rAS induce (*a*) quantitative and (*b*) qualitative transcriptomic changes, respectively. (*c*) Core and auxiliary determinants necessary for rGE. Core *cis*-regulatory elements [i.e., the promoter, TSS, and PAS] and the main core *trans*-acting factor (RNA Pol II) are depicted in green. Auxiliary *cis*-regulatory elements (enhancers/silencers) and auxiliary *trans*-acting factors (TFs) appear in purple. The legend is in common with panel *d*. (*d*) Core and auxiliary determinants necessary for rAS of a cassette exon (*light blue*). Core *cis*-regulatory elements (5′ ss and 3′ ss, BP, and PPT) and the core spliceosomal components (U1, U2, and U4/U6.U5) are depicted in green. Auxiliary determinants appear in purple. The box depicts all types of auxiliary *cis*-regulatory elements (ISS, ESE, ESS, and ISE), while auxiliary *trans*-acting factors are mainly SFs. For simplicity, only TFs and SFs that positively regulate transcription or inclusion (i.e., by binding to enhancers and ESEs/ISEs, respectively) are represented. However, TFs and SFs can also negatively regulate transcription and inclusion by binding to silencers and ESSs/ISSs, respectively. Abbreviations: BP, branch point; ESE, exonic splicing enhancer; ESS, exonic splicing silencer; ISE, intronic splicing enhancer; ISS, intronic splicing silencer; PAS, polyadenylation site; PPT, polypyrimidine tract; rAS, regulated alternative splicing; rGE, regulated gene expression; RNA Pol II, RNA polymerase II; SF, splicing factor; ss, splice sites; TF, transcription factor; TSS, transcription start site.

2.1. The Molecular Basis of rGE Programs

Gene expression normally refers to the molecular process of transcription, by which a DNA region (a gene) serves as a template for the production of a pre-mRNA (a transcript) by an RNA polymerase. The pre-mRNA is further processed into mature mRNA and, in the case of protein-coding genes, translated into a protein. rGE is normally a tightly regulated process, resulting in considerable variation in the pool of expressed genes and their relative levels across biological contexts. rGE is mediated by a set of *cis*-regulatory elements that are recognized by a series of *trans*-acting factors (55), both of which can be conceptually classified into core and auxiliary determinants (**Figure 2*c***).

In complex biological contexts, core determinants are necessary for transcription to occur but are usually not sufficient to drive gene expression, while auxiliary determinants alone cannot initiate transcription but are normally responsible for the precise spatiotemporal and quantitative regulation of the process (107). Auxiliary *cis*-regulatory elements are mainly classified into enhancers and silencers depending on the effect they exert on gene expression (activation and repression, respectively) (74), and in bilaterian genomes they are usually located at great distance from their target gene. Auxiliary *cis*-regulatory elements affect gene expression only when recognized by auxiliary *trans*-factors, mainly TFs. TFs often have a DNA-binding domain, which allows them to bind the *cis*-regulatory elements, and one or more regulatory domains, which interact with other proteins driving either activation or repression of the transcriptional process, depending on the nature of the TF (for a detailed review of binding patterns and modalities of TFs, see 107). Notably, each gene is usually controlled by the combinatorial activity of several TFs and cofactors (106).

2.2. The Molecular Basis of rAS Programs

Eukaryotic genes are organized into exons (from the term expressed regions) and introns (from the term intragenic regions) (4, 27), which are both transcribed into the pre-mRNA but usually encounter different fates upon splicing: While introns are normally excised from the pre-mRNA, exons are joined together to form the mature transcript. Alternative splicing occurs when different transcript isoforms, characterized by distinct combinations of exons and introns, are generated from the same gene (27). Depending on how exons and introns are combined in the final transcript, four major types of alternative splicing are defined (49): (*a*) exon skipping, when an exon (known as a cassette exon) is fully removed from the mature transcript together with its flanking introns; (*b*) alternative 5′ splice site selection or (*c*) alternative 3′ splice site selection, when two or more splice sites are present at the 3′ or 5′ end of the exon, allowing the incorporation of slightly different versions of the same exon; and (*d*) intron retention, where an entire intron is maintained in the processed mRNA. Regardless of the type, alternative splicing can exert two major molecular functions. First, it can generate protein isoforms with slightly different sequences, with the potential to alter or fine-tune the canonical gene function in particular cellular contexts (122). These functional adjustments range from alterations of mRNA and protein molecular properties (e.g., subcellular location, stability, and DNA-/RNA-/protein-binding specificity and affinity) to modulation of multiple cell behaviors (e.g., proliferation, apoptosis, migration, and adhesion) (48). Second, rAS events can induce gene expression downregulation: The inclusion or skipping of specific sequences might disrupt the open reading frame, producing truncated proteins or introducing premature termination codons, which in turn trigger mRNA degradation through the nonsense-mediated decay pathway (34). Thus, through its qualitative regulation, rAS may in fact produce quantitative effects. While all types of rAS can potentially exert all these functions, we limit the remaining discussion to exon skipping, since it is the most common and studied rAS type in bilaterians and more often gives rise to alternative functional protein isoforms.

rAS can be defined, in a neat parallelism with rGE, by its dependency on both core and auxiliary determinants (i.e., *cis*-regulatory elements and *trans*-acting factors): Core determinants are necessary for splicing to occur, while auxiliary determinants mediate differences in splicing patterns across biological contexts (**Figure 2*d***). Auxiliary *cis*-regulatory elements are located in either exons or introns and are able to either promote (enhancers) or repress (silencers) splicing of their target regions (23). The need for auxiliary elements is particularly strong for exons with weak core splicing signals, which are more often associated with inefficient or noisy alternative splicing but can become part of highly regulated rAS programs in the presence of the right enhancers and/or silencers (19, 65, 77, 114). Splicing enhancers and silencers fulfill their regulatory potential on the splicing reaction when bound by auxiliary *trans*-acting factors, which can be roughly

subdivided into two different types: general or tissue-specific SFs. General SFs are expressed across most conditions and tissues but show varying levels of expression and activity across biological contexts. Serine- and arginine-rich (SR) proteins (63) and heterogeneous nuclear ribonucleoproteins (hnRNPs) (73) are the most relevant protein superfamilies of general SFs and contribute to the generation of unique spatiotemporal rAS patterns by establishing complex synergistic and antagonistic interactions. However, tissue-specific SFs are expressed in a tissue-restricted manner and contribute to the establishment of rAS programs exclusively in particular tissue contexts, as exemplified in the next sections.

2.3. Transcription Factors and Splicing Factors as Master Regulators

TFs and SFs share the ability to trigger unique rGE or rAS molecular programs through coordinated regulation of hundreds of functionally related targets (genes and exons, respectively), thus acting as master regulators in particular biological contexts. Remarkably, TFs and SFs can influence the behavior of their targets with two main modalities, namely (*a*) positive regulation, resulting in increased expression of regulated genes or inclusion of target exons, and (*b*) negative regulation, decreasing gene expression or inclusion levels of specific gene/exon targets (**Figure 3**). However, TFs and SFs greatly differ in their ability to carry out these two molecular functions.

On one hand, TFs are broadly classified as either activators or repressors. Thus, they usually trigger rGE programs composed of numerous genes that all change their expression in a precise direction (either increased or decreased). Examples for each of these two categories of TFs include ASCL1 (91) and REST (13, 14), an activator and a repressor, respectively (**Figure 3**). Despite the general trend, some TFs can act as either positive or negative regulators depending on the cellular context, the external stimuli, or their own expression levels (68). One example is PAX6, a homeobox TF that simultaneously activates and represses different sets of targets in the various contexts where it is expressed (112, 117).

On the other hand, SFs are normally associated with both molecular functions (i.e., positive and negative regulation). In fact, the same SFs can either promote or repress exon inclusion depending on the position of their auxiliary *cis*-regulatory elements around the target exon. This positional dependency was first reported for NOVA1 and NOVA2, two paralog SFs that were shown to generally promote or repress exon inclusion when binding downstream or upstream and within the target exons, respectively (61, 123). Since then, most SFs have been shown to have this positional regulatory effect (131, 135). However, there are some exceptions to this rule, as some SFs (nearly) exclusively act as positive (e.g., SRRM4 and its paralog SRRM3) (44) or negative (e.g., PTBP1) (40) regulators of exon inclusion (**Figure 3**). Still, the widespread dual regulatory nature of SFs implies a fundamental difference in the rAS programs triggered by SFs compared to the rGE programs orchestrated by TFs: While rGE programs generally refer to a set of genes that are either activated or repressed by a given TF, rAS programs promoted by SFs normally include target exons that change their inclusion levels in both directions (i.e., increased or decreased inclusion), becoming respectively more or less represented in the resulting pool of isoforms.

