



HHS Public Access

Author manuscript

Wiley Interdiscip Rev RNA. Author manuscript; available in PMC 2024 September 01.

Published in final edited form as:

Wiley Interdiscip Rev RNA. 2023 ; 14(5): e1778. doi:10.1002/wrna.1778.

The DNA binding high mobility group box (HMGB) protein family functionally binds RNA

Desmond J. Hamilton,

University of Colorado Boulder

Abigail E. Hein,

University of Colorado Boulder

Deborah S. Wuttke*,

University of Colorado Boulder

Robert T. Batey*

University of Colorado Boulder

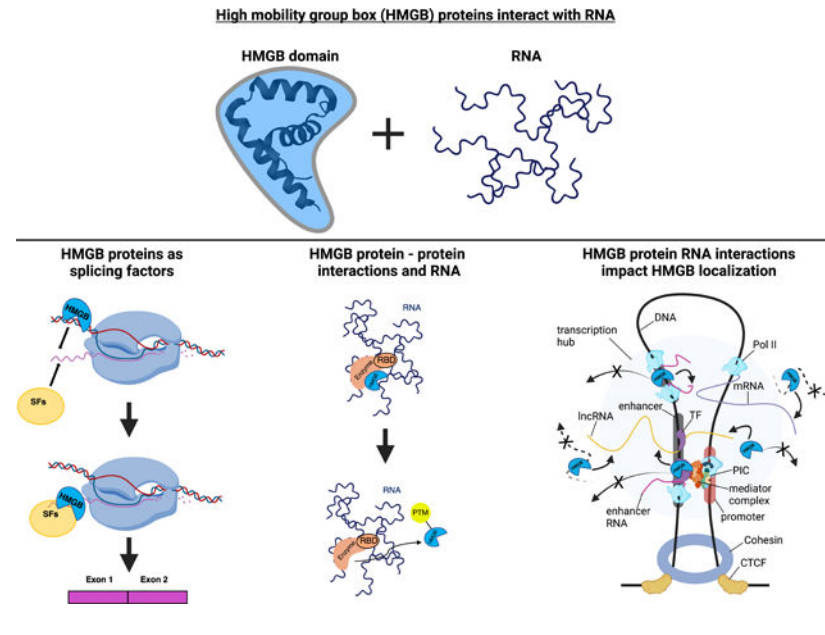
Abstract

Nucleic-acid binding proteins regulate transcription, splicing, RNA stability, RNA localization and translation, together tailoring gene expression in response to stimuli. Upon discovery, these proteins are typically classified as either DNA or RNA binding as defined by their *in vivo* functions; however, recent evidence suggests dual DNA and RNA binding by many of these proteins. High mobility group box (HMGB) proteins have a DNA binding HMGB domain, act as transcription factors and chromatin remodeling proteins, and are increasingly understood to interact with RNA as means to regulate gene expression. Herein, multiple layers of evidence that the HMGB family are dual DNA and RNA binding proteins is comprehensively reviewed. For example, HMGB proteins directly interact with RNA *in vitro* and *in vivo*, are localized to RNP granules involved in RNA processing, and their protein interactors are enriched in RNA binding proteins involved in RNA metabolism. Importantly, in cell-based systems, HMGB protein-RNA interactions facilitate protein-protein interactions, impact splicing outcomes, and modify HMGB protein genomic or cellular localization. Mis-regulation of these HMGB-RNA interactions are also likely involved in human disease. This review brings to light that as a family, HMGB proteins are likely to bind RNA which is essential to HMGB protein biology.

Graphical Abstract

*Authors to whom correspondence should be addressed. Deborah.wuttke@colorado.edu, Robert.batey@colorado.edu.

Conflict of Interest: No conflict of interest to report.



1. INTRODUCTION

The interdependence of how an active and pervasive transcriptome regulates the activity of DNA-binding transcription factors (TF) and chromatin remodeling proteins (ChRP) or how these TFs and ChRPs control multiple levels of gene expression by interacting with RNA is a deep, unsolved problem (D et al., 2016; Guo & Guttman, 2022; Hudson & Ortlund, 2014; Long, Wang, et al., 2017). TF- and ChRP-RNA interactions appear to operate through many distinct mechanisms that impact a large number of gene expression pathways which could allow them to regulate diverse cellular functions within numerous cell types by binding RNA (Akhade et al., 2017; Brown, 2005; D et al., 2016; Hentze et al., 2018; Hudson & Ortlund, 2014; Langst & Manelyte, 2015; Long, Bolanos, et al., 2017; See et al., 2019; Sigova et al., 2015; Skalska et al., 2017; Voong et al., 2021). The depth of known TF/ChRP-RNA interactions reveal an immense gap in TF and ChRP knowledge. The initial challenge of developing a comprehensive understanding of the relationship between the transcriptome and TFs/ChRPs is the identification and categorization of TFs and ChRPs that interact with RNA. One family of TFs and ChRPs that bind DNA through their high mobility group box (HMGB) DNA-binding domain (HMGB-DBD) have yet to be categorized as an RNA binding family; however, extensive family-wide RNA binding evidence suggests they bind RNA. An analysis of this evidence is the subject of this review.

HMGBs are ubiquitous, essential proteins present in all eukaryotic species with at least 48 human proteins possessing an HMGB-DBD (Malarkey & Churchill, 2012; Stros, 2010; Stros et al., 2007). HMGB-DBDs are implicated in diverse cellular processes related to their DNA-binding capabilities, including, but not limited to, stem cell biology, DNA repair, nucleosome and chromatin remodeling, and transcriptional regulation with mis-regulation of HMGB function resulting in human disease (Bianchi & Agresti, 2005; Malarkey & Churchill, 2012; McCauley et al., 2019; Reeves, 2015; Stros, 2010; Stros et al., 2007; Voong et al., 2021). HMGB-DBDs are composed of ~65–85 residues that form three

alpha-helices which fold into an ~80° L-shaped wedge, producing a concave surface that acts as the primary DNA binding surface (Figure 1) (Hou et al., 2017; Malarkey & Churchill, 2012; Remenyi et al., 2003; Weiss, 2001). Via extensive interactions with the minor groove, HMGB-DBDs bend DNA which may recruit other proteins or unravel nucleosomal DNA (Catez et al., 2004; Papantonis, 2021). HMGBs are divided into two subfamilies based predominantly on their DNA binding preferences: sequence specific (SS) and non-sequence specific (NSS) DNA or structure specific DNA interactors (Bowles et al., 2000; Phochanukul & Russell, 2010; Schepers et al., 2002; Malarkey & Churchill, 2012; Ner, 1992). SS and NSS HMGBs generally have distinct biological functions: SS are TFs and NSS generally act as ChRPs (Malarkey & Churchill, 2012).

Here, we present a synthesis of current knowledge from a diverse set of literature suggesting HMGB-DBD containing proteins are an RNA-binding family. Supporting evidence that HMGBs productively interact with RNA is reviewed in the following categories: HMGB-DBDs have RNA binding protein (RBP) properties, HMGBs directly interact with RNA through their HMGB-DBD *in vitro* and *in vivo*, HMGB-DBD conservation suggests RNA binding potential, HMGBs localize to sites of RNP granules, HMGBs' interacting protein partners are enriched with RBPs, and HMGB-RNA interactions *in vivo* impact HMGB driven cellular processes. To stimulate future discovery of HMGB-RNA interaction biology, we discuss HMGB family wide trends, potential mechanisms for how HMGB-RNA interactions impact HMGB processes, and the presence of HMGB-RNA interactions in human disease. We also highlight important unanswered questions that are essential to understand HMGB-RNA biology.

2. HMGB DOMAIN PROPERTIES FACILITATE RNA INTERACTIONS

The physical characteristics of HMGB-DBDs confer a broad ability to interact with diverse DNA structures and sequences by inducing DNA bends and/or by stabilizing non-B form DNA ligands, including but not limited to, cruciform DNA, bent or kinked DNA, supercoiled DNA, and nucleosomal DNA (Agresti & Bianchi, 2003; Bianchi & Beltrame, 1998; JR et al., 1998; Malarkey & Churchill, 2012; Murugesapillai et al., 2017; Ner, 1992; Reeves, 2015; Stros, 2010; Stros et al., 2007). The DNA-binding properties of HMGBs have been thoroughly reviewed previously and will not be a major component of this review (Hou et al., 2017; Malarkey & Churchill, 2012; Reeves, 2015; Stros, 2010), significantly though, several physical properties of HMGB-DBDs that allow binding to diverse non-B form DNA structures may facilitate binding to RNA. First, HMGB-DBDs are partially disordered in the unbound state and transition to a more ordered state when bound to DNA (Masse et al., 2002; Weiss, 2001). This is important because intrinsically disordered regions are commonly found in RBPs which is thought to increase binding affinity of RBPs to RNA when undergoing an unordered to ordered state transition upon binding RNA (Balcerak et al., 2019; Corley et al., 2020; Cruz-Gallardo et al., 2019; Jarvelin et al., 2016; Leulliot & Varani, 2001). Further, HMGBs contain between 1–6 HMGB-DBDs with many in tandem. Tandem RNA binding domains are a common feature of RBPs and are thought to impart a malleable and adaptive RNA binding interface which facilitates protein-RNA interactions (Hollmann et al., 2020; Lunde et al., 2007). Finally, HMGBs of both subfamilies interact

with non-B form DNA structures with high affinity, thus HMGBs might interact with non-B form RNA structures as well.

3. EXPERIMENTAL EVIDENCE THAT HMGBS ARE AN RNA BINDING FAMILY

3.1 HMGBs bind and associate with RNA.

HMGBs are pervasively and compellingly shown to be RBPs. When many studies are evaluated together, there is clear and unambiguous evidence that RNA binding is a general feature of HMGBs. One caveat however is that the many experimental approaches used to discover RBPs and/or to investigate HMGB-RNA biology do not always differentiate between direct or indirect HMGB-RNA interactions, yet this distinction is important to understand mechanistic details more completely. Thus, in this review we clearly distinguish between binding and association. We define binding as an HMGB-RNA zero distance direct interaction obtained from either ultra-violet crosslinking *in vivo* and/or *in vitro* biochemical binding assays (Budowsky et al., 1986; Lee & Ule, 2018; Li et al., 2014; Pashev et al., 1991; Wheeler et al., 2018). In contrast, association is defined as an observed RNA interaction, but the experimental approach used is not able to differentiate between a direct or indirect interaction, with indirect interactions likely occurring through protein-protein-RNA and/or DNA-RNA interactions. HMGB-RNA association designations are assigned from methods that use formaldehyde crosslinking or non-UV pull down experiments such as RNA immunoprecipitation from cell lysates (Ramanathan et al., 2019; Wheeler et al., 2018). It is key to note that interactions identified in these ways could be direct, just that there remains the possibility that they are indirect in the absence of further characterization.

Numerous HMGBs bind and/or associate with RNA *in vivo* and/or *in vitro* (Table 1). Experimental evidence supports that a significant portion of HMGBs directly bind (30/48, Table 1) and/or associate (19/48, Table 1) with RNA across multiple conditions, cell types, and species. As a portion of this information is obtained from broad or indiscriminate RBP discovery studies, which are more likely to detect abundant proteins or proteins that are commonly expressed across many cell types, the real number of species-wide HMGBs that bind RNA is likely much higher. For instance, HMGB1 is expressed across many cell types and tissues whereas Sox protein family expression is more cell-type specific (Muller et al., 2004; Uhlen et al., 2015). Additionally, the more well studied HMGBs are more likely to have been investigated as an RBP whereas RNA binding potential may not yet have been explored for the lesser studied HMGBs. The likelihood for other HMGB-RNA interactions is reinforced by prediction software which suggests the majority of HMGBs are also RBPs (CatRAPID, ~80% of full length human HMGBs) (Livi et al., 2016). Both SS TF and NSS ChRP HMGB subfamilies bind RNA, further reinforcing that general features of the HMGB-DBD facilitate binding to RNA. Most of the HMGBs identified to associate with RNA also directly bind RNA (Table 1) suggesting that techniques that identify HMGB-RNA associations, likely also identify RNA binding. Thus, at least 75% of these HMGBs possess experimental evidence for direct interactions with RNA.

3.2 Direct HMGB-RNA interactions

In studies that analyze direct RNA binding, it appears that HMGBs bind to RNAs that are diverse in both sequence and structure, as also seen for their DNA-binding preferences (Table 2). Thus far, biochemical characterization of HMGB-RNA interactions are restricted to a subset of HMGBs. However, clear trends can be discerned from the available studies.

In vitro, Sox2 interacts with a variety of RNA ligands with this interaction likely mediated by the HMGB-DBD (Table 2) (Hamilton et al., 2022; Holmes et al., 2020; Hou et al., 2020; Tung et al., 2010; Cajigas et al., 2021). Sox2 HMGB-DBD interacts with G-quadruplex RNA (1 nM), 4-way junction RNA (5 nM), fragment of lncRNA ES2 (3 nM), and varied stem loop RNAs (~10–100 nM) (Hamilton et al., 2022; Holmes et al., 2020). Sox2-HMGB-DBD's affinity for RNA is comparable to dsDNA containing the ChIP validated Sox2 consensus motif ($K_D = \sim 1$ nM), but much tighter than to dsDNA without the Sox consensus motif ($K_D = \sim 600$ nM) (Hamilton et al., 2022; Holmes et al., 2020).

Several features of Sox2's interactions with RNA have been established from various studies. First, a core RNA sequence motif was identified for Sox2 both *in vitro* (GCCCCX, X= A or U) and *in vivo* (YCCCCZG, Y= U/G, Z= A/G) implying sequence preferences (Hou et al., 2020). Second, an identified nearby side motif (UCGCGWU, V= C/U/G) may promote the formation of Sox2 preferred secondary structures. Third, it was observed that Sox2 HMGB-DBD binding to DNA and RNA is mutually exclusive and uses similar electrostatic contributions (Holmes et al., 2020). A comprehensive alanine scan revealed that overlapping but distinct residues are important for Sox2-HMGB-DBD interactions with DNA or RNA (Holmes et al., 2020). Thus, DNA and RNA binding are mediated, at least in part, by the same face of the HMGB-DBD. Finally, Sox2 was observed to bind multiple sites within EVF2 RNAs with similar affinities (Cajigas et al., 2021). When ~60–100 residue regions C-terminus of the HMGB-DBD was deleted, Sox2's preference for RNA ligands was altered or removed, suggesting that a region adjacent to the canonical HMGB-DBD could be important for imbuing Sox2 with RNA specificity through direct RNA interactions or through protein-protein interactions (Hou et al., 2020; Jing et al., 2020).

Other Sox family proteins bind RNA through their HMGB-DBD with Sox family proteins being split into nine groups based on HMGB-DBD sequence similarities (Table 2). Significantly, except for Sox30, one Sox HMGB-DBD from each group binds with high affinity to a variety of RNA ligands *in vitro* (affinities range from ~1 to 360 nM) (Hamilton et al., 2022). These affinities are comparable to dsDNA that contains the Sox family consensus sequence and are much tighter than to dsDNA without the consensus sequence. Further, the Sox family shows selectivity for distinct RNA features from G-quadruplex RNA, stem-loop RNA with an internal bulge, a fragment of ES2 lncRNA, and four-way junction RNA. The Sox proteins predominately bound the G-quadruplex RNA the tightest and the stem-loop with internal bulge RNA the weakest.