2.4. Interconnections Between rGE and rAS

TFs and SFs act as master regulators by establishing rGE and rAS programs, respectively, but the correct specification of phenotypic features often relies on the crosstalk between the two molecular layers: In fact, many SFs have been known to determine the generation of TFs transcripts, and TFs in turn regulate the transcription of target SFs (35, 64, 87). However, although cross-regulation between rGE and rAS layers is frequent, intralayer regulation is favored for both TFs and SFs (53). Moreover, the set of effector genes that undergo either one type of regulation or

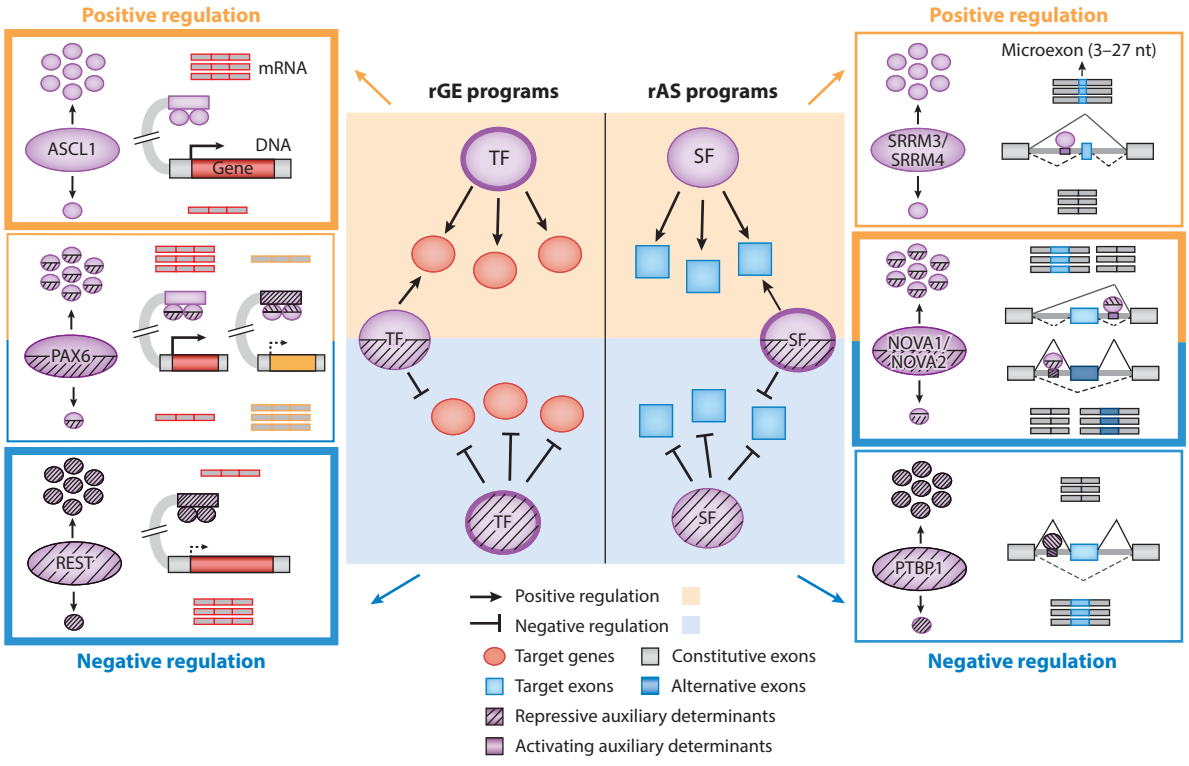

Figure 3

TF and SF master regulators orchestrate rGE and rAS programs through positive and/or negative regulation of their target genes and exons. (*Center*) Scheme of the two different types of regulation, positive (*top*) and negative (*bottom*), that master regulators, TFs (*left*) and SFs (*right*), can exert on the relative targets in order to coordinate complex regulatory programs. Some TFs and SFs exert both regulatory potentials. Thicker borders indicate the most common regulatory nature of each class of master regulators (either positive or negative for TFs, both positive and negative for SFs). Examples of master regulators for each category are shown on the left (rGE) and right (rAS) sides. (*Top left*) High levels of ASCL1 promote expression of its target genes. (*Center left*) PAX6 simultaneously activates and represses different sets of target genes. (*Bottom left*) High levels of REST inhibit the expression of its target genes. (*Top right*) SRRM3 and SRRM4 promote microexon inclusion. (*Center right*) NOVA1 and NOVA2 exhibit both positive and negative regulatory potential depending on the location of the target auxiliary *cis*-regulatory elements. (*Bottom right*) PTBP1 is a splicing repressor that negatively regulates exon inclusion of its target exons. Abbreviations: rAS, regulated alternative splicing; rGE, regulated gene expression; SF, splicing factor; TF, transcription factor.

another seems to be largely independent. In *Arabidopsis thaliana*, the sets of genes with rGE or rAS activated under stress conditions or in specific tissue contexts were shown to be highly nonoverlapping and characterized by distinctive architectural features (72). Compared to the genomic average, rGE/rAS targets present (*a*) significantly stronger/weaker splice sites, (*b*) higher/lower number of introns, and (*c*) longer/shorter promoter regions with more/fewer TF binding sites, respectively. Thus, even if the crosstalk between rGE and rAS regulatory layers is required for the correct generation of transcriptomic programs, TFs and SFs seem to preferentially regulate nonoverlapping sets of genes.

3. THE TISSUE AXIS

In this section, we move along the tissue axis of the 3D-evo space, highlighting the connections between tissue-specific TF and SF master regulators (hereafter referred to as TFs and SFs), the

regulatory programs they orchestrate, and the variety of tissue-related biological phenotypes they determine. The logic underlying these tissue-restricted regulatory programs is elegant, as it implies that the phenotypic features specified by particular master regulators are triggered only in those biological contexts where the master regulators are specifically expressed. Most of the master regulators described in the following paragraphs are conserved across bilaterians, but for simplicity we mainly discuss the tissue phenotypes they specify in human and/or mouse. However, we adopt a more general, pan-bilaterian perspective in later sections.

3.1. Tissue-Specific rGE Programs

An exemplary case of a tissue-specific rGE program is represented by the transcriptional cascade activated by ASCL1, a proneural TF that coordinates rGE programs in both proliferating and differentiating neural progenitors and facilitates expression of its target genes by inducing an increase in chromatin accessibility (91). ASCL1 plays a key role in neuronal differentiation, and it is one of three TFs able to generate functional neurons from mouse fibroblasts (126) or from human pluripotent stem cells (85) (**Figure 4*a***). The precise definition of the nervous system also relies on REST, a transcriptional repressor highly expressed in nonneuronal tissues where it inhibits genes involved in the specification of neuronal-like traits (13, 14) (**Figure 4*a***). Similar to the nervous system, the correct development and preservation of other somatic and germline tissues depend on the controlled expression of a few TFs and on the rGE programs they coordinate, as illustrated in **Figure 4*a***.

While many human TFs indeed show preferential expression exclusively in one tissue (55), many others are actually activated in a few tissue contexts where they can regulate distinct rGE programs depending on the cellular conditions. These conditions include (*a*) the pool of available TFs/cofactors, which act combinatorially to determine target activation or repression, and (*b*) the accessibility and sensitivity of the cognate *cis*-regulatory elements, which are known to be highly cell type specific (38). For instance, PAX6 (**Figure 4*b***) is highly expressed in developing and mature neurons where it is necessary for the correct specification of both the retina (37) and the central nervous system (103), often in cooperation with SOX2 TFs, which co-occupy several PAX6-bound promoters (117). However, PAX6 is also required for the precise specification of pancreatic islet cell types (100, 108), where its endocrine-specific role is mediated by the interaction with other TFs and cofactors such as MAF and NEUROD1 (47) as well as the islet-specific accessibility of several PAX6-responsive *cis*-regulatory elements (116).

3.2. Tissue-Specific rAS Programs

rAS programs are mainly established through strict modulation of the SF expression patterns, which can then regulate their respective target exons exclusively in well-defined tissue contexts. Tissue-specific programs of alternative exons are particularly abundant in the neural tissue, muscle, and testis but are also highly represented in embryonic stem, adipose, or immune cells (among others) (3, 20, 28, 115). Additionally, the tissues most enriched for rAS events are also those with the highest number of highly differentially expressed SFs (30), showing a nice correlation between tissue-specific SF expression and the emergence of tissue-specific rAS programs.

SRRM3 and SRRM4 are canonical examples of tissue-specific SFs, as they are prevalently expressed at the neural level, where they activate a rAS program of neuronal microexons (**Figure 4*c***). Neuronal microexons are very short (3–27 nucleotides long) and highly conserved exons specifically included in neural transcripts, and their misregulation has been associated with autism spectrum disorder in humans (44, 88) and impairment of central nervous system development in mice (81, 89, 90). On the contrary, PTBP1 is a negative splicing regulator

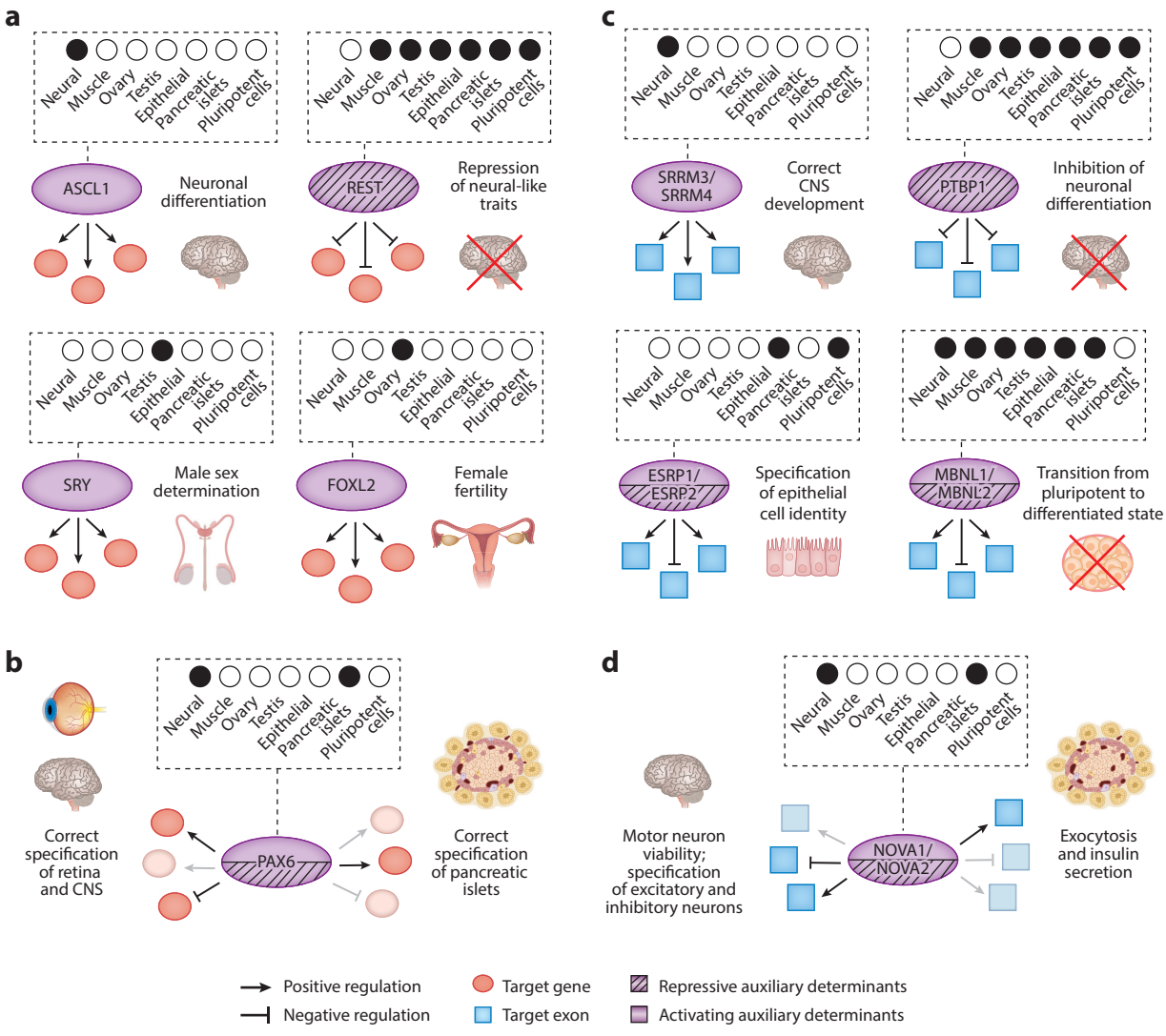

Figure 4

Tissue-specific TFs and SFs orchestrate rGE and rAS programs linked to tissue-specific phenotypes. (*a*) rGE programs orchestrated by TFs with tissue-specific expression. (*b*) A TF that orchestrates distinct rGE programs in different tissue contexts. (*c*) rAS programs orchestrated by SFs with tissue-specific expression. (*d*) An SF that orchestrates distinct rAS programs in different tissue contexts. The dashed boxes represent the expression profiles of each master regulator across tissues, where black and white circles indicate domains with high and low/no expression levels, respectively. Transparent target genes/exons represent downstream effectors not regulated in that particular context. Abbreviations: CNS, central nervous system; rAS, regulated alternative splicing; rGE, regulated gene expression; SF, splicing factor; TF, transcription factor.

highly expressed in embryonic stem cells and neuronal progenitors (as well as in most other nonneural cells), where it represses the inclusion of a specific set of neurally included exons and microexons and triggers a rAS program crucial to the inhibition of neuronal differentiation (64) (**Figure 4*c***). Additional cases of tissue-specific splicing regulators outside the nervous system are illustrated in **Figure 4*c***, including the epithelial-associated ESRP1/ESRP2 (130) or the differentiation-associated MBNL1/MBNL2 (36) SFs.

Some SFs might drive diverse rAS programs when expressed in different tissues; in fact, in a fine parallelism with rGE programs, distinct tissue-related regulatory outcomes depend on the cellular context. For instance, NOVA1 and NOVA2 (**Figure 4*d***) are highly expressed at the neural level (8, 132) and essential for motor-neuronal viability (46) (among many other functions). However, NOVA1 is also expressed in pancreatic β cells, where it orchestrates an rAS program of targets enriched for exocytosis and insulin secretion, two functions impaired upon NOVA1 silencing in both human and mouse β cells (128). Similarly, NOVA2 is expressed during angiogenesis, where it controls the formation of the vascular lumen by contributing to correct endothelial cell polarity (26). NOVA1 and NOVA2 target exons are also sometimes coregulated by other SFs, which contribute to the generation of cell type variation in rAS programs. For instance, NOVA1 and NOVA2 orchestrate different rAS programs between inhibitory and excitatory neurons, which are partially assisted by the interaction with PTBP2 (99). In other cases, NOVA1 and NOVA2 target exons are coregulated by RBFOX proteins, which share the same positional dependency and thus, depending on the relative location of their binding sites, can impact exon inclusion either synergistically or antagonistically; however, in all cases, successful regulation of the target exons only occurs in those cellular contexts where both SFs are expressed (134).

4. THE EVOLUTION AXIS

In this section, we conceptually move along the evolution axis to explore how a few key features affect the evolutionary fate of tissue-specific rGE and rAS programs controlled by master regulators. Again, we can draw a neat parallelism between the two regulatory mechanisms, as the evolution of rGE and rAS programs can be evaluated by assessing four common features. Such features are based on a conceptualization of rGE and rAS programs as networks (**Figure 5*a***) in which a central hub (the regulator) is connected to multiple terminal nodes (targets) through edges (regulatory interactions). These features thus either relate to the nature of the master regulators (hub features A and B) or are associated with their relative targets (node and edge features C and D). In particular, conservation of rGE and rAS programs occurs when diverging species conserve master regulators (TFs and SFs) with the same binding specificity and regulatory activity (hub feature A) and regulated expression pattern (hub feature B), plus the same pool of target genes/exons (node feature C) and the relative adequate *cis*-regulatory elements (edge feature D). Importantly, the master regulator–related features further strengthen the connection among all components of the 3D-evo space, as they directly refer to the regulatory nature of the TFs and SFs, which is described in the molecular axis (hub feature A), and their tissue-specific expression profiles, which are the foundation of the tissue axis (hub feature B). While the preservation of these four features is the sine qua non condition for the full evolutionary conservation of an rGE/rAS program, they also represent the raw material to act upon in order to rewire existing networks and eventually define and redefine biological traits. Importantly, a large body of work indicates that, for both rGE and rAS, changes involving master regulators (hub features A and B), often referred to as *trans* changes, are expected to be highly pleiotropic, resulting in widespread phenotype effects, and are thus less likely to occur. By contrast, changes involving individual targets, particularly their *cis*-regulatory elements (edge feature D) usually known as *cis* changes, can have more subtle and gradual effects and are thus more common. In the next paragraphs, we describe how conservation or modifications at the level of these four features shaped extant bilaterian rGE/rAS programs.

4.1. Evolution of rGE Programs

The first two features determining the evolution of rGE programs depend on the nature of the master regulators (TFs) themselves. First of all, conservation of rGE programs across species

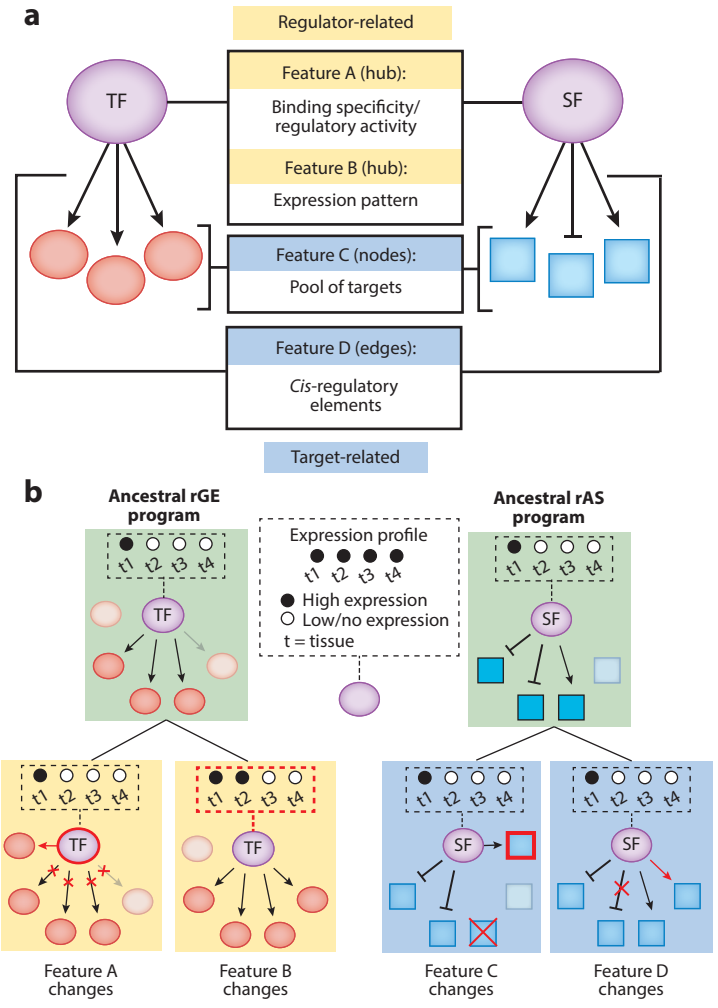

Figure 5

Evolutionary determinants of rGE and rAS programs. (*a*) Features influencing the evolution of rGE and rAS programs. Features A and B (*yellow*) depend on the master regulator's nature; features C and D (*blue*) are related to the targets. (*b*) Examples of how changes at the master regulator level [i.e., changes in binding specificity/regulatory activity (hub feature A) and expression pattern (hub feature B); *yellow*] or the target level [i.e., changes in the pool of targets (nodes feature C) and turnover of *cis*-regulatory elements (edges feature D); *blue*] induce rewiring of ancestral rGE and rAS programs (*green*). All master regulator–related changes and acquired or lost targets and regulatory interactions are highlighted in red. Transparent elements indicate existing targets that are not controlled by the master regulator and existing regulatory interactions that are not fulfilled in the tissue contexts where the master regulator is expressed. For simplicity, we represent changes in hub features A and B only for rGE programs and changes in node feature C and edge feature D only for rAS programs. However, changes in all features contribute to the evolution of both regulatory layers. Abbreviations: rAS, regulated alternative splicing; rGE, regulated gene expression; SF, splicing factor; TF, transcription factor.