Some of the TCF/LEF family transcription factors bind RNA *in vitro* with high affinity through their HMGB-DBDs (Table 2). The TCF7 HMGB-DBD binds RNA *in vitro* with a preference for RNAs with stem-loops containing internal bulges with the interaction occurring predominately with internal bulged regions (Lee & Jeong, 2004; Park et al., 2005).

As observed with Sox2, the TCF7-HMGB-DBD interaction with RNA is competitive with DNA. An RNase foot printing analysis revealed protected and deprotected regions in the presence of TCF7, (Park et al., 2005) suggesting TCF7 may change the structure of RNA. LEF1 HMGB-DBD binds to a variety of RNA ligands with affinities (4–630 nM) that are ~2–45 fold weaker than the consensus dsDNA but much tighter than to dsDNA without consensus motif (Hamilton et al., 2022). Since the TCF/LEF family HMGB-DBDs have nearly 100% sequence conservation, the other members of the family are highly likely to bind RNA.

One interesting HMGB is TFAM which has tandem HMGB-DBDs and is thought to bind both SS and NSS DNA. *In vitro*, TFAM binds hybrid RNA:DNA 4-way junctions ($K_D = 300$ nM), RNA 4-way junctions ($K_D = 70$ –300 nM), mitochondrial tRNAs ($K_D = 20$ nM), hybrid DNA/RNA G-quadruplexes ($K_D = 1.9$ nM), and RNA G-quadruplexes ($K_D = 1.3$ nM) (Table 2) (Brown et al., 2015; Lyonnais et al., 2017). TFAM does not interact tightly with ssRNA (poly A/C or poly U), dsRNA, dsRNA hairpin loops, or linear RNA:DNA hybrids ligands; however, TFAM interacts weakly with dsRNA with an internal loop with a K_D of 2 μ M (Brown et al., 2015). While the natural RNA substrate of TFAM remains unknown, since it binds an assortment of RNA ligands with a range of affinities it is likely to preferentially bind a subset of cellular RNA targets.

Drosophila HMGD is another HMGB that has strong evidence for direct RNA binding (Table 2). This protein binds to the transactivation response region (TAR) RNA of HIV-1 with a 29 nM K_D , the dsRNA stem loop Rbe RNA of HIV-1, and did not bind to ssRNA (Arimondo et al., 2000). HMGD binds and maintains high affinity to a variety of TAR RNA mutants including: deletion of a 3-base bulge ($K_D = 36$ nM), mutation of terminal loop sequence from wild type CUGGGA to UUCG ($K_D = 30$ nM), deletion of the stem loop and a reduction in size by ~50% while maintaining 3 base bulge ($K_D = 53$ nM), and a fully paired dsRNA mutant lacking both the 3 base bulge and stem loop ($K_D = 70$ nM). RNase A foot-printing showed regions of RNA near the stem-loop/bulge were protected in the presence of HMGD with increased cleavage at a site opposing one of the protected regions signifying that HMGD may impart structural changes. Together, this study shows a strong ability to interact with high affinity to various dsRNA stem loop elements.

In another example of broad specificity RNA binding, HMGB1 interacts with RNA G-quadruplexes (Serikawa et al., 2018), a specific stem loop of hepatitis C virus RNA (Yu et al., 2015), and 5S *E. coli* rRNA which contains stem loops, internal bulges, and dsRNA regions (Bell et al., 2008) (Table 2). The 5S *E. coli* rRNA was able to compete 4-way junction DNA from HMGB1; however, HMGB1 did not bind tRNA. The HMGB1-A domain and dual HMGB-DBDs bind to siRNAs (Choi et al., 2020; Lee et al., 2012; Oh & Lee, 2014). HMGB1 binds RNA *in vivo* with an identified motif of 5'- MWGRA-3' (M=A/C, W=A/U, R=A/G) (Sofiadis et al., 2021). From the same subfamily as HMGB1, HMGB3 binds to specific sites within HIV-1 Tat mRNA, including stem loops and internal bulges (Khoury et al., 2021).

Many non-human HMGBs also bind RNA *in vitro* (Table 2). Mouse maelstrom (MAEL) is an HMGB important for piRNA biogenesis and processing (Genzor & Bortvin, 2015; Sato

& Siomi, 2015). MAEL bound four-way junction RNA with a K_D of 640 nM and an alanine scan demonstrated that residues involved in binding to DNA and RNA 4-way junctions did not completely overlap (Genzor & Bortvin, 2015). MAEL did not bind ssRNA and weakly bound dsRNA and stem-loop RNA. Interestingly, in another species, MAEL from *Bombyx* (silkworm) bound ssRNA (Chen et al., 2015). Another *Drosophila* HMGB, DssRP, is homologous by sequence comparison to human SSRP1, an HMGB that binds RNA. DssRP binds homopolymers of poly(U) and poly(G) ssRNA (Hsu et al., 1993). A plant HMGB, CsHMG, interacts with poly(U) and poly(G) RNA with a preference for uridine through the HMGB domain (de Souza et al., 2012). Further, it was shown that CsHMG did not bind to AU-rich probes and that CsHMG bound ssRNA ligands more tightly than dsDNA ligands.

While the comprehensive biochemical characterization of HMGB-DBD-RNA interactions is limited, there are trends that emerge. HMGB-DBDs from human, *Drosophila*, mouse, *Bombyx*, and plants bind to a diverse set of RNA structures and sequences, suggesting RNA interactions are a common feature of all HMGBs across multiple species (Table 2). HMGBs have been observed to interact with high affinity to a variety of dsRNA stem-loop elements which contain non-base paired regions; however, this activity is not observed uniformly for all HMGBs. Sox2 and MAEL both use overlapping and yet distinct residues when binding DNA or RNA, suggesting a separation of function between DNA and RNA binding may exist. Importantly, the binding affinity for some HMGBs to their RNA ligands is comparable to the HMGBs DNA ligands, suggesting RNA and DNA may compete for interactions with HMGBs *in vivo*. Finally, multiple HMGBs appear to be capable of modifying the RNA structure. Taken together, these studies indicate that DNA-binding HMGB-DBDs are extensively involved in RNA recognition as well and that this is a widespread property of this family of proteins.

3.3 HMGB domain conservation suggests RNA binding

While all HMGB-DBDs appear to possess the same L-shaped structure, their amino acid sequence varies between the different HMGB subfamilies, as shown by a sequence alignment of human HMGB-DBDs (Figure 2). HMGB-DBDs from the same subfamily are generally highly conserved, such as the TCF/LEF family in which the HMGB-DBD is nearly 100% identical or the Sox family which are at least ~50% conserved. This is important because at least one HMGB from each of these sub-families binds to RNA (asterisk, Figure 2). The patterns of amino acid conservation across families along with evidence of RNA binding in each family suggests that the other HMGBs from the same family may bind RNA as well.

Residues within the HMGB-DBD important for RNA binding or RNA associations were identified using amino acid mutational analysis of Sox2, Sox9, and mouse MAEL HMGB-DBDs (Figure 2) (Genzor & Bortvin, 2015; Girardot et al., 2018; Holmes et al., 2020). A few of the positively charged residues important for RNA binding in the Sox2 HMGB-DBD are highly conserved throughout a majority of HMGBs (purple, Figure 2). In contrast, residues from Sox9 (blue, Figure 2) and MAEL (magenta, Figure 2) generally show less

conservation throughout the HMGB family. Conserved residues that are important for RNA interactions hint that other HMGBs may also utilize these residues to bind RNA.

3.4 HMGBs localize to RNA-protein granules

Protein localization is another layer of evidence that assists in identifying a protein's biological function (Itzhak et al., 2016; Scott et al., 2005). In that light, HMGBs found within RNA-protein (RNP) granules may be localized to these regions through interactions with RNA. A substantial portion (20/48) of human HMGBs have been reported to localize to RNP granules broadly involved in RNA metabolism processes (Figure 3). For example, UBTF, Sox15, HMGB2, HMGB3, TOX, and TOX4 have been observed within nucleoli which are centers of rRNA assembly and processing and with a separate function of assembling RNP complexes distinct from rRNA (Iarovaia et al., 2019; Jellbauer & Jansen, 2008; Lafontaine et al., 2021; Thul et al., 2017). SRY, Sox6, Sox7, TOX4, Smarce1, and HBP1 are localized within nuclear speckles, which are important centers for multiple steps of RNA metabolism such as splicing (Chen & Belmont, 2019; Faber et al., 2022; Galganski et al., 2017; Ilik & Aktas, 2021; Ohe et al., 2002; Saitoh et al., 2004; Sato et al., 2011; Spector & Lamond, 2011; Thul et al., 2017). Furthermore, UBTF, SSRP1, MAEL, HMGB3, TFAM, and HMGB1 localize to stress granules, which are cytoplasmic RNP granules that form in response to cellular stress and are thought to protect RNAs and/or RNA processing machinery from damage (Campos-Melo et al., 2021; Corbet & Parker, 2019; Jain et al., 2016; Markmiller et al., 2018; Youn et al., 2019; Y. Zhang et al., 2021). Other HMGBs localize to nuclear bodies including paraspeckles which may regulate gene expression by retaining dsRNAs with adenosine to inosine edits (Chaoui et al., 2015; Hata et al., 2008; Penrad-Mobayed et al., 2018; Pisani & Baron, 2019; Thul et al., 2017; Wang & Chen, 2020). However, with the exception of UBTF which is known to regulate rRNA gene expression (Ueshima et al., 2017), the role(s) HMGBs play in these RNP granules is largely unknown, but given the current interest in RNP granules, it is likely that the specific roles HMGBs play in RNA metabolism will be uncovered.

3.5 HMGB protein interactors are enriched with RBPs

Protein-protein interaction (PPI) networks are another useful means to infer the function of proteins. Examination of the PPI database, agile protein interactome data server (APID) (Alonso-Lopez et al., 2019; Alonso-Lopez et al., 2016), reveals that HMGB protein interactors are enriched with RBPs involved in RNA metabolism (Figure 4). To assess whether an HMGB protein partner is involved in a broad range of RNA-centric roles, a Gene Ontology (GO) enrichment analysis of each human HMGB's interacting proteins was applied (Ashburner et al., 2000; Gene Ontology, 2021; Mi et al., 2013; Mi et al., 2019). HMGBs are then labeled as "found" (green) or "unknown" (grey) (Figure 4A) dependent upon whether they interact with at least one protein with the following GO annotation: labeled as RNA binding (GO: molecular function); RNA metabolism process, i.e., splicing (GO: biological process); within RNP assemblies (GO: cellular component). Thus, the APID database reveals that ~54% of human HMGBs interact with proteins involved in RNA metabolism and ~80% interact with RBPs and/or with proteins localized within RNP assemblies (Figure 4). Significantly, many of the HMGBs interact with more RBPs than DNA binding proteins or with approximately similar amounts. While HMGBs interact

broadly with proteins involved in RNA metabolism, certain biological process assignments were enriched, including splicing, RNA transport, stability, and processing. Further, the cellular component aspect shows HMGB protein interactors are highly localized within RNP assemblies. Some of these HMGBs have only a few APID interactors, thus it is possible they interact with proteins involved in RNA metabolism, it's just not in the APID database or has not yet been experimentally shown yet. This PPI network evaluation provides strong circumstantial evidence that HMGBs are involved in RNA metabolism pathways, perhaps by binding RNA.

4. HMGBS INTERACTIONS WITH RNA IMPACT HMGB CELL BIOLOGY

HMGBs possess multiple properties of RBPs and bind RNA, thus, a major outstanding question is whether and how these RNA binding capabilities drive *in vivo* biological activities. While this is currently an active area of research, multiple lines of evidence support a variety of important roles for RNA binding by HMGBs *in vivo*. These include the regulation of transcription by modulating HMGB genomic and cellular localization, affecting splicing decisions of HMGB bound RNAs, and assisting in the formation of HMGB-protein-RNA assemblies to facilitate post-translational modifications and/or to establish topological chromatin structure through chromosomal interactions. When these various studies are viewed through the lens of non-canonical RNA interactions by members of the HMGB family, a multitude of possible mechanisms of action likely exist. Outlined in the next few subsections is an examination of the experimental evidence for how HMGB-RNA interactions impact HMGB cell biology followed by a discussion of potential mechanisms.

4.1 HMGB-RNA interactions influence HMGB-protein interactions

Protein-RNA interactions are increasingly understood to be important for PPIs, including RNA scaffolds which can play roles in regulation of chromosome topology and/or recruiting proteins to facilitate PPIs to promote enzymatic reactions (Bouwman et al., 2022; Castello et al., 2015; Chen et al., 2014; Fernandes & Buchan, 2021; Razin & Gavrilyov, 2021). Since protein interactors of HMGBs are highly enriched in RBPs, HMGBs bind DNA with protein partners, and their post translation modification (PTM) landscapes are important for their many functions (Andersson et al., 2014; Bernard & Harley, 2010; Goos et al., 2022; Richard et al., 2017; Williams et al., 2020), RNAs that regulate HMGB-protein interactions are likely important for HMGB family biology.

4.1.1 Experimental evidence for HMGB-RNA interactions regulating HMGB-protein interactions—

Sox2 binds the enhancer lncRNA EVF2, which likely facilitates PPIs and intrachromosomal interactions to regulate gene expression in mouse cell lines (Cajigas et al., 2021). Sox2, EVF2 and an ultra-conserved element enhancer, through intrachromosomal interactions, occupy overlapping genomic regions. Knockdown of wild type EVF2 results in loss of Sox2 genomic occupancy which was in part rescued with a truncated 5'-EVF2 RNA lacking the 3' end fragment. Knockdown of Sox2 or EVF2 changes the ultra-conserved element enhancer chromosome interaction landscape and impacts expression of the overlapped genes. Together, EVF2 likely acts in *cis* as an RNP

scaffold which allows Sox2 and other proteins to facilitate intrachromosomal interactions to repress or activate genes.

Sox2 associates with the super enhancer lncRNA NEAT1 which increases Sox2 serine phosphorylation resulting in impaired maintenance of pluripotency in bone marrow mesenchymal stem cells (H. Zhang et al., 2021). CDK2 interacts with and phosphorylates serine residues of Sox2 and knockdown of NEAT1 reduced CDK2-Sox2 interactions with NEAT1 overexpression increasing the CDK2-Sox2 interactions (Ouyang et al., 2015; H. Zhang et al., 2021). Sox2 interacts with Oct4 to regulate expression of pluripotency associated genes (Rizzino & Wuebben, 2016); conversely, the CDK2 phosphorylation of Sox2 is thought to reduce the Sox2-Oct4 interaction (H. Zhang et al., 2021). Sox2 and CDK2 both associate with NEAT1 which may form an RNP complex, and together, implies NEAT1 can act as a scaffold RNA that facilitates enzymatic complex formation for CDK2 driven serine phosphorylation of Sox2.

One protein that interacts with the lncRNA RMST is Sox9 which facilitates Sox9 degradation in lung cancer cells (Pei et al., 2021). RMST knockdown or overexpression did not significantly impact Sox9 mRNA levels; however, RMST knockdown or overexpression increased and decreased Sox9 protein levels respectively. Further, RMST promotes the interaction between Sox9 and the ubiquitin ligase, FBW7. Upon knockdown of FBW7, Sox9 protein expression increased, and ubiquitination levels decreased. Importantly, Sox9 and FBW7 were both identified to associate with RMST. Collectively, this suggests that RMST may act as a scaffold to promote enzymatic ubiquitination of Sox9.