implies that homologous TFs are able to bind the same set of genomic sequences and to carry out the same regulatory activity (hub feature A). The preferential recognition of a set of sequences (i.e., the binding specificity) is mediated by the DNA-binding domain, while the regulatory activity (i.e., the ultimate positive or negative regulation of the target genes) is generally carried out by the regulatory domains. The binding specificity of TFs has often been under purifying selection during bilaterian evolution (6), as demonstrated by hundreds of *Drosophila* TFs showing highly conserved binding specificities in mammals (82). However, evolutionary changes in TF binding preferences have indeed been adopted to rewire existing rGE programs, as attested by several TF families (e.g., *CH2H*) that have had extensive divergence in binding specificities in bilaterian evolution (56, 80). From a mechanistic point of view, differences in binding specificities are mainly due to structural changes in TF DNA-binding domains, which mostly arise through point mutations (80) or following transposable element (TE) insertion (16). These events can give rise to a new master regulator that preserves the original TF's regulatory activity but acquires a completely novel set of targets (**Figure 5*b***). Apart from the binding specificity, full conservation of hub feature A implies conservation of a TF's regulatory activity. Changes in a TF's regulatory activity, as well as in its ability to form protein–protein interactions, are mainly due to structural modifications to the TF's regulatory domains (50). Such modifications ultimately drive the evolution of rGE programs, but they might cause disruption of a plethora of necessary regulatory interactions and consequently be negatively selected. However, gene duplication, in the form of either single-gene duplication or whole-genome duplication (WGD), is a major force that allows the expansion of rGE programs while preserving all the original regulatory interactions, thus releasing the evolutionary constraints imposed by pleiotropy. The extra TF copies are in fact free to acquire alternative binding specificities and regulatory activities (66).

Homologous tissues: tissues that share a common ancestry, either between species (e.g., the muscle system in human and mouse) or within species (e.g., cortex and cerebellum in human)

The second master regulator–related feature affecting the evolution of tissue rGE programs is the regulated expression of the TF (hub feature B), as conservation of the TF expression profile in homologous tissues is key to the conservation of ancestral rGE programs across species. However, specific expansions of the ancestral TF expression profiles have been known to drive rGE program divergence without necessarily affecting the ancestral regulatory programs and are likely one of the major drivers of phenotypic evolution (109, 110). In many cases, new *cis*-regulatory elements enable the TF to be transcribed in novel tissues, expanding the rGE program it orchestrates while conserving all ancestral interactions (102, 133). As per feature A, this type of rGE program expansion is usually boosted by gene duplication. Gene duplicates normally undergo one of three possible fates in terms of expression profiles across tissues: (*a*) redundancy, i.e., all duplicates preserve the ancestral expression profile; (*b*) subfunctionalization, where the ancestral tissues are partitioned between the duplicates; and (*c*) specialization, with one or some of the duplicates conserving the ancestral profile but one or some of the others restricting their expression to specific tissues. Remarkably, specialization seems to be the prevalent scenario for gene duplicates derived from WGDs throughout vertebrate evolution (62, 69). In summary, divergence at the level of hub features A and B (i.e., TF binding specificities/regulatory activity and expression patterns) are ways through which homologous master regulators can potentially end up coordinating quite different molecular programs (schematized in **Figure 5*b***), particularly when gene duplication grants the preservation of ancestral networks and allows TFs to explore novel evolutionary trajectories.

The evolutionary fate of rGE programs also relies on target-related features. In principle, full conservation of an rGE program requires identical pools of target genes (node feature C); therefore, loss of ancestral targets and/or integration of novel ones generates molecular diversity between species. Indeed, gene gain and loss have been shown to be important genome remodeling forces throughout metazoan and bilaterian evolution (21, 33, 86). Notably, a fully conserved TF (hub features A and B) regulates conserved gene targets (node feature C) only when all these

targets are surrounded by the adequate auxiliary *cis*-regulatory elements, i.e., elements bearing the binding motif recognized by the master regulator (edge feature D). In some cases, *cis*-regulatory elements are conserved across large evolutionary distances (15, 98), but changes in *cis*-regulatory landscapes are arguably the most powerful and common determinants of rGE program rewiring (32, 54), often driving co-option of existing genes into novel regulatory networks (1). In general, single auxiliary *cis*-regulatory elements (enhancers/silencers) are less conserved and more plastic compared to core elements (promoters), even at short evolutionary distances (such as within mammals) (84, 127). Most recent enhancers originated through exaptation of existing sequence (127), but TE expansion also contributed to rGE programs' divergence. In particular, some TEs harbor TF binding sites that can potentially act as enhancers/silencers to fine-tune the expression of the genes surrounding the insertion locus. Remarkably, a substantial fraction of mammalian auxiliary *cis*-regulatory elements were derived from TEs, with clear genomic evidence that their TF binding sites were already present at the moment of insertion and could thus be immediately recruited into extant programs (67, 70, 95, 101). Another way through which auxiliary *cis*-regulatory elements can originate is the conversion of promoters into enhancers, which seems to have been a meaningful mechanism that was active at least throughout mammalian evolution (10, 95). Of note, the plasticity at the level of individual auxiliary *cis*-regulatory elements does not necessarily imply divergent regulatory outcomes, as it was shown that mammalian genes associated with complex regulatory landscapes (i.e., with a high number of *cis*-regulatory elements) present conserved gene expression levels independently from the conservation of the individual elements (5).

Exaptation: any adaptation that performs a function different from the original one; in the molecular context, exaptation of sequence implies the use of an existing sequence to serve a distinct purpose

4.2. Evolution of rAS Programs

As TFs and SFs establish tissue-specific molecular programs following the same fundamental principles, it is not surprising that the evolution of rAS programs is determined by the same key features that impact the evolution of rGE programs. First of all, conservation of rAS programs requires homologous SFs to preserve their ancestral binding specificity and regulatory activity (hub feature A). In this respect, a study characterizing the binding specificities for more than 200 RNA-binding proteins (RBPs) from different eukaryotic species showed that RBPs with similar RNA-binding domains (i.e., up to 70% of sequence identity) tended to recognize analogous binding motifs (92). This result indicated that RNA-binding specificities can be deduced from sequence homologies, providing useful practical knowledge to distinguish between cases in which homologous SFs possibly direct homologous rAS programs and cases where sequence divergence at the RNA-binding domain level led to divergence of the regulated targets. Conservation or divergence of the ancestral rAS programs also depends on whether homologous SFs preserve their regulated expression profile (hub feature B). Similarly to rGE programs, gene duplication offers good opportunities in terms of rAS program rewiring, although these changes in master regulators are likely scarce. In this respect, *LS2* is an interesting case of a *Drosophila*-specific SF derived from the duplication of the general splicing factor *U2AF2*, which developed a distinct, testis-specific expression profile (113).

As per rGE, the evolution of rAS programs is particularly shaped at the target level (**Figure 5*b***). Conservation of ancestral rAS programs implies conservation of the target exons (node feature C) and of their aptitude to be regulated by a particular SF (edge feature D). Features C and D in rAS programs are much more plastic compared to their rGE counterparts and are the main drivers of rAS programs' evolution. Conservation of target exons strongly depends on the conservation of exon–intron structures. Importantly, both introns and exons can be gained and lost, with different consequences for the ancestral mRNA sequence. On one hand, although intron gains and losses leave the ancestral mRNA sequence unaltered, they give rise to novel exon entities that

differentially impact isoform function when alternatively spliced. For example, an intron gain within an existing exon would generate two novel exon entities, even if the sequence they are composed of is ancestral; in the same way, intron loss mediates the fusion of two ancestral exons into a unique, novel exon entity with distinct effects upon inclusion or skipping. The frequency of changes in exon–intron architectures caused by intron gain and loss are quite variable across bilaterians, but they can be very common in certain clades (17). On the other hand, ancestral exons can be lost, and additional exons, which were not contained in the ancestral mRNA sequence, can arise in a gene following the incorporation of extra sequence. The latter mainly occurs with three modalities: (*a*) exon shuffling, through which a gene receives an existing exon (or sets of exons) from another gene; (*b*) exon duplication, where one or more of its exons undergo in loco duplication; and (*c*) exonization of genomic sequences, which often involves TEs (e.g., Alus) (49). These novel exons can be then recruited to preexistent rAS programs.

Exon entities (either novel or ancestral) can be integrated into or removed from an rAS program through the acquisition or loss of specific SF-responsive *cis*-regulatory elements that determine their predisposition to be alternatively spliced in a regulated manner. In some cases, novel or ancestral exons can be directly recruited into an rAS program through the acquisition of specific SF-responsive elements (45). In other cases, the predisposition of an exon to be alternatively spliced (for example, because of weak splice sites) precedes its integration into specific SF-directed rAS programs (9). Thus, to fully shed light on the evolutionary histories of these rAS targets, it is important to first understand how exons can eventually become alternatively spliced. For ancestral constitutive exons, this involves either weakening splice site mutations, which induce suboptimal spliceosome recognition and thus exon skipping (49), or changing other *cis*-regulatory elements. This process, referred to as exon alternativization, has been shown to occur recurrently across vertebrate evolution (76). Notably, there are several ways through which the overall characteristics of the exon–intron structure are linked to exon skipping emergence (25, 29). On the contrary, for novel exons, their genomic origin and alternative splicing pattern are usually tightly linked. The modalities through which novel exons are generated (see above) can be good predictors of their regulatory fates: Duplicated exons tend to preserve the splicing status of the original exons (i.e., duplicates of alternative exons also tend to be alternatively spliced), and Alu elements that undergo exonization generally originate directly as alternative or cryptic exons (105). In fact, these two patterns are in line with the more general observation that exon skipping is often associated with recent exon creation (58, 78, 83, 118). Regardless of their evolutionary history, alternatively spliced exons (i.e., either novel or evolved ancestrally constitutive exons) can be integrated into particular rAS programs only through a tight, SF-driven regulation mediated by the acquisition of SF-responsive *cis*-regulatory elements. These *cis*-regulatory elements underwent fast turnovers throughout bilaterian evolution, leading to rapid changes in rAS programs and several instances of homologous exons with divergent regulatory patterns (120).

5. THE THREE-DIMENSIONAL-EVO SPACE

In this section, we conceptually step into the abstract volume defined by the molecular, tissue, and evolution axes (the 3D-evo space) and combine all of their perspectives. First, we discuss general trends of rGE and rAS evolution in the context of bilaterian tissues; then, we focus on specific examples for each regulatory mechanism.

5.1. General Trends of rGE and rAS Evolution

The advent of RNA sequencing (RNA-seq) technologies in the last decade has allowed the unbiased, genome-wide reconstruction of transcriptomic landscapes across a wide range of tissues

and species, even if some microarray studies had previously provided information in this respect (31, 60). The investigation of rGE and rAS evolution involves the comparison of transcriptional profiles of orthologous entities (genes and exons, respectively) typically across homologous tissues. While many of the existing multispecies and multitissue studies investigating trends of rGE and rAS evolution are restricted to vertebrate clades, especially within mammals, they still provide unique insights that set the basis for the same investigation at the bilaterian level.

Expression shift: expression changes from the ancestral expression domain(s) to novel domain(s) that a gene can undergo in the course of evolution

Brawand et al. (7) performed one of the first evolutionary studies built upon a multispecies (10 Amniota species) and multitissue (6 tissues) RNA-seq data set. Among other groundbreaking observations, they highlighted the greater transcriptomic similarity between homologous tissues in different species compared to that between nonhomologous tissues within the same species [previously observed more limitedly (11)]. This finding implies high evolutionary conservation of rGE tissue programs and was later replicated (111) and recently expanded (12). In particular, Chen et al. (12) focused on a set of 17 mammalian species and 7 tissues across which they also modeled expression evolution. They discovered that ~95% of the mammalian-conserved genes evolve under stabilizing selection in at least one of the tissues. This finding further supports the notion that homologous tissues present highly conserved rGE programs, and researchers actually demonstrated that the more recent the separation between two tissues (i.e., the more evolutionarily related they are), the higher the level of correlated transcriptome evolution (i.e., the impact of genes that drive codependent evolution) (59).

Notwithstanding the high general conservation of rGE tissue programs among vertebrates, different tissues indeed experience distinct rates of transcriptomic evolution, with neural and testis being the slowest- and fastest-evolving tissues, respectively, in amniotes (7). These evolutionary rates are directly related to the number of genes that undergo expression shifts, as both modular expression shifts (e.g., groups of tissue-specific genes that coordinately change expression patterns) and single-gene expression shifts are overall highly abundant in testis and scarce in brain (7). Notably, modular expression shifts are likely due to alterations at the level of the master regulator (hub features A and B), which can simultaneously affect a wide range of targets, while single-gene expression shifts are probably triggered by modifications at the *cis*-regulatory level of the gene target (edge feature D). Similar frameworks led to the characterization of hundreds of tissue-specific expression shifts between primates and rodents (12) and among vertebrates (24). This characterization also highlighted the existence of tissue-specific propensities—pairs of tissues among which expression shifts occur more often than expected by chance (e.g., testis–ovary, testis–brain, and liver–kidney) (24). Importantly, the characterization of the expression shifts in all of these studies is impaired by the use of undersampled phylogenies, which led to false discovery rates of ~20% or higher, according to the authors' estimates (12, 24). However, they can still provide valuable knowledge regarding the evolutionary rewiring of tissue rGE programs, setting a starting point for case-by-case validations.

In summary, the main, common results of all of these studies is that the rGE programs of homologous tissues tend to be highly conserved in the course of vertebrate evolution, even if expression shifts and partial rewiring of tissue rGE networks could indeed be detected. Interestingly, similar investigations at the level of rAS programs led to quite different conclusions. In two milestone articles, researchers analyzed the evolution of alternative splicing profiles in 10 Tetrapoda species and 6 tissues (2) and in 10 Amniota species and 9 tissues (76), reaching the concordant conclusion that these profiles tended to be more similar between nonhomologous tissues in the same species rather than between homologous tissues in different species. These results imply low conservation of alternative splicing profiles across vertebrates and highlight the more plastic and/or noisy nature of the process compared to gene expression patterns (75). However, restricting the analysis to cassette exons that are alternatively spliced across species (i.e., most likely finely

controlled rAS events) revealed clear conservation of alternative splicing profiles across homologous tissues (2), similar to what has been observed for rGE. This trend was later confirmed for larger exon sets conserved among human, mouse, and chicken (115) and among primates (93). All together, these findings point to a meaningful degree of conservation of rAS programs, probably associated with the definition of relevant biological traits.

5.2. Cases of rGE Program Evolution

As we reported in the previous paragraph, many studies have exploited multitissue and multispecies RNA-seq data sets to identify general trends of rGE tissue evolution. While several such trends have been characterized, considerably fewer studies have systematically evaluated the evolution of distinct rGE programs in terms of the four key features we previously described (see Section 4), especially at the bilaterian level. However, there are a few works that partially fit our framework: For instance, Odom et al. (84) investigated the evolution of the rGE programs orchestrated by four tissue-specific TFs (FOXA2, HNF1A, HNF4A, and HNF6) by comparing chromatin immunoprecipitation sequencing (ChIP-seq) data between human and mouse. These four TFs present conserved binding specificity/regulatory activity (hub feature A) and conserved liver-specific expression (hub feature B) between primates and rodents, but the rGE programs that they determine strongly differ at the target level: In fact, a high percentage of orthologous regions (41–89%) bound by the human TF were not recognized by their murine counterpart (or vice versa), implying a large divergence in the pool of regulated targets among homologous genes (edge feature D). Moreover, even when orthologous target genes were regulated by the same factors in the two species, only in one-third of cases did the TF binding sites actually align, suggesting conservation of ancestral regulation (again related to edge feature D). In the other two-thirds of cases, the sites bound by homologous TFs in human and mouse do not align, hinting at a potentially convergent evolution of TF-responsive *cis*-regulatory elements.

5.3. Cases of rAS Program Evolution

Compared to rGE, fewer studies have been aimed at characterizing genome-wide patterns of rAS tissue evolution, but we possess substantially more information regarding the evolutionary history of individual tissue-specific rAS programs in terms of the four key features previously described. In this section, we focus on three SF master regulators (*SRRM3/SRRM4*, *NOVA1/NOVA2*, and *ESRP1/ESRP2*) and discuss the evolution of the tissue rAS programs they orchestrate among bilaterians.

SRRM4 is highly expressed in neurons and necessary and sufficient for the implementation of an rAS program of neuronal microexons (see Section 3). The ability of *SRRM4* to regulate the inclusion of short exons relies on the enhancer of microexons (eMIC) domain, which is shared by its vertebrate paralog *SRRM3* but not *SRRM2* (119). Analysis of the *SRRM2/SRRM3/SRRM4* ortholog in nonvertebrate bilaterian species provided insights into the evolutionary history of these master regulators and their relative rAS programs. Most nonvertebrate bilaterians have only one such ortholog (hereafter *Srrm234*), which presents an eMIC domain. Notably, the nonvertebrate *Srrm234* eMIC domain preserves the same binding specificity and regulatory activity as the vertebrate eMIC domain (119), suggesting that the bilaterian ancestors already possessed an SF master regulator with the ability to preferentially include short exons (hub feature A). Interestingly, *Srrm234* does not constitutively express the eMIC in nonvertebrate bilaterians but has an alternative C-terminal exon isoform preferentially expressed at the neural level that indeed allows the expression of the domain in neural tissues (119, 120) (hub feature B). However, conservation at the SF master regulator level does not imply conservation of ancestral rAS programs at the

target level. As we previously pointed out, changes in exon–intron structures are quite widespread among distant bilaterian lineages, even if gene conservation itself is high. Surprisingly, only 12% of fruit fly genes containing eMIC-dependent exons present a mouse ortholog also bearing an eMIC-regulated exon; moreover, only 2.5% of such exons share the same position between the two species, suggesting that they conserve an ancestral eMIC regulation (edge feature D), while the others probably represent cases of convergent exon target evolution within orthologous genes (120). Additionally, these results shed light on the high dynamicity of tissue rAS programs at the target level between distant bilaterian clades, which evolve in a parallel way even when the SF master regulator conserves its binding specificity, regulatory activity, and expression patterns. In fact, both *SRRM4* and its nonvertebrate ortholog *Srrm234* regulate few targets hosted by cytoskeleton genes, but mouse eMIC-dependent exons are preferentially located in genes involved in vesicle transport, and fruit fly eMIC-regulated exons are mainly hosted by channel-related genes (120).