Conversely, Sox4 associates with lncRNA TGLC15 which reduced Sox4 protein degradation in hepatocellular carcinoma cells with Sox4 only associating with a specific TGLC15 fragment (Chen et al., 2020). Overexpression of TGLC15 reduced Sox4 ubiquitination levels and when TGLC15 was knocked down, Sox4 ubiquitination levels increased with proteasome mediated degradation of Sox4 protein. Together this suggests that the Sox4-TGLC15 association protects Sox4 from being ubiquitinated; however, the mechanism of protection was not further explored.

Like Sox4, LEF1's association with lncRNA OXCT1-AS1 prevents proteasome mediated degradation of LEF1 (B. Li et al., 2021). LEF1 predominately associates with a specific site of OXCT1-AS1. When OXCT1-AS1 was knocked down, LEF1 ubiquitination increased and LEF1 protein levels decreased. Further, OXCT1-AS1 knockdown increased the interaction between LEF1 and NARF (a ubiquitin-ligase known to interact with TCF/LEF proteins) (Yamada et al., 2006). Interestingly, LEF1 ubiquitination sites overlap with predicted LEF1-OXCT1-AS1 binding sites, together suggesting that LEF1's association with OXCT1-AS1 sterically blocks NARF's ability to ubiquitinate LEF1.

HMGB1 likely also uses RNA to form complex assemblies. HMGB1 associates with oncogenic lncRNA MALAT-1 in a multiple myeloma cell line and when MALAT-1 was knocked down, HMGB1 ubiquitination was increased resulting in increased HMGB1 protein degradation (Gao et al., 2017). Whether the HMGB1-MALAT-1 association is directly responsible for the protection of HMGB1 or through another pathway was not explored

(Gao et al., 2017). HMGB1 also associates with the brain-specific lncRNA, BS-DRL1, in neurons which is essential for the formation of DNA damage repair complexes and genomic stability (Lou et al., 2021). HMGB1 associates with BS-DRL1 through the N-terminal HMGB1-A domain and upon knockdown of BS-DRL1, HMGB1 was found to associate with significantly less chromatin, had reduced genomic occupancy at sites of DNA damage, and reduced interaction with a DNA repair protein. This suggests BS-DRL1 acts as a scaffold to recruit and form DNA repair protein complexes which include HMGB1. Together, this suggests that HMGB1 may use RNA to increase PPIs.

4.1.2. Possible mechanisms for how RNA impacts HMGB-protein interactions

—Many cellular processes require formation of complexes, yet the process for how individual complex components find one another remains under investigation. One proposed mechanism is that first RBPs interact with RNA which then utilize PPIs to recruit proteins together (Chen & Mayr, 2022). RNAs that contain multiple RBP binding sites appear to also be important for maintaining and/or increasing the interaction with RBPs (Chen & Mayr, 2022). RBP-RBP interactions can also be increased by binding to the same RNA which may facilitate enzymatic PTMs as many enzymes have RNA binding activity (Albihlal & Gerber, 2018; Ciesla, 2006; Curtis & Jeffery, 2021; Hildebrandt et al., 2017). Current literature supports HMGBs likely utilize a mixture of these mechanisms.

While the knowledge about HMGB-RNA scaffold biology is limited in scope, there are available mechanistic insights. Thus far, we know that HMGBs can bind to multiple sites with similar affinities within the same RNA which are important for RNP assemblies that contain both RBPs and non-RBPs (Cajigas et al., 2021; Hamilton et al., 2022; Holmes et al., 2020; Khoury et al., 2021). Sox2 interacts with scaffold RNAs to recruit and promote PPIs which regulate intrachromosomal contacts (Cajigas et al., 2021). HMGB1's interaction with DNA repair proteins at sites of DNA damage were also promoted by an RNA acting as a scaffold (Lou et al., 2021). Together this suggests HMGB-RNA interactions are likely essential to regulate processes that manipulate DNA, such as DNA repair (Figure 5A-C). Sox2, Sox9, and HMGB1 all associate with RNAs that impact their PTM landscape further supporting that HMGBs likely use RNAs as scaffolds to facilitate interactions with enzymatic machinery, perhaps forming enzyme PTM factories (Figure 5D-E). It's also possible that RNA may sterically prevent PPIs to adjust enzymatic modifications impacting the PTM landscape of HMGBs as in the case of LEF1 or Sox4. How often HMGBs utilize RNA for either mechanism remains unclear; however, evidence described in previous sections suggests RNA is central for regulating HMGB PPIs.

4.2 HMGBs are involved in RNA splicing

The HMGB family may also be a family of splicing factors, as HMGBs bind and associate with RNA, are highly localized within RNP granules where RNA processing occurs, and many HMGB partners are RBPs directly involved in splicing. Importantly, multiple lines of evidence show that several HMGBs directly impact splicing outcomes, including a survey that identifies HMGBs as involved in alternative splicing (Han et al., 2017).

4.2.1 Experimental evidence for HMGBs as splicing factors—Sox2 binds to RNA to directly promote alternative splicing during induced pluripotent stem cell generation (Hou et al., 2020). Removal of the HMGB-DBD or a 60-residue region C-terminal of the HMGB-DBD alters splicing, reduces the ability to reprogram somatic cells to induced pluripotent stem cells, and impacts exon selection during reprogramming. Significantly, Sox2 binds some of the alternatively spliced genes *in vivo*, providing strong evidence that Sox2 binding RNA directly impacts splicing outcomes. Mechanistically, Sox2 has proximal genomic occupancy to 91% of the alternatively spliced genes, indicating Sox2 may bind these RNAs during transcription for co-transcriptional RNA processing. In a study where Sox2 binds RNA *in vitro*, ectopic expression of Sox2 in transitional carcinoma cell lines increases 5' splice site selection of a non-endogenous splicing reporter RNA compared to the cell lines with lower Sox2 expression (Tung et al., 2010). Sox2 also associates with the lncRNA MIAT and with the MIAT promoter DNA to regulate MIAT transcription in esophageal squamous carcinoma cells (Zhang et al., 2020); however, the mechanistic details or functions of Sox2-MIAT associations are not known, suggesting Sox2 interacts with nascent RNA.

Sox6 has yet to be shown to interact with RNA *in vivo*; however, circumstantial evidence suggests Sox6 impacts splicing outcomes by binding RNA. Sox6 co-localizes to nuclear speckles with splicing factors and U snRNAs (Ohe et al., 2002). Splicing activity was inhibited when Sox6 was depleted from nuclear extracts, which may be attributed to spliceosome assembly as without Sox6 present, assembly of spliceosome complexes was inhibited. When just the HMGB-DBD of Sox6 was added to nuclear extracts depleted of native Sox6, splicing activity was partially restored, suggesting that the HMGB-DBD is in part responsible for Sox6's role in splicing.

Sox6 also reduces inhibition of pre-mRNA splicing by Dax-1 (Ohe et al., 2009). When Dax-1 was added to *in vitro* splicing assays, splicing was inhibited in a Dax-1 dose dependent manner and when the Sox6 HMGB-DBD was added to the assays, inhibition of splicing was reduced. Dax-1 and full length Sox6 interact through Sox6's coiled-coiled domain and not through the HMGB-DBD (Ohe et al., 2009), suggesting that the Sox6-Dax-1 interaction is not required to remove splicing inhibition. Dax-1 inhibition of splicing and the subsequent rescue of splicing by Sox6 was also observed *in vivo*. Dax-1 directly interacts with the splicing factor U2AF65, and upon co-transfection of the HMGBs Sox6, Sox9, or SRY with Dax-1, the Dax-1-U2AF65 interaction was greatly reduced. This suggests that the HMGB-DBD common to the Sox proteins may be responsible as only Sox6 has a coiled-coiled domain that can interact with Dax-1. This suggests that HMGB-DBD binding to RNA may block the Dax-1-U2AF65 interaction.

Sox9 associates with RNA to impact splicing outcomes (Girardot et al., 2018). In cell-based assays, knockdown of Sox9 results in hundreds of splicing changes including exon inclusion and exclusion, alternative 5' - and 3' -splice site selection, and inclusion of introns (Girardot et al., 2018; Rahmoun et al., 2017). Sox9's transcriptional regulation and impact on alternative splicing appear to be separate functions as there is little overlap between the genes impacted by Sox9 knockdown. A mutational analysis of Sox9 with 15 mutants and two mRNA splicing reporters revealed that the HMGB-DBD was directly involved in

splicing as many of the mutations within the HMGB-DBD impacted splicing outcomes (Girardot et al., 2018). Additionally, certain mutants abolished transcriptional activity but not splicing activity, suggesting separate DNA and RNA binding functions. Further, Sox9 binds near the transcription start site of a subset of the alternatively spliced RNAs yet did not change the mRNA expression levels of these genes, suggesting a role of DNA binding other than promoting or repressing transcription. Perhaps Sox9's DNA binding activity helps to recruit Sox9 to operate as a splicing factor in nascent RNA processing. Of the 154 mRNAs with differential splicing between wild type and a Sox9 knockout, Sox9 ChIP data shows that 45% were genes bound by Sox9 further suggesting Sox9 may bind nascent RNAs as a consequence of localization via DNA binding (Rahmoun et al., 2017).

Sox9 interacts with multiple RNA binding proteins known to be splicing factors, including Y14, a component of the exon junction complex which is involved in splicing (Girardot et al., 2018; Hata et al., 2008). Knockdown of Sox9 or Y14 results in ~60% overlap of alternatively spliced events with an enrichment of changes to exon cassettes and of these, 80% of exon inclusion or exclusion were in the same direction. Another report shows Sox9 colocalizes to nuclear paraspeckles with RNA binding protein p54nrb, and when either Sox9 or p54nrb are overexpressed, an alternatively spliced variant of Col2a1 mRNA was detected (Hata et al., 2008). Sox9 also localizes to lateral loops of lamp brush chromosomes which are transcriptionally active and are coated in nascent RNP fibrils with Sox9's localization being RNA dependent (Penrad-Mobayed et al., 2018). Altogether, these studies suggest that Sox9 is a splicing factor, possibly operating through a co-transcriptional mechanism.

HMGB1 acts as a splicing factor by binding RNA which impact splicing outcomes of senescence related RNAs (Sofiadis et al., 2021). When HMGB1 was knocked down, ~6000 different splicing events were observed. HMGB1 binds ~9% of ~4000 alternatively spliced transcripts during senescence entry, with splicing mediated through alternative transcription start sites, alternative polyadenylation, core exon usage, and alternative 5' - or 3' -splice site selection. HMGB1 binds genomic regions that overlap with ~14% of the transcripts bound by HMGB1, suggesting these interactions could be with nascent RNA during co-transcriptional splicing. In another example, HMGB1 binds the L-3 ribozyme and inhibits cleavage of a small nine nucleotide RNA substrate *in vitro*; however, the mechanism of inhibition was never addressed (Bell et al., 2008). HMGB3 binds to HIV-1 *Tat* mRNA and when HMGB3 was knocked down, splicing at a site directly bound by HMGB3 was increased by ~9 fold (Khoury et al., 2021). Due to high sequence similarity, this suggests that HMGB2 and HMGB4 could also be directly involved in splicing by binding RNA.

4.2.2 Possible HMGB mediated splicing mechanisms—Transcription and RNA processing are increasingly understood to be coupled events (Boumpas et al., 2022; Brody et al., 2011; Brody & Shav-Tal, 2011; Kornblihtt, 2006; Kornblihtt et al., 2004; Neugebauer, 2019; Tellier et al., 2020). Co-transcriptional splicing is estimated to occur at least ~80% of the time with many mRNAs maintaining their localization at the site of transcription while undergoing processing (Brody et al., 2011; Neugebauer, 2019; Rambout & Maquat, 2020). SRY, Sox2, Sox6, Sox9, HMGB1 and HMGB3 are implicated as splicing factors with HMGB-RNA associations likely increased by the nearby localization of HMGBs to RNAs undergoing processing. Sox2, Sox9, and HMGB1 all associate with both the genomic

DNA and the alternatively spliced RNAs transcribed from these genomic regions, suggesting these HMGB-RNA interactions occur co-transcriptionally with nascent RNA (Figure 6). In support of this idea, multiple HMGBs bind to newly transcribed or nascent RNA (Bao et al., 2018). Unfortunately, it is not yet known whether the HMGB-RNA interactions observed directly impact splicing outcomes by binding RNA or are simply observed because of proximity.

No direct HMGB-mediated *in vivo* splicing mechanism has been demonstrated to date. Thus, non-HMGB dual DNA and RNA binding proteins that impact splicing and other known splicing mechanisms are used to infer potential HMGB splicing mechanisms such as by modifying nascent RNA structure or the ability to form splicing complexes (Auboeuf et al., 2005; Georgakopoulos-Soares et al., 2022; Han et al., 2017; Pandit et al., 2008; Rambout & Maquat, 2020; Saha et al., 2020; Scharfen & Neugebauer, 2021; Tang et al., 2021). As HMGBs likely interact with nascent RNA and may change the structure of RNA, it's possible that HMGBs direct manipulation of nascent RNA structure and/or the steric blocking or recruitment of splicing factors could alter splicing outcomes (Figure 6A-C). Sox family proteins, HMGB1, HMGB3, and other HMGBs all bind stem loop RNAs which are common RNA elements known to act as protein binding hubs during RNA processing, including splicing decisions (Georgakopoulos-Soares et al., 2022; Kralovicova et al., 2015; Saha et al., 2020; Tang et al., 2021). Further, Sox6 appears to be important for spliceosome assembly. Perhaps, HMGBs bind and change the structure of nascent RNA, which impacts splicing factor assembly to influence splicing outcomes (Ohe et al., 2002). It should be noted that TFs and ChRPs can also impact splicing by interacting with DNA or splicing factors near the nascent RNA through a variety of mechanisms suggesting HMGBs could also impact splicing without binding RNA (Allemand et al., 2016; Kornblihtt, 2006; Ramanouskaya & Grinev, 2017; Zraly & Dingwall, 2012). Together, certain HMGBs likely act as splicing factors by binding RNA; however, determining the mechanism(s) HMGBs use to drive splicing decisions is an important question for further study.

4.3 HMGB interactions with RNA are important for HMGB localization

A protein's function is intimately tied to their cellular localization. As HMGBs are localized to RNP assemblies, interact with RNA and RBPs, we expect that HMGB interactions with RNA will be linked to HMGB localization. Multiple HMGBs from both the sequence specific TF and non-sequence specific ChRPs subfamilies interact with RNA, resulting in modifications to their cellular and/or genomic localization, suggesting similar HMGB localization mechanisms may be utilized throughout the family.

4.3.1 Experimental evidence suggests RNA is essential for HMGB genomic localization—Sox2 associates with the lncRNA LincQ to promote the maintenance of pluripotency in mouse embryonic stem cells (Jing et al., 2020). Sox2 associates with the promoter regions (–1500–0 bp) of the *LincQ* gene to regulate LincQ expression. Further, LincQ knockdown reduced Sox2 and RNA polymerase II genomic co-occupancy with downregulation of genes bound by Sox2, suggesting that the Sox2-LincQ interaction is important for Sox2's genomic localization to activate gene expression. Sox2 specifically associates with Exon 1 of LincQ, suggesting specificity for sequence or structural features

within Exon 1. A biotinylated-RNA *in vitro* cell lysate pull-down experiment reveals that Sox2 associates with LincQ through a region ~100 residues C-terminal of the HMGB domain; however, this region could also promote protein-protein interactions (PPI) that facilitate specific Sox2 associations with LincQ (Cox et al., 2010; Jing et al., 2020).