Other SF master regulators useful to illustrate a case of parallel and convergent evolution of tissue rAS programs between distant bilaterian clades are *NOVA1/NOVA2*. *NOVA1* and *NOVA2* and their nonvertebrate orthologs indeed present conserved binding specificities and regulatory activities (hub feature A) and adult expression patterns [i.e., they are mainly expressed at the neural level in adults, despite broader and more divergent expression patterns during development (hub feature B)] across many bilaterians (41). However, the rAS programs they regulate are highly nonoverlapping between vertebrate and nonvertebrate species due to large divergence at the target level. While virtually 100% of vertebrate genes hosting *NOVA1/NOVA2* exon targets are conserved in closely related nonvertebrate species, only 40% of those target exons are indeed conserved (node feature C). Moreover, the nonvertebrate orthologous exons present equivalent *NOVA1/NOVA2* binding motifs (edge feature D) in only ~10% of cases. Altogether, these results indicate that the majority of the vertebrate *NOVA1/NOVA2* rAS program was not assembled before the origin of vertebrates, even if several ancestral exons have been co-opted as rAS targets (41). In addition to this parallel evolution between bilaterian clades, there was a significant level of convergent acquisition of *NOVA1/NOVA2* exon targets between vertebrates and fruit fly orthologous genes: In fact, more than 75% of the mouse *NOVA2*-dependent exons are hosted by genes whose fruit fly ortholog also contains *Nova*-regulated exons, but such exon targets were independently acquired. In other words, there is no orthology between the mouse and fruit fly *Nova*-regulated exon entities found in most orthologous genes (71). This observation of both parallel and convergent evolution was in line with what was previously observed for the rAS program orchestrated by *ESRP1/ESRP2* (9). *ESRP1* and *ESRP2* also conserve their binding specificity, regulatory activity, and epithelial-associated expression (hub features A and B) across bilaterians; however, analyses of target conservation shed further light on the distinct strategies of exon recruitment adopted at different evolutionary distances. In fact, long-distance comparisons (i.e., human versus sea urchin) show results consistent with those of *SRRM3/SRRM4* and *NOVA1/NOVA2* (i.e., mostly changes at the level of node feature C), where a significant fraction of these independently evolved target exons were convergently acquired by homologous genes. On the other hand, short-scale evolutionary comparisons (i.e., human versus mouse and zebrafish) suggest mainly the recruitment of preexisting alternative exons (i.e., changes in edge feature D).

6. CONCLUSIONS

In this review, we introduced the 3D-evo space as an innovative framework to thoroughly compare rGE and rAS from different perspectives. In the context of the molecular axis, we described the molecular basis of rGE and rAS programs in the light of their common fundamental principles and unique features, particularly focusing on TFs and SFs as master regulators of biologically

relevant programs. Through the tissue axis, we described how tissue-specific TFs and SFs orchestrate rGE and rAS programs necessary for the definition of multiple tissue traits in bilaterian animals. Moving along the evolution axis, we conceptualized rGE/rAS programs as networks and analyzed four network features (related to either the TF/SF master regulators or their target gene/exon) that determine the evolution of ancestral molecular programs. Finally, we conceptually stepped into the 3D-evo volume and unified the perspectives emerging from the three axes by discussing general trends and specific examples of rGE and rAS program evolution across bilaterian animals. We propose that future evolutionary studies of rGE and rAS will benefit from using the 3D-evo space framework, as it provides a formal infrastructure to investigate several crucial aspects of these molecular programs.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Sophie Bonnal for her critical reading of the manuscript and members of the TransDEvo lab for giving feedback on the figures. F.M. holds an FPI PhD fellowship from the Spanish Ministerio de Ciencia, associated with the grant BFU2017-89201-P to M.I.

LITERATURE CITED

1. Arnoult L, Su KFY, Manoel D, Minervino C, Magriña J, et al. 2013. Emergence and diversification of fly pigmentation through evolution of a gene regulatory module. *Science* 339(6126):1423–26
2. Barbosa-Morais NL, Irimia M, Pan Q, Xiong HY, Gueroussov S, et al. 2012. The evolutionary landscape of alternative splicing in vertebrate species. *Science* 338:1587–93
3. Bebee TW, Park JW, Sheridan KI, Warzecha CC, Cieply BW, et al. 2015. The splicing regulators Esrp1 and Esrp2 direct an epithelial splicing program essential for mammalian development. *eLife* 4:e08954
4. Berk AJ. 2016. Discovery of RNA splicing and genes in pieces. *PNAS* 113(4):801–5
5. Berthelot C, Villar D, Horvath JE, Odom DT, Flicek P. 2018. Complexity and conservation of regulatory landscapes underlie evolutionary resilience of mammalian gene expression. *Nat. Ecol. Evol.* 2(1):152–63
6. Boyle AP, Araya CL, Brdlik C, Cayting P, Cheng C, et al. 2014. Comparative analysis of regulatory information and circuits across distant species. *Nature* 512(7515):453–56
7. Brawand D, Soumillon M, Necsulea A, Julien P, Csárdi G, et al. 2011. The evolution of gene expression levels in mammalian organs. *Nature* 478(7369):343–48
8. Buckanovich RJ, Yang YY, Darnell RB. 1996. The onconeural antigen Nova-1 is a neuron-specific RNA-binding protein, the activity of which is inhibited by paraneoplastic antibodies. *J. Neurosci.* 16(3):1114–22
9. Burguera D, Marquez Y, Racioppi C, Permanyer J, Torres-Méndez A, et al. 2017. Evolutionary recruitment of flexible *Esrp*-dependent splicing programs into diverse embryonic morphogenetic processes. *Nat. Commun.* 8(1):1799
10. Carelli FN, Liechti A, Halbert J, Warnefors M, Kaessmann H. 2018. Repurposing of promoters and enhancers during mammalian evolution. *Nat. Commun.* 9(1):4066
11. Chan ET, Quon GT, Chua G, Babak T, Trochesset M, et al. 2009. Conservation of core gene expression in vertebrate tissues. *J. Biol.* 8(3):33
12. Chen J, Swofford R, Johnson J, Cummings BB, Rogel N, et al. 2019. A quantitative framework for characterizing the evolutionary history of mammalian gene expression. *Genome Res.* 29(1):53–63
13. Chen Z-F, Paquette AJ, Anderson DJ. 1998. NRSF/REST is required *in vivo* for repression of multiple neuronal target genes during embryogenesis. *Nat. Genet.* 20(2):136–42
14. Chong JA, Tàpia-Ramirez J, Kim S, Toledo-Aral JJ, Zheng Y, et al. 1995. REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. *Cell* 80(6):949–57

15. Clarke SL, VanderMeer JE, Wenger AM, Schaar BT, Ahituv N, Bejerano G. 2012. Human developmental enhancers conserved between deuterostomes and protostomes. *PLOS Genet*. 8(8):e1002852
16. Cosby RL, Judd J, Zhang R, Zhong A, Garry N, et al. 2021. Recurrent evolution of vertebrate transcription factors by transposase capture. *Science* 371(6531):eabc6405
17. Csuros M, Rogozin IB, Koonin EV. 2011. A detailed history of intron-rich eukaryotic ancestors inferred from a global survey of 100 complete genomes. *PLOS Comput. Biol.* 7(9):e1002150
18. de Crécy-Lagard V, Hanson AD. 2013. Comparative genomics. In *Brenner's Encyclopedia of Genetics*, Vol. 2, ed. S Maloy, K Hughes, pp. 102–5. London: Academic. 2nd ed.
19. Eom T, Zhang C, Wang H, Lay K, Fak J, et al. 2013. NOVA-dependent regulation of cryptic NMD exons controls synaptic protein levels after seizure. *eLife* 2:e00178
20. Fagnani M, Barash Y, Ip JY, Misquitta C, Pan Q, et al. 2007. Functional coordination of alternative splicing in the mammalian central nervous system. *Genome Biol*. 8(6):R108
21. Fernández R, Gabaldón T. 2020. Gene gain and loss across the metazoan tree of life. *Nat. Ecol. Evol.* 4(4):524–33
22. Fortini ME, Skupski MP, Boguski MS, Hariharan IK. 2000. A survey of human disease gene counterparts in the *Drosophila* genome. *J. Cell Biol.* 150(2):F23–30
23. Fu X-D, Ares M Jr. 2014. Context-dependent control of alternative splicing by RNA-binding proteins. *Nat. Rev. Genet.* 15(10):689–701
24. Fukushima K, Pollock DD. 2020. Amalgamated cross-species transcriptomes reveal organ-specific propensity in gene expression evolution. *Nat. Commun.* 11(1):4459
25. Gelfman S, Burstein D, Penn O, Savchenko A, Amit M, et al. 2012. Changes in exon–intron structure during vertebrate evolution affect the splicing pattern of exons. *Genome Res*. 22(1):35–50
26. Giampietro C, Deflorian G, Gallo S, Di Matteo A, Pradella D, et al. 2015. The alternative splicing factor Nova2 regulates vascular development and lumen formation. *Nat. Commun.* 6:8479
27. Gilbert W. 1978. Why genes in pieces? *Nature* 271(5645):501
28. Giudice J, Xia Z, Wang ET, Scavuzzo MA, Ward AJ, et al. 2014. Alternative splicing regulates vesicular trafficking genes in cardiomyocytes during postnatal heart development. *Nat. Commun.* 5:3603
29. Grau-Bové X, Ruiz-Trillo I, Irimia M. 2018. Origin of exon skipping-rich transcriptomes in animals driven by evolution of gene architecture. *Genome Biol*. 19(1):135
30. Grosso AR, Gomes AQ, Barbosa-Morais NL, Caldeira S, Thorne NP, et al. 2008. Tissue-specific splicing factor gene expression signatures. *Nucleic Acids Res*. 36(15):4823–32
31. Gu X, Su Z. 2007. Tissue-driven hypothesis of genomic evolution and sequence-expression correlations. *PNAS* 104(8):2779–84
32. Guerreiro I, Nunes A, Woltering JM, Casaca A, Nóvoa A, et al. 2013. Role of a polymorphism in a Hox/Pax-responsive enhancer in the evolution of the vertebrate spine. *PNAS* 110(26):10682–86
33. Guijarro-Clarke C, Holland PWH, Paps J. 2020. Widespread patterns of gene loss in the evolution of the animal kingdom. *Nat. Ecol. Evol.* 4(4):519–23
34. Hamid FM, Makeyev EV. 2014. Emerging functions of alternative splicing coupled with nonsense-mediated decay. *Biochem. Soc. Trans.* 42(4):1168–73
35. Han H, Braunschweig U, Gonatopoulos-Pournatzis T, Weatheritt RJ, Hirsch CL, et al. 2017. Multi-layered control of alternative splicing regulatory networks by transcription factors. *Mol. Cell* 65(3):539–53.e7
36. Han H, Irimia M, Ross PJ, Sung H-K, Alipanahi B, et al. 2013. MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature* 498(7453):241–45
37. Hanson I, Van Heyningen V. 1995. Pax6: more than meets the eye. *Trends Genet*. 11(7):268–72
38. Heinz S, Romanoski CE, Benner C, Glass CK. 2015. The selection and function of cell type-specific enhancers. *Nat. Rev. Mol. Cell Biol.* 16(3):144–54
39. Hill MS, Vande Zande P, Wittkopp PJ. 2021. Molecular and evolutionary processes generating variation in gene expression. *Nat. Rev. Genet.* 22(4):203–15
40. Hu J, Qian H, Xue Y, Fu X-D. 2018. PTB/nPTB: master regulators of neuronal fate in mammals. *Biophys. Rep.* 4(4):204–14
41. Irimia M, Denuc A, Burguera D, Somorjai I, Martín-Durán JM, et al. 2011. Stepwise assembly of the *Nova*-regulated alternative splicing network in the vertebrate brain. *PNAS* 108(13):5319–24