Sox2 associates with another lncRNA, linc1614, and the PRC2 component EZH2, which co-occupy genomic regions and together promote the repression of developmental genes to maintain pluripotency in mouse embryonic stem cells (Guo et al., 2018). Sox2 binds the promoter of linc1614 and regulates the expression of linc1614 with linc1614 downregulated upon Sox2 knockdown. A linc1614 knockdown reduced Sox2 and EZH2 genomic localization at specific sites, although it is not known if linc1614 occupancy changes upon Sox2 knockdown. Sox2 associates specifically with the 5' end of linc1614 with only this fragment being required to maintain stem cell morphology upon wild type linc1614 knockdown. Sox2, Linc1614, and EZH2 co-occupy promoters/enhancers of developmental genes, suggesting linc1614 may act as a scaffold to repress genes.

Another lncRNA, RMST, associates with Sox2 to promote neuronal differentiation in human embryonic stem cells (Ng et al., 2013). RMST and Sox2 regulate the expression of overlapping genes in differentiating neuronal stem cells as shown by independent knockdowns of RMST and Sox2. Sox2 and RMST also associate with overlapping genomic regions, including the promoters of genes involved in neurogenesis. RMST knockdown results in loss of ~50% of Sox2's genomic binding sites. From another study, Sox2 associates with RMST in neural crest cells and upon knockdown of RMST, Sox2's genomic localization was modified (Stamou et al., 2020). Another study identified a Sox2-RMST association which was thought to co-regulate the expression of microRNA; however, the impact on localization was not explored (Zhou et al., 2022). The lncRNA, linc1548, also associates with Sox2 and upon knockdown of linc1548, Sox2's localization was lost at genes important for neuronal differentiation (Bai et al., 2022). Taken together, RMST and linc1548 appear to be necessary for Sox2 genomic localization to drive neuronal differentiation.

Sox2 associates with lncRNA MEG3 which likely increases bone morphogenetic protein 4 (BMP4) transcription in mesenchymal stem cells obtained from multiple myeloma patients (Zhuang et al., 2015). Sox2 associates with the BMP4 promoter and represses BMP4 transcription as Sox2 knockdown resulted in increased BMP4 transcription. Sox2 and MEG3 associate with overlapping genomic regions within the BMP4 promoter and upon knockdown of MEG3, Sox2's occupancy was increased at this promoter. Overexpression of MEG3 resulted in increased BMP4 transcription and when MEG3 was knocked down, BMP4 expression was reduced. Together these data suggest that Sox2 functions as a repressor of BMP4 transcription and MEG3 increases transcription of BMP4 by sequestering Sox2 from BMP4 promoters or by blocking Sox2 access to the promoter.

Sox2 associates with the oncogenic enhancer lncRNA, CCAT1, which promotes tumorigenicity in squamous cell carcinoma cell lines (Jiang et al., 2018). Sox2 associates with tumor protein 63 (TP63) to activate CCAT1's transcription with knockdown of Sox2 or TP63 resulting in reduced CCAT1 expression. Sox2/TP63/CCAT1 form a complex and co-occupy the EGFR super-enhancer to activate EGFR expression. Upon knockdown of

CCAT1, Sox2 and TP63 occupancy at EGFR super enhancers was reduced; however, CCAT1 did not impact Sox2-TP63 interactions. Knockdown of Sox2 or TP63 reduced CCAT1 occupancy at the EGFR super enhancer. Sox2 also associates with Lnc-RP11 in colorectal cancer cells and when Lnc-RP11 was knocked down Sox2's localization at the USP7 promoter was reduced (Q. Li et al., 2021). Together, these studies suggest an interdependency between genomic localization of Sox2 and lncRNA; however, the mechanisms were not further explored.

Likewise, HMGB2 has been reported to have functional interactions with RNA that appear to occur through similar pathways. For example, HMGB2 associates with the *Lrp1*-antisense (*Lrp1*-AS) ncRNA resulting in a decrease in *Lrp1* mRNA expression (Yamanaka et al., 2015). HMGB2 and TF Srebp1a interact to regulate *Lrp1* gene expression by increasing Srebp1a occupancy at *Lrp1* promoter regions while *Lrp1*-AS RNA reduces *Lrp1* expression. Overexpression of *Lrp1*-AS reduces promoter occupancy of Srebp1a and a *Lrp1*-AS KD increases occupancy. HMGB2 likely associates with a specific region of *Lrp1*-AS as antisense oligonucleotides targeting a specific region reduced HMGB2-*Lrp1*-AS associations and increased *Lrp1* expression. Together, these data suggest that *Lrp1*-AS acts a molecular decoy to sequester HMGB2 from Srebp1a, reducing *Lrp1* expression.

4.3.2 Experimental evidence suggests RNA is essential for HMGB cellular localization—HMGB1-RNA associations also promote hepatitis C virus (HCV) replication in HCV infected cells (Yu et al., 2015). HMGB1 associates with the 5'-UTR stem loop 4 of the positive strand HCV RNA in cell lysates and binds the RNA *in vitro*, suggesting specificity for certain RNA features. HMGB1 is mostly localized to the nucleus; however, upon HCV infection, a population of HMGB1s are co-localized in the cytoplasm with the HCV RNA, suggesting the HCV RNA recruits or maintains HMGB1 proteins in the cytoplasm. Interestingly, the sole HMGB1-A domain binds the 5' UTR stem loop 4 while HMGB1-B domain appeared to not be required for associations in cell-based assays. Finally, the HMGB1-A domain is in part responsible for increasing the HCV replication in these cell-based assays. HMGB1 may also act as immune response proteins by binding to immunogenic nucleic acids, including poly (I:C) dsRNA and poly (U) ssRNA (Yanai et al., 2009). Full length HMGB1 and the individual HMGB1-A domain bind to siRNA, forming complexes with carrier molecules that allow for siRNA delivery into cells with this delivery method successfully reducing expression of the targeted genes (Choi et al., 2020; Lee et al., 2012; Oh & Lee, 2014). These functionally unrelated behaviors illustrate the multifaceted potential of HMGB1-RNA interactions. Together, these studies suggest HMGB1-RNA interactions impact HMGB1 cellular localization.

HMGB2 associates with the lncRNA CRCMSL which may impact HMGB localization and protein-protein interactions in colorectal cancer cell lines (Han et al., 2019). CRCMSL is primarily localized to the cytoplasm and when CRCMSL is over-expressed, HMGB2 localization is reduced in the nucleus and increased in the cytoplasm. When CRCMSL was downregulated, HMGB2 nuclear localization increases, while HMGB2 cytoplasm localization was decreased. Together this suggests the HMGB2-CRCMSL association impacts HMGB2 localization and protein-protein interactions; however, the direct mechanism was not further explored.

4.3.3 Possible HMGB-RNA mechanisms for HMGB localization—Mounting evidence shows that RNA is essential for cellular localization of TFs and ChRPs (Cassiday & Maher, 2002; Mourtada-Maarabouni et al., 2009; Noh et al., 2018; Sigova et al., 2015). One model suggests that stable, processed RNAs and/or nascent RNAs transcribed from regulatory regions such as enhancers promote formation of transcriptional hubs which can recruit or maintain proteins involved in transcription and RNA metabolism (Bouwman et al., 2022; Palacio & Taatjes, 2022; Quinodoz et al., 2021; Sharp et al., 2022). HMGBs' interactions with RNA appear to be involved, at least in part, with this model.

Currently, the transcription hub model is supported by multiple lines of evidence. First, HMGBs bind newborn or nascent RNA and associate with the genomic DNA and the RNAs transcribed nearby, further supporting HMGB-RNA interactions occur with nascent and/or newborn RNA. Significantly, knockdown of HMGB interacting RNAs directly impacts HMGBs genomic localization. Moreover, Sox2-RNA interactions promote and/or maintain chromosomal interactions which are thought to occur at sites of transcription hubs (Cajigas et al., 2021; Richter et al., 2022; Tang et al., 2022). Likewise, Sox2 enhancer clusters overlap with regions enriched with Pol II (Liu et al., 2014). Additionally, HMGBs form liquid condensates with transcriptional machinery *in vitro* (Boija et al., 2018; Feric et al., 2021) and HMGBs association with chromatin was modified when RNA Pol II driven nascent RNA transcription was inhibited (Skalska et al., 2021). Intriguingly, Sox family proteins are thought to be pioneer TFs that regulate gene expression by binding to enhancer regions (Kagawa & Kurumizaka, 2021; Soufi et al., 2015; Vanzan et al., 2021; Zaret & Carroll, 2011), perhaps Sox proteins regulate the expression of and bind to nascent enhancer RNAs as a means to jumpstart transcription hub formation. Together, this suggests HMGB-RNA interactions are essential for maintaining and/or recruiting HMGBs and other transcriptional machinery within or to transcriptional hubs (Figure 7). Notably, this model cannot describe all HMGB-RNA interaction mechanisms as HMGBs also act as transcriptional repressors and co-localize with RNAs at genomic sites trans from their transcription start sites.

Another, yet more speculative mechanism is that HMGBs and their interacting RNAs co-migrate throughout the cell to regulate gene expression. This notion is supported because Sox2 has genomic co-localization and associates with RMST, linc1614, MEG3 and CCAT1 RNAs. Excitingly, Sox2 regulates the expression of linc1614 and CCAT1 while also co-localizing with the RNAs at *cis* and *trans* genomic sites, suggesting that Sox2 may associate with linc1614 or CCAT1 as nascent/newborn RNAs and migrate as a complex to regulatory genomic sites. Perhaps these RNAs, comparable to mRNA export processes with messenger RNPs (Katahira, 2015), act as migratory scaffold RNPs that promotes or prevents enzymatic protein modifications or PPIs which results in a decrease or increase of HMGB TF/ChRP non-specific interactions with genomic DNA, RNA, or proteins. This could reduce protein search times, resulting in a more selective and responsive cellular environment.

Steady state RNA, which include stable processed ncRNAs and mRNAs, also appears to be important for HMGB localization. HMGB1 and HMGB2 proteins have decreased concentrations in the nucleus and increased in the cytoplasm in the presence of HCV RNA and CRCMSL, respectively. Whether these RNAs interact with HMGBs immediately after

translation in the cytoplasm or after nuclear export is not known as the HMGB PTM landscape is likely different and could impact interactions with RNA (Andersson et al., 2014; Richard et al., 2017; Williams et al., 2020). Collectively, HMGB-RNA interactions impact HMGB localization, likely through a multitude of mechanisms, possibly even within the same protein, suggesting a rich field of future discovery exists.

5. HMGB-RNA INTERACTIONS IN HUMAN DISEASE

HMGBs are involved in many human diseases that have been attributed to their DNA-binding activities; however, HMGBs interactions with RNA likely play a role as well (Lomeli & Castillo-Robles, 2016; Niu et al., 2020; Novak et al., 2020; Stevanovic et al., 2021; Taniguchi et al., 2018; Voong et al., 2021; Vozarikova et al., 2020). HMGB-RNA interactions occur in a variety of cancer cell lines with HMGB-RNA interactions likely impacting cancer progression in multiple myeloma cells (HMGB1(Gao et al., 2017) and Sox2 (Zhuang et al., 2015)), squamous cell carcinoma cells (Sox2, (Jiang et al., 2018)), transitional carcinoma cells (Sox2, (Tung et al., 2010)), hepatocellular carcinoma cells (Sox4, (Chen et al., 2020)), esophageal squamous carcinoma cells (Sox2, (Zhang et al., 2020)), lung cancer cells (LEF1(B. Li et al., 2021), Sox9 (Pei et al., 2021)) and colorectal cancer cells (Sox2,(Q. Li et al., 2021), HMGB2, (Han et al., 2019)). Multiple HMGBs are labeled as oncogenes, including the Sox and HMGB subfamilies (Moreno, 2020; Pouremamali et al., 2022; Weina & Utikal, 2014; Zhang & Sun, 2020; Zheng et al., 2022). Further, Sox2 and Sox10 associate with RNA in cells obtained from a Kallman syndrome patient (Stamou et al., 2020). HMGB1 also interacts with RNA from HCV infected cells (Yu et al., 2015). HMGBs are also involved in immune response pathways, perhaps their ability to bind RNA is important in these processes, like HMGB1 or HMGB3. All together this suggests HMGBs interactions with RNA and its related dysfunction likely directly impact human health, an important area for further study.

6. FUTURE CHALLENGES

Much remains unknown about HMGB-RNA biology and we hope that this review inspires a push to extend the boundaries of knowledge in this area. While the cumulative evidence clearly establishes that HMGBs interact with RNA to affect biological processes, key questions remain unanswered. Some of the missing pieces that need to be explored more fully to understand HMGB biology are as follows.

Certain questions are more easily understood using *in vitro* experimental approaches. For example, a central feature of the interaction of HMGBs with DNA is the induction of significant bends that are crucial for their function as TFs and ChRPs. Whether HMGBs induce structural changes in the RNA has not been established. No HMGB-RNA structure has been solved that would some shed light on potential structural changes; however, this is likely to be a particularly important and challenging problem. It's also possible that there's a connection between the sequence/structure of DNA and RNA that HMGBs prefer. While it appears that both SS and NSS HMGBs bind to stem loop and G-quadruplex RNAs, whether and how these *in vitro* specificities for RNA transfer to *in vivo* activities is not yet known.

Thus, it is essential to identify *in vivo*, site-specific HMGB RNA binding sites which will aid in building evidence-based models.

As shown in this review, some HMGBs have only circumstantial or association evidence that they are RBPs, thus, whether these proteins directly interact with RNA *in vivo* needs to be determined. Intriguingly, some HMGB studies suggest mechanistically unrelated DNA and RNA binding functions for individual proteins, yet there exists no “picture-perfect” separation of function mutant to date that would enable isolation of DNA and RNA binding functions *in vivo*. However, as they are so closely tied to the HMGB-DBD, this will be a very challenging problem to solve. These types of mutants will be invaluable for mechanistic studies of the RNA binding activity and function of HMGB-DBDs. A separate but interesting problem is why are HMGBs localized within RNP granules. Is this by chance when RNAs they are bound to form into RNP granules or is there an associated function? Finally, what is the role of partner proteins on the partitioning of HMGBs between RNA and DNA. In particular, isolated HMGB-DBDs appear to bind a broad spectrum of RNA motifs *in vitro*. Do partner proteins determine the specificity of an HMGB for RNA targets?

Some other more speculative, but interesting, questions remain. What is the distribution of HMGBs bound to DNA or RNA, is it possible that HMGBs are stochastically bound to RNA more than DNA *in vivo*? Could HMGB PTMs or RNA modifications impact RNA binding potential and/or selectivity? HMGB-RNA interactions appear to utilize several distinct mechanisms and excitingly, it’s possible these mechanisms may be used by other dual DNA and RNA binding proteins, further illustrating the importance of understanding HMGB-RNA interactions more in depth in the future.

Conclusion

Collectively, this review summarizes the current body of evidence that demonstrates the HMGB family is an RNA binding family. The finding that HMGBs widely and robustly directly interact with RNAs *in vitro* and *in vivo*, are localized to regions of RNA processing, and their protein partners are enriched with RBPs has been established from many studies using a broad spectrum of experimental approaches. This observation, however, raises many important questions as to the functional role of HMGB-RNA interactions. Furthermore, it also increasingly clear that other TFs similarly have both robust DNA and RNA binding activities that are crucial for their cellular function. This review thus illustrates a key challenge in understanding the regulation of the genome by the transcriptome by highlighting the breadth of evidence that HMGBs are dual DNA and RNA binding proteins whose full range of biological activities are yet to be fully understood.

Acknowledgements.

We thank Dr. Dylan Taatjes and Dr. Stephanie Moon for their feedback and guidance. Figures 3, 5, 6, and 7 were created in part using [Biorender.com](https://biorender.com).

Funding Information.

This review was sustained by a grant from the National Institutes of Health to D.S.W. and R.T.B. (R01 GM120347). D.J.H. was further supported through the Molecular Biophysics Training Program, funded through the

National Institutes of Health (T32 GM065103) and A.E.H. was further supported with the University of Colorado Undergraduate Research Opportunity Program and Biological Sciences Initiative.

References

- Agresti A, & Bianchi ME (2003, Apr). HMGB proteins and gene expression. *Curr Opin Genet Dev*, 13(2), 170–178. 10.1016/s0959-437x(03)00023-6 [PubMed: 12672494]
- Akhade VS, Pal D, & Kanduri C. (2017). Long Noncoding RNA: Genome Organization and Mechanism of Action. *Adv Exp Med Biol*, 1008, 47–74. 10.1007/978-981-10-5203-3_2 [PubMed: 28815536]
- Albihlal WS, & Gerber AP (2018, Sep). Unconventional RNA-binding proteins: an uncharted zone in RNA biology. *FEBS Lett*, 592(17), 2917–2931. 10.1002/1873-3468.13161 [PubMed: 29908064]
- Allemand E, Myers MP, Garcia-Bernardo J, Harel-Bellan A, Krainer AR, & Muchardt C. (2016, Sep). A Broad Set of Chromatin Factors Influences Splicing. *Plos Genetics*, 12(9). 10.1371/journal.pgen.1006318
- Alonso-Lopez D, Campos-Laborie FJ, Gutierrez MA, Lambourne L, Calderwood MA, Vidal M, & De Las Rivas J. (2019, Jan 1). APID database: redefining protein-protein interaction experimental evidences and binary interactomes. *Database (Oxford)*, 2019. 10.1093/database/baz005
- Alonso-Lopez D, Gutierrez MA, Lopes KP, Prieto C, Santamaria R, & De Las Rivas J. (2016, Jul 8). APID interactomes: providing proteome-based interactomes with controlled quality for multiple species and derived networks. *Nucleic Acids Res*, 44(W1), W529–535. 10.1093/nar/gkw363 [PubMed: 27131791]
- Andersson U, Antoine DJ, & Tracey KJ (2014, Nov). The functions of HMGB1 depend on molecular localization and post-translational modifications. *J Intern Med*, 276(5), 420–424. 10.1111/joim.12309 [PubMed: 25346011]
- Arimondo PB, Gelus N, Hamy F, Payet D, Travers A, & Bailly C. (2000, Nov 17). The chromosomal protein HMG-D binds to the TAR and RBE RNA of HIV-1. *FEBS Lett*, 485(1), 47–52. 10.1016/s0014-5793(00)02183-9 [PubMed: 11086163]
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, & Sherlock G. (2000, May). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*, 25(1), 25–29. 10.1038/75556 [PubMed: 10802651]
- Auboeuf D, Dowhan DH, Dutertre M, Martin N, Bergeret SM, & O'Malley BW (2005, Jul). A subset of nuclear receptor coregulators act as coupling proteins during synthesis and maturation of RNA transcripts. *Molecular and Cellular Biology*, 25(13), 5307–5316. 10.1128/Mcb.25.13.5307-5316.2005 [PubMed: 15964789]
- Bai ML, Li GP, Jiapaer Z, Guo XD, Xi JJ, Wang GY, Ye D, Chen W, Duan BY, & Kang JH (2022, Mar 3). Linc1548 Promotes the Transition of Epiblast Stem Cells Into Neural Progenitors by Engaging OCT6 and SOX2. *Stem Cells*, 40(1), 22–34. 10.1093/stmcls/sxab003 [PubMed: 35511866]
- Balcerak A, Trebinska-Stryjewska A, Konopinski R, Wakula M, & Grzybowska EA (2019, Jun 28). RNA-protein interactions: disorder, moonlighting and junk contribute to eukaryotic complexity. *Open Biol*, 9(6), 190096. 10.1098/rsob.190096
- Baltz AG, Munschauer M, Schwanhausser B, Vasile A, Murakawa Y, Schueler M, Youngs N, Penfold-Brown D, Drew K, Milek M, Wylter E, Bonneau R, Selbach M, Dieterich C, & Landthaler M. (2012, Jun 8). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell*, 46(5), 674–690. 10.1016/j.molcel.2012.05.021 [PubMed: 22681889]
- Bao X, Guo X, Yin M, Tariq M, Lai Y, Kanwal S, Zhou J, Li N, Lv Y, Pulido-Quetglas C, Wang X, Ji L, Khan MJ, Zhu X, Luo Z, Shao C, Lim DH, Liu X, Li N, Wang W, He M, Liu YL, Ward C, Wang T, Zhang G, Wang D, Yang J, Chen Y, Zhang C, Jauch R, Yang YG, Wang Y, Qin B, Anko ML, Hutchins AP, Sun H, Wang H, Fu XD, Zhang B, & Esteban MA (2018, Mar). Capturing the interactome of newly transcribed RNA. *Nat Methods*, 15(3), 213–220. 10.1038/nmeth.4595 [PubMed: 29431736]
- Beckmann BM, Horos R, Fischer B, Castello A, Eichelbaum K, Alleaume AM, Schwarzl T, Curk T, Foehr S, Huber W, Krijgsveld J, & Hentze MW (2015, Dec 3). The RNA-binding proteomes from

- yeast to man harbour conserved enigmRBPs. *Nat Commun*, 6, 10127. 10.1038/ncomms10127 [PubMed: 26632259]
- Bell AJ Jr., Chauhan S, Woodson SA, & Kallenbach NR (2008, Dec 5). Interactions of recombinant HMGB proteins with branched RNA substrates. *Biochem Biophys Res Commun*, 377(1), 262–267. 10.1016/j.bbrc.2008.09.131 [PubMed: 18845125]
- Bernard P, & Harley VR (2010, Mar). Acquisition of SOX transcription factor specificity through protein-protein interaction, modulation of Wnt signalling and post-translational modification. *Int J Biochem Cell Biol*, 42(3), 400–410. 10.1016/j.biocel.2009.10.017 [PubMed: 19854293]
- Bianchi ME, & Agresti A. (2005, Oct). HMG proteins: dynamic players in gene regulation and differentiation. *Curr Opin Genet Dev*, 15(5), 496–506. 10.1016/j.gde.2005.08.007 [PubMed: 16102963]
- Bianchi ME, & Beltrame M. (1998, Dec). Flexing DNA: HMG-box proteins and their partners. *Am J Hum Genet*, 63(6), 1573–1577. 10.1086/302170 [PubMed: 9837808]
- Boija A, Klein IA, Sabari BR, Dall'Agnesse A, Coffey EL, Zamudio AV, Li CH, Shrinivas K, Manteiga JC, Hannett NM, Abraham BJ, Afeyan LK, Guo YE, Rimel JK, Fant CB, Schuijers J, Lee TI, Taatjes DJ, & Young RA (2018, Dec 13). Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell*, 175(7), 1842–1855 e1816. 10.1016/j.cell.2018.10.042 [PubMed: 30449618]
- Boumpas P, Merabet S, & Carnesecchi J. (2022, Jul 28). Integrating transcription and splicing into cell fate: Transcription factors on the block. *Wiley Interdisciplinary Reviews-Rna*. 10.1002/wrna.1752
- Bouwman BAM, Crosetto N, & Bienko M. (2022, Feb). RNA gradients: Shapers of 3D genome architecture. *Curr Opin Cell Biol*, 74, 7–12. 10.1016/jceb.2021.12.001 [PubMed: 34998095]
- Bowles J, Schepers G, & Koopman P. (2000, Nov 15). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol*, 227(2), 239–255. 10.1006/dbio.2000.9883 [PubMed: 11071752]
- Brieno-Enriquez MA, Moak SL, Abud-Flores A, & Cohen PE (2019, Apr 1). Characterization of telomeric repeat-containing RNA (TERRA) localization and protein interactions in primordial germ cells of the mousedagger. *Biol Reprod*, 100(4), 950–962. 10.1093/biolre/iy0243 [PubMed: 30423030]
- Brody Y, Neufeld N, Bieberstein N, Causse SZ, Bohnlein EM, Neugebauer KM, Darzacq X, & Shav-Tal Y. (2011, Jan). The In Vivo Kinetics of RNA Polymerase II Elongation during Co-Transcriptional Splicing. *Plos Biology*, 9(1). 10.1371/journal.pbio.1000573
- Brody Y, & Shav-Tal Y. (2011). Transcription and splicing When the twain meet. *Transcription-Austin*, 2(5), 216–220. 10.4161/trns.2.5.17273
- Brown RS (2005, Feb). Zinc finger proteins: getting a grip on RNA. *Curr Opin Struct Biol*, 15(1), 94–98. 10.1016/j.sbi.2005.01.006 [PubMed: 15718139]
- Brown TA, Tkachuk AN, & Clayton DA (2015). Mitochondrial Transcription Factor A (TFAM) Binds to RNA Containing 4-Way Junctions and Mitochondrial tRNA. *PLoS One*, 10(11), e0142436. 10.1371/journal.pone.0142436
- Budowsky EI, Axentyeva MS, Abdurashidova GG, Simukova NA, & Rubin LB (1986, Aug 15). Induction of polynucleotide-protein cross-linkages by ultraviolet irradiation. Peculiarities of the high-intensity laser pulse irradiation. *Eur J Biochem*, 159(1), 95–101. 10.1111/j.1432-1033.1986.tb09837.x [PubMed: 2427338]
- Cajigas I, Chakraborty A, Lynam M, Swyter KR, Bastidas M, Collens L, Luo H, Ay F, & Kohtz JD (2021, Mar 15). Sox2-Evf2 lncRNA-mediated mechanisms of chromosome topological control in developing forebrain. *Development*, 148(6). 10.1242/dev.197202
- Campos-Melo D, Hawley ZCE, Droppelmann CA, & Strong MJ (2021). The Integral Role of RNA in Stress Granule Formation and Function. *Front Cell Dev Biol*, 9, 621779. 10.3389/fcell.2021.621779
- Cassiday LA, & Maher LJ (2002, Oct 1). Having it both ways: transcription factors that bind DNA and RNA. *Nucleic Acids Research*, 30(19), 4118–4126. 10.1093/nar/gkf512 [PubMed: 12364590]
- Castello A, Fischer B, Eichelbaum K, Horos R, Beckmann BM, Strein C, Davey NE, Humphreys DT, Preiss T, Steinmetz LM, Krijgsvelde J, & Hentze MW (2012, Jun 8). Insights into RNA

- biology from an atlas of mammalian mRNA-binding proteins. *Cell*, 149(6), 1393–1406. 10.1016/j.cell.2012.04.031 [PubMed: 22658674]
- Castello A, Fischer B, Frese CK, Horos R, Alleaume AM, Foehr S, Curk T, Krijgsveld J, & Hentze MW (2016, Aug 18). Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol Cell*, 63(4), 696–710. 10.1016/j.molcel.2016.06.029 [PubMed: 27453046]
- Castello A, Hentze MW, & Preiss T. (2015, Dec). Metabolic Enzymes Enjoying New Partnerships as RNA-Binding Proteins. *Trends in Endocrinology and Metabolism*, 26(12), 746–757. 10.1016/j.tem.2015.09.012 [PubMed: 26520658]
- Catez F, Yang H, Tracey KJ, Reeves R, Misteli T, & Bustin M. (2004, May). Network of dynamic interactions between histone H1 and high-mobility-group proteins in chromatin. *Mol Cell Biol*, 24(10), 4321–4328. 10.1128/MCB.24.10.4321-4328.2004 [PubMed: 15121851]
- Caudron-Herger M, Rusin SF, Adamo ME, Seiler J, Schmid VK, Barreau E, Kettenbach AN, & Diederichs S. (2019, Jul 11). R-DeeP: Proteome-wide and Quantitative Identification of RNA-Dependent Proteins by Density Gradient Ultracentrifugation. *Mol Cell*, 75(1), 184–199 e110. 10.1016/j.molcel.2019.04.018 [PubMed: 31076284]
- Chaoui A, Kavou A, Baral V, Watanabe Y, Lecerf L, Colley A, Mendoza-Londono R, Pingault V, & Bondurand N. (2015, Sep 1). Subnuclear re-localization of SOX10 and p54NRB correlates with a unique neurological phenotype associated with SOX10 missense mutations. *Hum Mol Genet*, 24(17), 4933–4947. 10.1093/hmg/ddv215 [PubMed: 26060192]
- Chen KM, Campbell E, Pandey RR, Yang Z, McCarthy AA, & Pillai RS (2015, May). Metazoan Maelstrom is an RNA-binding protein that has evolved from an ancient nuclease active in protists. *RNA*, 21(5), 833–839. 10.1261/rna.049437.114 [PubMed: 25778731]
- Chen R, Chen Q, Kim H, Siu KH, Sun Q, Tsai SL, & Chen W. (2014, Aug). Biomolecular scaffolds for enhanced signaling and catalytic efficiency. *Curr Opin Biotechnol*, 28, 59–68. 10.1016/j.copbio.2013.11.007 [PubMed: 24832076]
- Chen X, & Mayr C. (2022, Jan). A working model for condensate RNA-binding proteins as matchmakers for protein complex assembly. *RNA*, 28(1), 76–87. 10.1261/rna.078995.121 [PubMed: 34706978]
- Chen Y, & Belmont AS (2019, Apr). Genome organization around nuclear speckles. *Curr Opin Genet Dev*, 55, 91–99. 10.1016/j.gde.2019.06.008 [PubMed: 31394307]
- Chen Y, Huang F, Deng L, Tang Y, Li D, Wang T, Fan Y, Tao Q, & Tang D. (2020, Jan). Long non-coding RNA TGLC15 advances hepatocellular carcinoma by stabilizing Sox4. *J Clin Lab Anal*, 34(1), e23009. 10.1002/jcla.23009
- Choi M, Jeong H, Kim S, Kim M, Lee M, & Rhim T. (2020, Jan 2). Targeted delivery of Chil3/Chil4 siRNA to alveolar macrophages using ternary complexes composed of HMG and oligoarginine micelles. *Nanoscale*, 12(2), 933–943. 10.1039/c9nr06382j [PubMed: 31840707]
- Ciesla J. (2006). Metabolic enzymes that bind RNA: yet another level of cellular regulatory network? *Acta Biochim Pol*, 53(1), 11–32. <https://www.ncbi.nlm.nih.gov/pubmed/16410835> [PubMed: 16410835]
- Connor F, Cary PD, Read CM, Preston NS, Driscoll PC, Denny P, Crane-Robinson C, & Ashworth A. (1994, Aug 25). DNA binding and bending properties of the post-meiotically expressed Sry-related protein Sox-5. *Nucleic Acids Res*, 22(16), 3339–3346. 10.1093/nar/22.16.3339 [PubMed: 8078769]
- Corbet GA, & Parker R. (2019). RNP Granule Formation: Lessons from P-Bodies and Stress Granules. *Cold Spring Harb Symp Quant Biol*, 84, 203–215. 10.1101/sqb.2019.84.040329 [PubMed: 32482896]
- Corley M, Burns MC, & Yeo GW (2020, Apr 2). How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Mol Cell*, 78(1), 9–29. 10.1016/j.molcel.2020.03.011 [PubMed: 32243832]
- Cox JL, Mallanna SK, Luo X, & Rizzino A. (2010, Nov 12). Sox2 uses multiple domains to associate with proteins present in Sox2-protein complexes. *PLoS One*, 5(11), e15486. 10.1371/journal.pone.0015486
- Cruz-Gallardo I, Martino L, Kelly G, Atkinson RA, Trotta R, De Tito S, Coleman P, Ahdash Z, Gu Y, Bui TTT, & Conte MR (2019, May 7). LARP4A recognizes polyA RNA via a novel binding

mechanism mediated by disordered regions and involving the PAM2w motif, revealing interplay between PABP, LARP4A and mRNA. *Nucleic Acids Res*, 47(8), 4272–4291. 10.1093/nar/gkz144 [PubMed: 30820564]