42. Irimia M, Rukov JL, Penny D, Roy SW. 2007. Functional and evolutionary analysis of alternatively spliced genes is consistent with an early eukaryotic origin of alternative splicing. *BMC Evol. Biol.* 7:188
43. Irimia M, Rukov JL, Penny D, Vinther J, Garcia-Fernandez J, Roy SW. 2008. Origin of introns by "intronization" of exonic sequences. *Trends Genet.* 24(8):378–81
44. Irimia M, Weatheritt RJ, Ellis JD, Parikshak NN, Gonatopoulos-Pournatzis T, et al. 2014. A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* 159(7):1511–23
45. Jelen N, Ule J, Živin M, Darnell RB. 2007. Evolution of Nova-dependent splicing regulation in the brain. *PLOS Genet.* 3(10):e173
46. Jensen KB, Dredge BK, Stefani G, Zhong R, Buckanovich RJ, et al. 2000. Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. *Neuron* 25(2):359–71
47. Katz LS, Gosmain Y, Marthinet E, Philippe J. 2009. Pax6 regulates the proglucagon processing enzyme PC2 and its chaperone 7B2. *Mol. Cell. Biol.* 29(8):2322–34
48. Kelemen O, Convertini P, Zhang Z, Wen Y, Shen M, et al. 2013. Function of alternative splicing. *Gene* 514(1):1–30
49. Keren H, Lev-Maor G, Ast G. 2010. Alternative splicing and evolution: diversification, exon definition and function. *Nat. Rev. Genet.* 11(5):345–55
50. Khor JM, Ettensohn CA. 2017. Functional divergence of paralogous transcription factors supported the evolution of biomineralization in echinoderms. *eLife* 6:e32728
51. King M-C, Wilson AC. 1975. Evolution at two levels in humans and chimpanzees. *Science* 188(4184):107–16
52. Koonin EV. 2006. The origin of introns and their role in eukaryogenesis: a compromise solution to the introns-early versus introns-late debate? *Biol. Direct* 1:22
53. Kosti I, Radivojac P, Mandel-Gutfreund Y. 2012. An integrated regulatory network reveals pervasive cross-regulation among transcription and splicing factors. *PLOS Comput. Biol.* 8(7):e1002603
54. Kwan CW, Gavin-Smyth J, Ferguson EL, Schmidt-Ott U. 2016. Functional evolution of a morphogenetic gradient. *eLife* 5:e20894
55. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, et al. 2018. The human transcription factors. *Cell* 172(4):650–65
56. Lambert SA, Yang AWH, Sasse A, Cowley G, Albu M, et al. 2019. Similarity regression predicts evolution of transcription factor sequence specificity. *Nat. Genet.* 51(6):981–89
57. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. 2001. Initial sequencing and analysis of the human genome. *Nature* 409(6822):860–921
58. Lev-Maor G, Sorek R, Shomron N, Ast G. 2003. The birth of an alternatively spliced exon: 3'splice-site selection in *Alu* exons. *Science* 300(5623):1288–91
59. Liang C, Musser JM, Cloutier A, Prum RO, Wagner GP. 2018. Pervasive correlated evolution in gene expression shapes cell and tissue type transcriptomes. *Genome Biol. Evol.* 10(2):538–52
60. Liao B-Y, Zhang J. 2006. Low rates of expression profile divergence in highly expressed genes and tissue-specific genes during mammalian evolution. *Mol. Biol. Evol.* 23(6):1119–28
61. Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, et al. 2008. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 456(7221):464–69
62. Lien S, Koop BF, Sandve SR, Miller JR, Kent MP, et al. 2016. The Atlantic salmon genome provides insights into rediploidization. *Nature* 533(7602):200–5
63. Lin S, Fu X-D. 2007. SR proteins and related factors in alternative splicing. *Adv. Exp. Med. Biol.* 623:107–22
64. Linares AJ, Lin C-H, Damianov A, Adams KL, Novitch BG, Black DL. 2015. The splicing regulator PTBP1 controls the activity of the transcription factor Pbx1 during neuronal differentiation. *eLife* 4:e09268
65. Ling JP, Chhabra R, Merran JD, Schaughency PM, Wheelan SJ, et al. 2016. PTBP1 and PTBP2 repress nonconserved cryptic exons. *Cell Rep.* 17(1):104–13
66. Louis EJ. 2007. Making the most of redundancy. *Nature* 449(7163):673–74
67. Lynch VJ, Leclerc RD, May G, Wagner GP. 2011. Transposon-mediated rewiring of gene regulatory networks contributed to the evolution of pregnancy in mammals. *Nat. Genet.* 43(11):1154–59

68. Ma J. 2005. Crossing the line between activation and repression. *Trends Genet*. 21(1):54–59
69. Marlétaz F, Firbas PN, Maeso I, Tena JJ, Bogdanovic O, et al. 2018. Amphioxus functional genomics and the origins of vertebrate gene regulation. *Nature* 564(7734):64–70
70. Marnetto D, Mantica F, Molineris I, Grassi E, Pesando I, Provero P. 2018. Evolutionary rewiring of human regulatory networks by waves of genome expansion. *Am. J. Hum. Genet.* 102(2):207–18
71. Márquez Y, Mantica F, Cozzuto L, Burguera D, Hermoso-Pulido A, et al. 2021. ExOrthist: a tool to infer exon orthologies at any evolutionary distance. *Genome Biol*. 22(1):239
72. Martín G, Márquez Y, Mantica F, Duque P, Irimia M. 2021. Alternative splicing landscapes in *Arabidopsis thaliana* across tissues and stress conditions highlight major functional differences with animals. *Genome Biol.* 22(1):35
73. Martinez-Contreras R, Cloutier P, Shkreta L, Fisette J-F, Revil T, Chabot B. 2007. hnRNP proteins and splicing control. *Adv. Exp. Med. Biol.* 623:123–47
74. Maston GA, Evans SK, Green MR. 2006. Transcriptional regulatory elements in the human genome. *Annu. Rev. Genom. Hum. Genet.* 7:29–59
75. Melé M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, et al. 2015. The human transcriptome across tissues and individuals. *Science* 348(6235):660–65
76. Merkin J, Russell C, Chen P, Burge CB. 2012. Evolutionary dynamics of gene and isoform regulation in mammalian tissues. *Science* 338(6114):1593–99
77. Miyawaki S, Kuroki S, Maeda R, Okashita N, Koopman P, Tachibana M. 2020. The mouse *Sry* locus harbors a cryptic exon that is essential for male sex determination. *Science* 370(6512):121–24
78. Modrek B, Lee CJ. 2003. Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss. *Nat. Genet.* 34(2):177–80
79. Moore MJ, Wang Q, Kennedy CJ, Silver PA. 2010. An alternative splicing network links cell-cycle control to apoptosis. *Cell* 142(4):625–36
80. Nadimpalli S, Persikov AV, Singh M. 2015. Pervasive variation of transcription factor orthologs contributes to regulatory network evolution. *PLOS Genet*. 11(3):e1005011
81. Nakano Y, Wiechert S, Bánfi B. 2019. Overlapping activities of two neuronal splicing factors switch the GABA effect from excitatory to inhibitory by regulating REST. *Cell Rep*. 27(3):860–71.e8
82. Nitta KR, Jolma A, Yin Y, Morgunova E, Kivioja T, et al. 2015. Conservation of transcription factor binding specificities across 600 million years of bilateria evolution. *eLife* 4:e04837
83. Nurtdinov RN, Artamonova II, Mironov AA, Gelfand MS. 2003. Low conservation of alternative splicing patterns in the human and mouse genomes. *Hum. Mol. Genet.* 12(11):1313–20
84. Odom DT, Dowell RD, Jacobsen ES, Gordon W, Danford TW, et al. 2007. Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat. Genet.* 39(6):730–32
85. Pang ZP, Yang N, Vierbuchen T, Ostermeier A, Fuentes DR, et al. 2011. Induction of human neuronal cells by defined transcription factors. *Nature* 476(7359):220–23
86. Paps J, Holland PWH. 2018. Reconstruction of the ancestral metazoan genome reveals an increase in genomic novelty. *Nat. Commun.* 9(1):1730
87. Park JW, Yang J, Xu R-H. 2018. PAX6 alternative splicing and corneal development. *Stem Cells Dev*. 27(6):367–77
88. Parras A, Anta H, Santos-Galindo M, Swarup V, Elorza A, et al. 2018. Autism-like phenotype and risk gene mRNA deadenylation by CPEB4 mis-splicing. *Nature* 560(7719):441–46
89. Quesnel-Vallières M, Dargaei Z, Irimia M, Gonatopoulos-Pournatzis T, Ip JY, et al. 2016. Misregulation of an activity-dependent splicing network as a common mechanism underlying autism spectrum disorders. *Mol. Cell* 64(6):1023–34
90. Quesnel-Vallières M, Irimia M, Cordes SP, Blencowe BJ. 2015. Essential roles for the splicing regulator nSR100/SRRM4 during nervous system development. *Genes Dev*. 29(7):746–59
91. Raposo AASF, Vasconcelos FF, Drechsel D, Marie C, Johnston C, et al. 2015. Ascl1 coordinately regulates gene expression and the chromatin landscape during neurogenesis. *Cell Rep*. 10(9):1544–56
92. Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS, et al. 2013. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499(7457):172–77
93. Reyes A, Anders S, Weatheritt RJ, Gibson TJ, Steinmetz LM, Huber W. 2013. Drift and conservation of differential exon usage across tissues in primate species. *PNAS* 110(38):15377–82