Curtis NJ, & Jeffery CJ (2021, Jun 30). The expanding world of metabolic enzymes moonlighting as RNA binding proteins. *Biochem Soc Trans*, 49(3), 1099–1108. 10.1042/BST20200664 [PubMed: 34110361]

D GH, Kelley DR, Tenen D, Bernstein B, & Rinn JL (2016, Feb 16). Widespread RNA binding by chromatin-associated proteins. *Genome Biol*, 17, 28. 10.1186/s13059-016-0878-3 [PubMed: 26883116]

de Souza TA, Soprano AS, de Lira NP, Quaresma AJ, Pauletti BA, Paes Leme AF, & Benedetti CE (2012). The TAL effector PthA4 interacts with nuclear factors involved in RNA-dependent processes including a HMG protein that selectively binds poly(U) RNA. *PLoS One*, 7(2), e32305. 10.1371/journal.pone.0032305

Du Z, Wen X, Wang Y, Jia L, Zhang S, Liu Y, Zhou L, Li H, Yang W, Wang C, Chen J, Hao Y, Salgado Figueroa D, Chen H, Li D, Chen N, Celik I, Zhu Y, Yan Z, Fu C, Liu S, Jiao B, Wang Z, Zhang H, Gulsoy G, Luo J, Qin B, Gao S, Kapranov P, Esteban MA, Zhang S, Li W, Ay F, Chen R, Hoffman AR, Cui J, & Hu JF (2021, Aug 19). Chromatin lncRNA Platr10 controls stem cell pluripotency by coordinating an intrachromosomal regulatory network. *Genome Biol*, 22(1), 233. 10.1186/s13059-021-02444-6 [PubMed: 34412677]

Dvir S, Argoetti A, Lesnik C, Roytblat M, Shriki K, Amit M, Hashimshony T, & Mandel-Gutfreund Y. (2021, Jun 1). Uncovering the RNA-binding protein landscape in the pluripotency network of human embryonic stem cells. *Cell Rep*, 35(9), 109198. 10.1016/j.celrep.2021.109198

Faber GP, Nadav-Eliyahu S, & Shav-Tal Y. (2022, Jul 1). Nuclear speckles - a driving force in gene expression. *J Cell Sci*, 135(13). 10.1242/jcs.259594

Feric M, Demarest TG, Tian J, Croteau DL, Bohr VA, & Misteli T. (2021, Mar 15). Self-assembly of multi-component mitochondrial nucleoids via phase separation. *EMBO J*, 40(6), e107165. 10.15252/embj.2020107165

Fernandes N, & Buchan JR (2021, Apr 9). RNAs as Regulators of Cellular Matchmaking. *Frontiers in Molecular Biosciences*, 8. 10.3389/fmolb.2021.634146

Gaillard C, & Strauss F. (2000). High affinity binding of proteins HMG1 and HMG2 to semicatenated DNA loops. *BMC Mol Biol*, 1, 1. 10.1186/1471-2199-1-1 [PubMed: 11041984]

Galganski L, Urbanek MO, & Krzyzosiak WJ (2017, Oct 13). Nuclear speckles: molecular organization, biological function and role in disease. *Nucleic Acids Res*, 45(18), 10350–10368. 10.1093/nar/gkx759 [PubMed: 28977640]

Gao D, Lv AE, Li HP, Han DH, & Zhang YP (2017, Oct). LncRNA MALAT-1 Elevates HMGB1 to Promote Autophagy Resulting in Inhibition of Tumor Cell Apoptosis in Multiple Myeloma. *J Cell Biochem*, 118(10), 3341–3348. 10.1002/jcb.25987 [PubMed: 28295550]

Gene Ontology C. (2021, Jan 8). The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res*, 49(D1), D325–D334. 10.1093/nar/gkaa1113 [PubMed: 33290552]

Genzor P, & Bortvin A. (2015). A unique HMG-box domain of mouse Maelstrom binds structured RNA but not double stranded DNA. *PLoS One*, 10(3), e0120268. 10.1371/journal.pone.0120268

Georgakopoulos-Soares I, Parada GE, & Hemberg M. (2022). Secondary structures in RNA synthesis, splicing and translation. *Comput Struct Biotechnol J*, 20, 2871–2884. 10.1016/j.csbj.2022.05.041 [PubMed: 35765654]

Giese K, Amsterdam A, & Grosschedl R. (1991, Dec). DNA-binding properties of the HMG domain of the lymphoid-specific transcriptional regulator LEF-1. *Genes Dev*, 5(12B), 2567–2578. 10.1101/gad.5.12b.2567 [PubMed: 1752444]

Girardot M, Bayet E, Maurin J, Fort P, Roux P, & Raynaud P. (2018, Sep 28). SOX9 has distinct regulatory roles in alternative splicing and transcription. *Nucleic Acids Res*, 46(17), 9106–9118. 10.1093/nar/gky553 [PubMed: 29901772]

Goos H, Kinnunen M, Salokas K, Tan Z, Liu X, Yadav L, Zhang Q, Wei GH, & Varjosalo M. (2022, Feb 9). Human transcription factor protein interaction networks. *Nat Commun*, 13(1), 766. 10.1038/s41467-022-28341-5 [PubMed: 35140242]

- Guo JK, & Guttman M. (2022, Oct). Regulatory non-coding RNAs: everything is possible, but what is important? *Nat Methods*, 19(10), 1156–1159. 10.1038/s41592-022-01629-6 [PubMed: 36203023]
- Guo X, Wang Z, Lu C, Hong W, Wang G, Xu Y, Liu Z, & Kang J. (2018, Apr 1). LincRNA-1614 coordinates Sox2/PRC2-mediated repression of developmental genes in pluripotency maintenance. *J Mol Cell Biol*, 10(2), 118–129. 10.1093/jmcb/mjx041 [PubMed: 28992244]
- Hamilton DJ, Hein AE, Holmes ZE, Wuttke DS, & Batey RT (2022, May 5). The DNA-Binding High-Mobility Group Box Domain of Sox Family Proteins Directly Interacts with RNA In Vitro. *Biochemistry*. 10.1021/acs.biochem.2c00218
- Han H, Braunschweig U, Gonatopoulos-Pournatzis T, Weatheritt RJ, Hirsch CL, Ha KCH, Radovani E, Nabeel-Shah S, Sterne-Weiler T, Wang JL, O'Hanlon D, Pan Q, Ray D, Zheng H, Vizeacoumar F, Datti A, Magomedova L, Cummins CL, Hughes TR, Greenblatt JF, Wrana JL, Moffat J, & Blencowe BJ (2017, Feb 2). Multilayered Control of Alternative Splicing Regulatory Networks by Transcription Factors. *Molecular Cell*, 65(3), 539–+. 10.1016/j.molcel.2017.01.011
- Han Q, Xu L, Lin W, Yao X, Jiang M, Zhou R, Sun X, & Zhao L. (2019, Apr). Long noncoding RNA CRCMSL suppresses tumor invasive and metastasis in colorectal carcinoma through nucleocytoplasmic shuttling of HMGB2. *Oncogene*, 38(16), 3019–3032. 10.1038/s41388-018-0614-4 [PubMed: 30575817]
- Hata K, Nishimura R, Muramatsu S, Matsuda A, Matsubara T, Amano K, Ikeda F, Harley VR, & Yoneda T. (2008, Sep). Paraspeckle protein p54nrb links Sox9-mediated transcription with RNA processing during chondrogenesis in mice. *J Clin Invest*, 118(9), 3098–3108. 10.1172/JCI31373 [PubMed: 18677406]
- He C, Sidoli S, Warneford-Thomson R, Tatomer DC, Wilusz JE, Garcia BA, & Bonasio R. (2016, Oct 20). High-Resolution Mapping of RNA-Binding Regions in the Nuclear Proteome of Embryonic Stem Cells. *Mol Cell*, 64(2), 416–430. 10.1016/j.molcel.2016.09.034 [PubMed: 27768875]
- Hentze MW, Castello A, Schwarzl T, & Preiss T. (2018, May). A brave new world of RNA-binding proteins. *Nat Rev Mol Cell Biol*, 19(5), 327–341. 10.1038/nrm.2017.130 [PubMed: 29339797]
- Hildebrandt A, Alanis-Lobato G, Voigt A, Zarnack K, Andrade-Navarro MA, Beli P, & Konig J. (2017, Nov 29). Interaction profiling of RNA-binding ubiquitin ligases reveals a link between posttranscriptional regulation and the ubiquitin system. *Sci Rep*, 7(1), 16582. 10.1038/s41598-017-16695-6 [PubMed: 29185492]
- Hollmann NM, Jagtap PKA, Masiewicz P, Guitart T, Simon B, Provaznik J, Stein F, Haberkant P, Sweetapple LJ, Villacorta L, Mooijman D, Benes V, Savitski MM, Gebauer F, & Hennig J. (2020, Jul 21). Pseudo-RNA-Binding Domains Mediate RNA Structure Specificity in Upstream of N-Ras. *Cell Rep*, 32(3), 107930. 10.1016/j.celrep.2020.107930
- Holmes ZE, Hamilton DJ, Hwang T, Parsonnet NV, Rinn JL, Wuttke DS, & Batey RT (2020, Apr 14). The Sox2 transcription factor binds RNA. *Nat Commun*, 11(1), 1805. 10.1038/s41467-020-15571-8 [PubMed: 32286318]
- Hou L, Srivastava Y, & Jauch R. (2017, Mar). Molecular basis for the genome engagement by Sox proteins. *Semin Cell Dev Biol*, 63, 2–12. 10.1016/j.semcdb.2016.08.005 [PubMed: 27521520]
- Hou L, Wei Y, Lin Y, Wang X, Lai Y, Yin M, Chen Y, Guo X, Wu S, Zhu Y, Yuan J, Tariq M, Li N, Sun H, Wang H, Zhang X, Chen J, Bao X, & Jauch R. (2020, Apr 17). Concurrent binding to DNA and RNA facilitates the pluripotency reprogramming activity of Sox2. *Nucleic Acids Res*, 48(7), 3869–3887. 10.1093/nar/gkaa067 [PubMed: 32016422]
- Hsu T, King DL, LaBonne C, & Kafatos FC (1993, Jul 15). A Drosophila single-strand DNA/RNA-binding factor contains a high-mobility-group box and is enriched in the nucleolus. *Proc Natl Acad Sci U S A*, 90(14), 6488–6492. 10.1073/pnas.90.14.6488 [PubMed: 7688122]
- Huang R, Han M, Meng L, & Chen X. (2018, Apr 24). Transcriptome-wide discovery of coding and noncoding RNA-binding proteins. *Proc Natl Acad Sci U S A*, 115(17), E3879–E3887. 10.1073/pnas.1718406115 [PubMed: 29636419]
- Hudson WH, & Ortlund EA (2014, Nov). The structure, function and evolution of proteins that bind DNA and RNA. *Nat Rev Mol Cell Biol*, 15(11), 749–760. 10.1038/nrm3884 [PubMed: 25269475]
- Iadevaia V, Wouters MD, Kanitz A, Matia-Gonzalez AM, Laing EE, & Gerber AP (2020, Jan). Tandem RNA isolation reveals functional rearrangement of RNA-binding proteins

on CDKN1B/p27(Kip1) 3'UTRs in cisplatin treated cells. *RNA Biol*, 17(1), 33–46. 10.1080/15476286.2019.1662268 [PubMed: 31522610]

- Iarovaia OV, Minina EP, Sheval EV, Onichtchouk D, Dokudovskaya S, Razin SV, & Vassetzky YS (2019, Aug). Nucleolus: A Central Hub for Nuclear Functions. *Trends Cell Biol*, 29(8), 647–659. 10.1016/j.tcb.2019.04.003 [PubMed: 31176528]
- Ignarski M, Rill C, Kaiser RWJ, Kaldirim M, Neuhaus R, Esmailie R, Li X, Klein C, Bohl K, Petersen M, Frese CK, Hohne M, Atanassov I, Rinschen MM, Hopker K, Schermer B, Benzing T, Dieterich C, Fabretti F, & Muller RU (2019, Apr). The RNA-Protein Interactome of Differentiated Kidney Tubular Epithelial Cells. *J Am Soc Nephrol*, 30(4), 564–576. 10.1681/ASN.2018090914 [PubMed: 30867249]
- Ilik IA, & Aktas T. (2021, Jul 10). Nuclear speckles: dynamic hubs of gene expression regulation. *FEBS J*. 10.1111/febs.16117
- Itzhak DN, Tyanova S, Cox J, & Borner GH (2016, Jun 9). Global, quantitative and dynamic mapping of protein subcellular localization. *Elife*, 5. 10.7554/eLife.16950
- Jain S, Wheeler JR, Walters RW, Agrawal A, Barsic A, & Parker R. (2016, Jan 28). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell*, 164(3), 487–498. 10.1016/j.cell.2015.12.038 [PubMed: 26777405]
- Jaouen S, de Koning L, Gaillard C, Muselikova-Polanska E, Stros M, & Strauss F. (2005, Nov 4). Determinants of specific binding of HMGB1 protein to hemicatenated DNA loops. *J Mol Biol*, 353(4), 822–837. 10.1016/j.jmb.2005.08.073 [PubMed: 16199053]
- Jarvelin AI, Noerenberg M, Davis I, & Castello A. (2016, Apr 6). The new (dis)order in RNA regulation. *Cell Commun Signal*, 14, 9. 10.1186/s12964-016-0132-3 [PubMed: 27048167]
- Jellbauer S, & Jansen RP (2008, Oct-Dec). A putative function of the nucleolus in the assembly or maturation of specialized messenger ribonucleoprotein complexes. *RNA Biol*, 5(4), 225–229. 10.4161/rna.7163 [PubMed: 19182521]
- Jiang Y, Jiang YY, Xie JJ, Mayakonda A, Hazawa M, Chen L, Xiao JF, Li CQ, Huang ML, Ding LW, Sun QY, Xu L, Kanojia D, Jeitany M, Deng JW, Liao LD, Soukiasian HJ, Berman BP, Hao JJ, Xu LY, Li EM, Wang MR, Bi XG, Lin DC, & Koeffler HP (2018, Sep 6). Co-activation of super-enhancer-driven CCAT1 by TP63 and SOX2 promotes squamous cancer progression. *Nat Commun*, 9(1), 3619. 10.1038/s41467-018-06081-9 [PubMed: 30190462]
- Jing R, Guo X, Yang Y, Chen W, Kang J, & Zhu S. (2020, Jul). Long noncoding RNA Q associates with Sox2 and is involved in the maintenance of pluripotency in mouse embryonic stem cells. *Stem Cells*, 38(7), 834–848. 10.1002/stem.3180 [PubMed: 32277787]
- JR P. o., Norman DG, Bramham J, Bianchi ME, & Lilley DM (1998, Feb 2). HMG box proteins bind to four-way DNA junctions in their open conformation. *EMBO J*, 17(3), 817–826. 10.1093/emboj/17.3.817 [PubMed: 9451006]
- Kagawa W, & Kurumizaka H. (2021, Dec). Structural basis for DNA sequence recognition by pioneer factors in nucleosomes. *Curr Opin Struct Biol*, 71, 59–64. 10.1016/j.sbi.2021.05.011 [PubMed: 34218163]
- Katahira J. (2015, Mar 31). Nuclear export of messenger RNA. *Genes (Basel)*, 6(2), 163–184. 10.3390/genes6020163 [PubMed: 25836925]
- Khoury G, Lee MY, Ramarathinam SH, McMahon J, Purcell AW, Sonza S, Lewin SR, & Purcell DFJ (2021). The RNA-Binding Proteins SRP14 and HMGB3 Control HIV-1 Tat mRNA Processing and Translation During HIV-1 Latency. *Front Genet*, 12, 680725. 10.3389/fgene.2021.680725
- Kornblihtt AR (2006, Jan). Chromatin, transcript elongation and alternative splicing. *Nature Structural & Molecular Biology*, 13(1), 5–7. 10.1038/nsmb0106-5
- Kornblihtt AR, De la Mata M, Fededa JP, Munoz MJ, & Nogues G. (2004, Oct). Multiple links between transcription and splicing. *RNA*, 10(10), 1489–1498. 10.1261/rna.7100104 [PubMed: 15383674]
- Kralovicova J, Patel A, Searle M, & Vorechovsky I. (2015). The role of short RNA loops in recognition of a single-hairpin exon derived from a mammalian-wide interspersed repeat. *Rna Biology*, 12(1), 54–69. 10.1080/15476286.2015.1017207 [PubMed: 25826413]

- Kwon SC, Yi H, Eichelbaum K, Fohr S, Fischer B, You KT, Castello A, Krijgsveld J, Hentze MW, & Kim VN (2013, Sep). The RNA-binding protein repertoire of embryonic stem cells. *Nature Structural & Molecular Biology*, 20(9), 1122–1130. 10.1038/nsmb.2638
- Lafontaine DLJ, Riback JA, Bascetin R, & Brangwynne CP (2021, Mar). The nucleolus as a multiphase liquid condensate. *Nat Rev Mol Cell Biol*, 22(3), 165–182. 10.1038/s41580-020-0272-6 [PubMed: 32873929]
- Langst G, & Manelyte L. (2015, Jun 12). Chromatin Remodelers: From Function to Dysfunction. *Genes (Basel)*, 6(2), 299–324. 10.3390/genes6020299 [PubMed: 26075616]
- Lee FCY, & Ule J. (2018, Feb 1). Advances in CLIP Technologies for Studies of Protein-RNA Interactions. *Mol Cell*, 69(3), 354–369. 10.1016/j.molcel.2018.01.005 [PubMed: 29395060]
- Lee S, Song H, Kim HA, Oh B, Lee DY, & Lee M. (2012, Jan). The box A domain of high mobility group box-1 protein as an efficient siRNA carrier with anti-inflammatory effects. *J Cell Biochem*, 113(1), 122–131. 10.1002/jcb.23334 [PubMed: 21898538]
- Lee SK, Park MW, Yang EG, Yu JH, & Jeong SJ (2005, Feb 4). An RNA aptamer that binds to the beta-catenin interaction domain of TCF-1 protein. *Biochemical and Biophysical Research Communications*, 327(1), 294–299. 10.1016/j.bbrc.2004.12.011 [PubMed: 15629461]
- Lee SY, & Jeong S. (2004, Feb 29). In vitro selection and characterization of TCF-1 binding RNA aptamers. *Mol Cells*, 17(1), 174–179. <https://www.ncbi.nlm.nih.gov/pubmed/15055546> [PubMed: 15055546]
- Leulliot N, & Varani G. (2001, Jul 10). Current topics in RNA-protein recognition: control of specificity and biological function through induced fit and conformational capture. *Biochemistry*, 40(27), 7947–7956. 10.1021/bi010680y [PubMed: 11434763]
- Li B, Zhu L, Li L, & Ma R. (2021). lncRNA OXCT1-AS1 Promotes Metastasis in Non-Small-Cell Lung Cancer by Stabilizing LEF1, In Vitro and In Vivo. *Biomed Res Int*, 2021, 4959381. 10.1155/2021/4959381
- Li Q, Sun H, Luo D, Gan L, Mo S, Dai W, Liang L, Yang Y, Xu M, Li J, Zheng P, Li X, Li Y, & Wang Z. (2021, Nov 5). Lnc-RP11–536 K7.3/SOX2/HIF-1alpha signaling axis regulates oxaliplatin resistance in patient-derived colorectal cancer organoids. *J Exp Clin Cancer Res*, 40(1), 348. 10.1186/s13046-021-02143-x [PubMed: 34740372]
- Li X, Song J, & Yi C. (2014, Apr). Genome-wide mapping of cellular protein-RNA interactions enabled by chemical crosslinking. *Genomics Proteomics Bioinformatics*, 12(2), 72–78. 10.1016/j.gpb.2014.03.001 [PubMed: 24747191]
- Liu Z, Legant WR, Chen BC, Li L, Grimm JB, Lavis LD, Betzig E, & Tjian R. (2014, Dec 24). 3D imaging of Sox2 enhancer clusters in embryonic stem cells. *Elife*, 3, e04236. 10.7554/eLife.04236
- Livi CM, Klus P, Delli Ponti R, & Tartaglia GG (2016, Mar 1). catRAPID signature: identification of ribonucleoproteins and RNA-binding regions. *Bioinformatics*, 32(5), 773–775. 10.1093/bioinformatics/btv629 [PubMed: 26520853]
- Lomeli H, & Castillo-Robles J. (2016, Jun). The developmental and pathogenic roles of BAF57, a special subunit of the BAF chromatin-remodeling complex. *Febs Letters*, 590(11), 1555–1569. 10.1002/1873-3468.12201 [PubMed: 27149204]
- Long Y, Bolanos B, Gong L, Liu W, Goodrich KJ, Yang X, Chen S, Gooding AR, Maegley KA, Gajiwala KS, Brooun A, Cech TR, & Liu X. (2017, Nov 29). Conserved RNA-binding specificity of polycomb repressive complex 2 is achieved by dispersed amino acid patches in EZH2. *Elife*, 6. 10.7554/eLife.31558
- Long Y, Wang X, Youmans DT, & Cech TR (2017, Sep). How do lncRNAs regulate transcription? *Sci Adv*, 3(9), eaao2110. 10.1126/sciadv.aao2110
- Lou MM, Tang XQ, Wang GM, He J, Luo F, Guan MF, Wang F, Zou H, Wang JY, Zhang Q, Xu MJ, Shi QL, Shen LB, Ma GM, Wu Y, Zhang YY, Liang AB, Wang TH, Xiong LL, Wang J, Xu J, & Wang WY (2021, Jul 1). Long noncoding RNA BS-DRL1 modulates the DNA damage response and genome stability by interacting with HMGB1 in neurons. *Nat Commun*, 12(1), 4075. 10.1038/s41467-021-24236-z [PubMed: 34210972]
- Lunde BM, Moore C, & Varani G. (2007, Jun). RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol*, 8(6), 479–490. 10.1038/nrm2178 [PubMed: 17473849]

- Lyonnais S, Tarres-Sole A, Rubio-Cosials A, Cuppari A, Brito R, Jaumot J, Gargallo R, Vilaseca M, Silva C, Granzhan A, Teulade-Fichou MP, Eritja R, & Sola M. (2017, Mar 9). The human mitochondrial transcription factor A is a versatile G-quadruplex binding protein. *Sci Rep*, 7, 43992. 10.1038/srep43992 [PubMed: 28276514]
- Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, Madhusoodanan N, Kolesnikov A, & Lopez R. (2022, Apr 12). Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res*. 10.1093/nar/gkac240
- Malarkey CS, & Churchill ME (2012, Dec). The high mobility group box: the ultimate utility player of a cell. *Trends Biochem Sci*, 37(12), 553–562. 10.1016/j.tibs.2012.09.003 [PubMed: 23153957]
- Mallam AL, Sae-Lee W, Schaub JM, Tu F, Battenhouse A, Jang YJ, Kim J, Wallingford JB, Finkelstein IJ, Marcotte EM, & Drew K. (2019, Oct 29). Systematic Discovery of Endogenous Human Ribonucleoprotein Complexes. *Cell Rep*, 29(5), 1351–1368 e1355. 10.1016/j.celrep.2019.09.060 [PubMed: 31665645]
- Markmiller S, Soltanieh S, Server KL, Mak R, Jin W, Fang MY, Luo EC, Krach F, Yang D, Sen A, Fulzele A, Wozniak JM, Gonzalez DJ, Kankel MW, Gao FB, Bennett EJ, Lecuyer E, & Yeo GW (2018, Jan 25). Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. *Cell*, 172(3), 590–604 e513. 10.1016/j.cell.2017.12.032 [PubMed: 29373831]
- Masse JE, Wong B, Yen YM, Allain FH, Johnson RC, & Feigon J. (2002, Oct 18). The *S. cerevisiae* architectural HMGB protein NHP6A complexed with DNA: DNA and protein conformational changes upon binding. *J Mol Biol*, 323(2), 263–284. 10.1016/S0022-2836(02)00938-5 [PubMed: 12381320]
- McCaughey MJ, Huo R, Becker N, Holte MN, Muthurajan UM, Rouzina I, Luger K, Maher LJ 3rd, Israeloff NE, & Williams MC (2019, Jan 25). Single and double box HMGB proteins differentially destabilize nucleosomes. *Nucleic Acids Res*, 47(2), 666–678. 10.1093/nar/gky1119 [PubMed: 30445475]
- Mertin S, McDowall SG, & Harley VR (1999, Mar 1). The DNA-binding specificity of SOX9 and other SOX proteins. *Nucleic Acids Res*, 27(5), 1359–1364. 10.1093/nar/27.5.1359 [PubMed: 9973626]
- Mi H, Muruganujan A, Casagrande JT, & Thomas PD (2013, Aug). Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc*, 8(8), 1551–1566. 10.1038/nprot.2013.092 [PubMed: 23868073]
- Mi H, Muruganujan A, Huang X, Ebert D, Mills C, Guo X, & Thomas PD (2019, Mar). Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat Protoc*, 14(3), 703–721. 10.1038/s41596-019-0128-8 [PubMed: 30804569]
- Moosa MM, Tsoi PS, Choi KJ, Ferreón ACM, & Ferreón JC (2018, Dec 4). Direct Single-Molecule Observation of Sequential DNA Bending Transitions by the Sox2 HMG Box. *Int J Mol Sci*, 19(12). 10.3390/ijms19123865
- Moreno CS (2020, Dec). SOX4: The unappreciated oncogene. *Seminars in Cancer Biology*, 67, 57–64. 10.1016/j.semcancer.2019.08.027 [PubMed: 31445218]
- Mourtada-Maarabouni M, Pickard MR, Hedge VL, Farzaneh F, & Williams GT (2009, Jan 15). GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene*, 28(2), 195–208. 10.1038/onc.2008.373 [PubMed: 18836484]
- Muller S, Ronfani L, & Bianchi ME (2004, Mar). Regulated expression and subcellular localization of HMGB1, a chromatin protein with a cytokine function. *J Intern Med*, 255(3), 332–343. 10.1111/j.1365-2796.2003.01296.x [PubMed: 14871457]
- Murugesapillai D, McCaughey MJ, Maher LJ 3rd, & Williams MC (2017, Feb). Single-molecule studies of high-mobility group B architectural DNA bending proteins. *Biophys Rev*, 9(1), 17–40. 10.1007/s12551-016-0236-4 [PubMed: 28303166]
- Ner SS (1992, Apr). HMGs everywhere. *Curr Biol*, 2(4), 208–210. 10.1016/0960-9822(92)90541-h [PubMed: 15335978]
- Neugebauer KM (2019, Aug). Nascent RNA and the Coordination of Splicing with Transcription. *Cold Spring Harbor Perspectives in Biology*, 11(8). 10.1101/cshperspect.a032227

- Ng SY, Bogu GK, Soh BS, & Stanton LW (2013, Aug 8). The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis. *Mol Cell*, 51(3), 349–359. 10.1016/j.molcel.2013.07.017 [PubMed: 23932716]
- Ng SY, Johnson R, & Stanton LW (2012, Feb 1). Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. *EMBO J*, 31(3), 522–533. 10.1038/emboj.2011.459 [PubMed: 22193719]
- Ngo HB, Kaiser JT, & Chan DC (2011, Oct 30). The mitochondrial transcription and packaging factor Tfam imposes a U-turn on mitochondrial DNA. *Nature Structural & Molecular Biology*, 18(11), 1290–1296. 10.1038/nsmb.2159
- Niu LR, Yang WL, Duan LL, Wang XQ, Li YD, Xu CC, Liu C, Zhang YJ, Zhou W, Liu JQ, Zhao QC, Han Y, Hong L, & Fan DM (2020, Nov). Biological functions and theranostic potential of HMGB family members in human cancers. *Therapeutic Advances in Medical Oncology*, 12. 10.1177/1758835920970850
- Noh JH, Kim KM, McClusky WG, Abdelmohsen K, & Gorospe M. (2018, May). Cytoplasmic functions of long noncoding RNAs. *Wiley Interdiscip Rev RNA*, 9(3), e1471. 10.1002/wrna.1471 [PubMed: 29516680]
- Novak D, Huser L, Elton JJ, Umansky V, Altevogt P, & Utikal J. (2020, Dec). SOX2 in development and cancer biology. *Seminars in Cancer Biology*, 67, 74–82. 10.1016/j.semcancer.2019.08.007 [PubMed: 31412296]
- Oh B, & Lee M. (2014, Feb 10). Combined delivery of HMGB-1 box A peptide and S1PLYase siRNA in animal models of acute lung injury. *J Control Release*, 175, 25–35. 10.1016/j.jconrel.2013.12.008 [PubMed: 24361371]
- Ohe K, Lalli E, & Sassone-Corsi P. (2002, Feb 5). A direct role of SRY and SOX proteins in pre-mRNA splicing. *Proc Natl Acad Sci U S A*, 99(3), 1146–1151. 10.1073/pnas.022645899 [PubMed: 11818535]
- Ohe K, Tamai KT, Parvinen M, & Sassone-Corsi P. (2009, Jun). DAX-1 and SOX6 molecular interplay results in an antagonistic effect in pre-mRNA splicing. *Dev Dyn*, 238(6), 1595–1604. 10.1002/dvdy.21957 [PubMed: 19384854]
- Ouyang J, Yu W, Liu J, Zhang N, Florens L, Chen J, Liu H, Washburn M, Pei D, & Xie T. (2015, Sep 11). Cyclin-dependent kinase-mediated Sox2 phosphorylation enhances the ability of Sox2 to establish the pluripotent state. *J Biol Chem*, 290(37), 22782–22794. 10.1074/jbc.M115.658195 [PubMed: 26139602]
- Palacio M, & Taatjes DJ (2022, Jan 15). Merging Established Mechanisms with New Insights: Condensates, Hubs, and the Regulation of RNA Polymerase II Transcription. *J Mol Biol*, 434(1), 167216. 10.1016/j.jmb.2021.167216
- Palasingam P, Jauch R, Ng CKL, & Kolatkar PR (2009, May 8). The Structure of Sox17 Bound to DNA Reveals a Conserved Bending Topology but Selective Protein Interaction Platforms. *Journal of Molecular Biology*, 388(3), 619–630. 10.1016/j.jmb.2009.03.055 [PubMed: 19328208]
- Pandit S, Wang D, & Fu XD (2008, Jun). Functional integration of transcriptional and RNA processing machineries. *Curr Opin Cell Biol*, 20(3), 260–265. 10.1016/j.ceb.2008.03.001 [PubMed: 18436438]
- Papantonis A. (2021, Nov). HMGs as rheostats of chromosomal structure and cell proliferation. *Trends Genet*, 37(11), 986–994. 10.1016/j.tig.2021.07.004 [PubMed: 34311989]
- Park MW, Choi KH, & Jeong S. (2005, Apr 29). Inhibition of the DNA binding by the TCF-1 binding RNA aptamer. *Biochem Biophys Res Commun*, 330(1), 11–17. 10.1016/j.bbrc.2005.02.119 [PubMed: 15781225]
- Pashev IG, Dimitrov SI, & Angelov D. (1991, Sep). Crosslinking proteins to nucleic acids by ultraviolet laser irradiation. *Trends Biochem Sci*, 16(9), 323–326. 10.1016/0968-0004(91)90133-g [PubMed: 1835191]
- Payet D, & Travers A. (1997, Feb 14). The acidic tail of the high mobility group protein HMG-D modulates the structural selectivity of DNA binding. *J Mol Biol*, 266(1), 66–75. 10.1006/jmbi.1996.0782 [PubMed: 9054971]