94. Rogozin IB, Babenko VN, Fedorova ND, Jackson JD, Jacobs AR, et al. 2003. Evolution of eukaryotic gene repertoire and gene structure: discovering the unexpected dynamics of genome evolution. *Cold Spring Harb. Symp. Quant. Biol.* 68:293–301
95. Roller M, Stamper E, Villar D, Izuogu O, Martin F, et al. 2021. LINE retrotransposons characterize mammalian tissue-specific and evolutionarily dynamic regulatory regions. *Genome Biol.* 22(1):62
96. Romero IG, Ruvinsky I, Gilad Y. 2012. Comparative studies of gene expression and the evolution of gene regulation. *Nat. Rev. Genet.* 13(7):505–16
97. Roy SW, Gilbert W. 2005. Complex early genes. *PNAS* 102(6):1986–91
98. Royo JL, Maeso I, Irimia M, Gao F, Peter IS, et al. 2011. Transphyletic conservation of developmental regulatory state in animal evolution. *PNAS* 108(34):14186–91
99. Saito Y, Yuan Y, Zucker-Scharff I, Fak JJ, Jereb S, et al. 2019. Differential NOVA2-mediated splicing in excitatory and inhibitory neurons regulates cortical development and cerebellar function. *Neuron* 101(4):707–20.e5
100. Sander M, Neubüser A, Kalamaras J, Ee HC, Martin GR, German MS. 1997. Genetic analysis reveals that PAX6 is required for normal transcription of pancreatic hormone genes and islet development. *Genes Dev.* 11(13):1662–73
101. Schmidt D, Schwalie PC, Wilson MD, Ballester B, Gonçalves A, et al. 2012. Waves of retrotransposon expansion remodel genome organization and CTCF binding in multiple mammalian lineages. *Cell* 148(1–2):335–48
102. Shirai LT, Saenko SV, Keller RA, Jerónimo MA, Brakefield PM, et al. 2012. Evolutionary history of the recruitment of conserved developmental genes in association to the formation and diversification of a novel trait. *BMC Evol. Biol.* 12:21
103. Simpson TI, Price DJ. 2002. Pax6; a pleiotropic player in development. *Bioessays* 24(11):1041–51
104. Singh P, Ahi EP. 2022. The importance of alternative splicing in adaptive evolution. *Mol. Ecol.* 31(7):1928–38
105. Sorek R, Ast G, Graur D. 2002. Alu-containing exons are alternatively spliced. *Genome Res.* 12(7):1060–67
106. Soutourina J. 2018. Transcription regulation by the Mediator complex. *Nat. Rev. Mol. Cell Biol.* 19(4):262–74
107. Spitz F, Furlong EEM. 2012. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* 13(9):613–26
108. St-Onge L, Sosa-Pineda B, Chowdhury K, Mansouri A, Gruss P. 1997. *Pax6* is required for differentiation of glucagon-producing α-cells in mouse pancreas. *Nature* 387(6631):406–9
109. Sucena E, Delon I, Jones I, Payre F, Stern DL. 2003. Regulatory evolution of *shavenbaby/ovo* underlies multiple cases of morphological parallelism. *Nature* 424(6951):935–38
110. Sucena E, Stern DL. 2000. Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by cis-regulatory evolution of *ovo/shaven-baby*. *PNAS* 97(9):4530–34
111. Sudmant PH, Alexis MS, Burge CB. 2015. Meta-analysis of RNA-seq expression data across species, tissues and studies. *Genome Biol.* 16:287
112. Swisa A, Avrahami D, Eden N, Zhang J, Feleke E, et al. 2017. PAX6 maintains β cell identity by repressing genes of alternative islet cell types. *J. Clin. Investig.* 127(1):230–43
113. Taliaferro JM, Alvarez N, Green RE, Blanchette M, Rio DC. 2011. Evolution of a tissue-specific splicing network. *Genes Dev.* 25(6):608–20
114. Tan Q, Yalamanchili HK, Park J, De Maio A, Lu H-C, et al. 2016. Extensive cryptic splicing upon loss of RBM17 and TDP43 in neurodegeneration models. *Hum. Mol. Genet.* 25(23):5083–93
115. Tapial J, Ha KCH, Sterne-Weiler T, Gohr A, Braunschweig U, et al. 2017. An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. *Genome Res.* 27(10):1759–68
116. Tennant BR, Robertson AG, Kramer M, Li L, Zhang X, et al. 2013. Identification and analysis of murine pancreatic islet enhancers. *Diabetologia* 56(3):542–52
117. Thakurela S, Tiwari N, Schick S, Garding A, Ivanek R, et al. 2016. Mapping gene regulatory circuitry of Pax6 during neurogenesis. *Cell Discov.* 2:15045

118. Thanaraj TA, Clark F, Muilu J. 2003. Conservation of human alternative splice events in mouse. *Nucleic Acids Res*. 31(10):2544–52
119. Torres-Méndez A, Bonnal S, Marquez Y, Roth J, Iglesias M, et al. 2019. A novel protein domain in an ancestral splicing factor drove the evolution of neural microexons. *Nat. Ecol. Evol.* 3(4):691–701
120. Torres-Méndez A, Pop S, Bonnal S, Almudi I, Avola A, et al. 2022. Parallel evolution of a splicing program controlling neuronal excitability in flies and mammals. *Sci. Adv.* 8(4):eabk0445
121. Tress ML, Abascal F, Valencia A. 2017. Alternative splicing may not be the key to proteome complexity. *Trends Biochem. Sci.* 42(2):98–110
122. Ule J, Blencowe BJ. 2019. Alternative splicing regulatory networks: functions, mechanisms, and evolution. *Mol. Cell* 76(2):329–45
123. Ule J, Stefani G, Mele A, Ruggiu M, Wang X, et al. 2006. An RNA map predicting Nova-dependent splicing regulation. *Nature* 444(7119):580–86
124. Venter JC, Adams MD, Myers EW, Li PW. 2001. The sequence of the human genome. *Science* 291(5507):1304–51
125. Verta J-P, Jacobs A. 2021. The role of alternative splicing in adaptation and evolution. *Trends Ecol. Evol.* 37(4):299–308
126. Vierbuchen T, Ostermeier A, Pang ZP, Kokubu Y, Südhof TC, Wernig M. 2010. Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* 463(7284):1035–41
127. Villar D, Berthelot C, Aldridge S, Rayner TF, Lukk M, et al. 2015. Enhancer evolution across 20 mammalian species. *Cell* 160(3):554–66
128. Villate O, Turatsinze J-V, Mascali LG, Grieco FA, Nogueira TC, et al. 2014. Nova1 is a master regulator of alternative splicing in pancreatic beta cells. *Nucleic Acids Res*. 42(18):11818–30
129. Waks Z, Klein AM, Silver PA. 2011. Cell-to-cell variability of alternative RNA splicing. *Mol. Syst. Biol.* 7:506
130. Warzecha CC, Shen S, Xing Y, Carstens RP. 2009. The epithelial splicing factors ESRP1 and ESRP2 positively and negatively regulate diverse types of alternative splicing events. *RNA Biol*. 6(5):546–62
131. Witten JT, Ule J. 2011. Understanding splicing regulation through RNA splicing maps. *Trends Genet*. 27(3):89–97
132. Yang YY, Yin GL, Darnell RB. 1998. The neuronal RNA-binding protein Nova-2 is implicated as the autoantigen targeted in POMA patients with dementia. *PNAS* 95(22):13254–59
133. Zattara EE, Busey HA, Linz DM, Tomoyasu Y, Moczek AP. 2016. Neofunctionalization of embryonic head patterning genes facilitates the positioning of novel traits on the dorsal head of adult beetles. *Proc. R. Soc. B* 283:20160824
134. Zhang C, Frias MA, Mele A, Ruggiu M, Eom T, et al. 2010. Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. *Science* 329(5990):439–43
135. Zhang C, Zhang Z, Castle J, Sun S, Johnson J, et al. 2008. Defining the regulatory network of the tissue-specific splicing factors Fox-1 and Fox-2. *Genes Dev*. 22(18):2550–63