- Pei YF, Zhou B, & Liu XQ (2021, Jan). The long non-coding RNA rhabdomyosarcoma 2-associated transcript exerts anti-tumor effects on lung adenocarcinoma via ubiquitination of SOX9. *Annals of Translational Medicine*, 10(1). 10.21037/atm-21-6052
- Penrad-Mobayed M, Perrin C, L'Hote D, Contremoulins V, Lepesant JA, Boizet-Bonhoure B, Poulat F, Baudin X, & Veitia RA (2018, May 8). A role for SOX9 in post-transcriptional processes: insights from the amphibian oocyte. *Sci Rep*, 8(1), 7191. 10.1038/s41598-018-25356-1 [PubMed: 29740094]
- Perez-Perri JI, Rogell B, Schwarzl T, Stein F, Zhou Y, Rettel M, Brosig A, & Hentze MW (2018, Oct 23). Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. *Nature Communications*, 9. 10.1038/s41467-018-06557-8
- Phochanukul N, & Russell S. (2010, Mar). No backbone but lots of Sox: Invertebrate Sox genes. *Int J Biochem Cell Biol*, 42(3), 453–464. 10.1016/j.biocel.2009.06.013 [PubMed: 19589395]
- Pisani G, & Baron B. (2019, Dec). Nuclear paraspeckles function in mediating gene regulatory and apoptotic pathways. *Noncoding RNA Res*, 4(4), 128–134. 10.1016/j.ncrna.2019.11.002 [PubMed: 32072080]
- Pouremamali F, Vahedian V, Hassani N, Mirzaei S, Pouremamali A, Kazemzadeh H, Faridvand Y, Jafari-Gharabaghloou D, Nouri M, & Maroufi NF (2022, Mar). The role of SOX family in cancer stem cell maintenance: With a focus on SOX2. *Pathol Res Pract*, 231, 153783. 10.1016/j.prp.2022.153783
- Queiroz RML, Smith T, Villanueva E, Marti-Solano M, Monti M, Pizzinga M, Mirea DM, Ramakrishna M, Harvey RF, Dezi V, Thomas GH, Willis AE, & Lilley KS (2019, Feb). Comprehensive identification of RNA-protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nature Biotechnology*, 37(2), 169–+. 10.1038/s41587-018-0001-2
- Quinodoz SA, Jachowicz JW, Bhat P, Ollikainen N, Banerjee AK, Goronzy IN, Blanco MR, Chovanec P, Chow A, Markaki Y, Thai J, Plath K, & Guttman M. (2021, Nov 11). RNA promotes the formation of spatial compartments in the nucleus. *Cell*, 184(23), 5775–5790 e5730. 10.1016/j.cell.2021.10.014 [PubMed: 34739832]
- Rahmoun M, Lavery R, Laurent-Chaballier S, Bellora N, Philip GK, Rossitto M, Symon A, Pailhoux E, Cammas F, Chung J, Bagheri-Fam S, Murphy M, Bardwell V, Zarkower D, Boizet-Bonhoure B, Clair P, Harley VR, & Poulat F. (2017, Jul 7). In mammalian foetal testes, SOX9 regulates expression of its target genes by binding to genomic regions with conserved signatures. *Nucleic Acids Res*, 45(12), 7191–7211. 10.1093/nar/gkx328 [PubMed: 28472341]
- Ramanathan M, Porter DF, & Khavari PA (2019, Mar). Methods to study RNA-protein interactions. *Nat Methods*, 16(3), 225–234. 10.1038/s41592-019-0330-1 [PubMed: 30804549]
- Ramanouskaya TV, & Grinev VV (2017, Dec). The determinants of alternative RNA splicing in human cells. *Molecular Genetics and Genomics*, 292(6), 1175–1195. 10.1007/s00438-017-1350-0 [PubMed: 28707092]
- Rambout X, & Maquat LE (2020, Sep 1). The nuclear cap-binding complex as choreographer of gene transcription and pre-mRNA processing. *Genes & Development*, 34(17–18), 1113–1127. 10.1101/gad.339986.120 [PubMed: 32873578]
- Razin SV, & Gavrillov AA (2021, Jul). Non-coding RNAs in chromatin folding and nuclear organization. *Cellular and Molecular Life Sciences*, 78(14), 5489–5504. 10.1007/s00018-021-03876-w [PubMed: 34117518]
- Reeves R. (2015, Dec). High mobility group (HMG) proteins: Modulators of chromatin structure and DNA repair in mammalian cells. *DNA Repair (Amst)*, 36, 122–136. 10.1016/j.dnarep.2015.09.015 [PubMed: 26411874]
- Remenyi A, Lins K, Nissen LJ, Reinbold R, Scholer HR, & Wilmanns M. (2003, Aug 15). Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev*, 17(16), 2048–2059. 10.1101/gad.269303 [PubMed: 12923055]
- Richard SA, Jiang Y, Xiang LH, Zhou S, Wang J, Su Z, & Xu H. (2017). Post-translational modifications of high mobility group box 1 and cancer. *Am J Transl Res*, 9(12), 5181–5196. <https://www.ncbi.nlm.nih.gov/pubmed/29312476> [PubMed: 29312476]

- Richter WF, Nayak S, Iwasa J, & Taatjes DJ (2022, Jun 20). The Mediator complex as a master regulator of transcription by RNA polymerase II. *Nat Rev Mol Cell Biol*. 10.1038/s41580-022-00498-3
- Rizzino A, & Wuebben EL (2016, Jun). Sox2/Oct4: A delicately balanced partnership in pluripotent stem cells and embryogenesis. *Biochim Biophys Acta*, 1859(6), 780–791. 10.1016/j.bbagr.2016.03.006 [PubMed: 26992828]
- Saha K, England W, Fernandez MM, Biswas T, Spitale RC, & Ghosh G. (2020, Jun 19). Structural disruption of exonic stem-loops immediately upstream of the intron regulates mammalian splicing. *Nucleic Acids Research*, 48(11), 6294–6309. 10.1093/nar/gkaa358 [PubMed: 32402057]
- Saitoh N, Spahr CS, Patterson SD, Bubulya P, Neuwald AF, & Spector DL (2004, Aug). Proteomic analysis of interchromatin granule clusters. *Mol Biol Cell*, 15(8), 3876–3890. 10.1091/mbc.e04-03-0253 [PubMed: 15169873]
- Samudyata Amaral P. P., Engstrom PG, Robson SC, Nielsen ML, Kouzarides T, & Castelo-Branco G. (2019, Aug 1). Interaction of Sox2 with RNA binding proteins in mouse embryonic stem cells. *Exp Cell Res*, 381(1), 129–138. 10.1016/j.yexcr.2019.05.006 [PubMed: 31077711]
- Sato K, & Siomi MC (2015, Jun 22). Functional and structural insights into the piRNA factor Maelstrom. *FEBS Lett*, 589(14), 1688–1693. 10.1016/j.febslet.2015.03.023 [PubMed: 25836734]
- Sato Y, Yano S, Ewis AA, & Nakahori Y. (2011, May). SRY interacts with ribosomal proteins S7 and L13a in nuclear speckles. *Cell Biol Int*, 35(5), 449–452. 10.1042/CBI20090201 [PubMed: 21114473]
- Scaffidi P, & Bianchi ME (2001, Dec 14). Spatially precise DNA bending is an essential activity of the Sox2 transcription factor. *Journal of Biological Chemistry*, 276(50), 47296–47302. 10.1074/jbc.M107619200 [PubMed: 11584012]
- Scharfen L, & Neugebauer KM (2021, Jul 9). Transcription Regulation Through Nascent RNA Folding. *J Mol Biol*, 433(14), 166975. 10.1016/j.jmb.2021.166975
- Schepers GE, Teasdale RD, & Koopman P. (2002, Aug). Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. *Dev Cell*, 3(2), 167–170. 10.1016/s1534-5807(02)00223-x [PubMed: 12194848]
- Scherer M, Levin M, Butter F, & Scheibe M. (2020, Feb 10). Quantitative Proteomics to Identify Nuclear RNA-Binding Proteins of Malat1. *Int J Mol Sci*, 21(3). 10.3390/ijms21031166
- Scott MS, Calafell SJ, Thomas DY, & Hallett MT (2005, Nov). Refining protein subcellular localization. *PLoS Comput Biol*, 1(6), e66. 10.1371/journal.pcbi.0010066 [PubMed: 16322766]
- See YX, Wang BZ, & Fullwood MJ (2019, Feb). Chromatin Interactions and Regulatory Elements in Cancer: From Bench to Bedside. *Trends Genet*, 35(2), 145–158. 10.1016/j.tig.2018.11.007 [PubMed: 30577989]
- Serikawa T, Spanos C, von Hacht A, Budisa N, Rappsilber J, & Kurreck J. (2018, Jan). Comprehensive identification of proteins binding to RNA G-quadruplex motifs in the 5' UTR of tumor-associated mRNAs. *Biochimie*, 144, 169–184. 10.1016/j.biochi.2017.11.003 [PubMed: 29129743]
- Sharp PA, Chakraborty AK, Henninger JE, & Young RA (2022, Jan). RNA in formation and regulation of transcriptional condensates. *RNA*, 28(1), 52–57. 10.1261/rna.078997.121 [PubMed: 34772787]
- Sigova AA, Abraham BJ, Ji X, Molinie B, Hannett NM, Guo YE, Jangi M, Giallourakis CC, Sharp PA, & Young RA (2015, Nov 20). Transcription factor trapping by RNA in gene regulatory elements. *Science*, 350(6263), 978–981. 10.1126/science.aad3346 [PubMed: 26516199]
- Skalska L, Begley V, Beltran M, Lukauskas S, Khandelwal G, Faull P, Bhamra A, Tavares M, Wellman R, Tvardovskiy A, Foster BM, Ruiz de Los Mozos I, Herrero J, Surinova S, Snijders AP, Bartke T, & Jenner RG (2021, Jul 15). Nascent RNA antagonizes the interaction of a set of regulatory proteins with chromatin. *Mol Cell*, 81(14), 2944–2959 e2910. 10.1016/j.molcel.2021.05.026 [PubMed: 34166609]
- Skalska L, Beltran-Nebot M, Ule J, & Jenner RG (2017, May). Regulatory feedback from nascent RNA to chromatin and transcription. *Nat Rev Mol Cell Biol*, 18(5), 331–337. 10.1038/nrm.2017.12 [PubMed: 28270684]

- Sofiadis K, Josipovic N, Nikolic M, Kargapolova Y, Ubelmesser N, Varamogianni-Mamatsi V, Zirkel A, Papadionysiou I, Loughran G, Keane J, Michel A, Gusmao EG, Becker C, Altmuller J, Georgomanolis T, Mizi A, & Papantonis A. (2021, Jun). HMGB1 coordinates SASP-related chromatin folding and RNA homeostasis on the path to senescence. *Mol Syst Biol*, 17(6), e9760. 10.15252/msb.20209760 [PubMed: 34166567]
- Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, & Zaret KS (2015, Apr 23). Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell*, 161(3), 555–568. 10.1016/j.cell.2015.03.017 [PubMed: 25892221]
- Spector DL, & Lamond AI (2011, Feb 1). Nuclear speckles. *Cold Spring Harb Perspect Biol*, 3(2). 10.1101/cshperspect.a000646
- Stamou M, Ng SY, Brand H, Wang H, Plummer L, Best L, Havlicek S, Hibberd M, Khor CC, Gusella J, Balasubramanian R, Talkowski M, Stanton LW, & Crowley WF (2020, Mar 1). A Balanced Translocation in Kallmann Syndrome Implicates a Long Noncoding RNA, RMST, as a GnRH Neuronal Regulator. *J Clin Endocrinol Metab*, 105(3). 10.1210/clinem/dgz011
- Stevanovic M, Kovacevic-Grujicic N, Mojsin M, Milivojevic M, & Drakulic D. (2021, Oct 26). SOX transcription factors and glioma stem cells: Choosing between stemness and differentiation. *World Journal of Stem Cells*, 13(10), 1417–1445. 10.4252/wjsc.v13.i10.1417 [PubMed: 34786152]
- Stros M. (2010, Jan-Feb). HMGB proteins: interactions with DNA and chromatin. *Biochim Biophys Acta*, 1799(1–2), 101–113. 10.1016/j.bbagr.2009.09.008 [PubMed: 20123072]
- Stros M, Launholt D, & Grasser KD (2007, Oct). The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins. *Cell Mol Life Sci*, 64(19–20), 2590–2606. 10.1007/s00018-007-7162-3 [PubMed: 17599239]
- Sysoev VO, Fischer B, Frese CK, Gupta I, Krijgsvelde J, Hentze MW, Castello A, & Ephrussi A. (2016, Jul 5). Global changes of the RNA-bound proteome during the maternal-to-zygotic transition in *Drosophila*. *Nat Commun*, 7, 12128. 10.1038/ncomms12128 [PubMed: 27378189]
- Tang SC, Vijayakumar U, Zhang Y, & Fullwood MJ (2022, Jun 10). Super-Enhancers, Phase-Separated Condensates, and 3D Genome Organization in Cancer. *Cancers (Basel)*, 14(12). 10.3390/cancers14122866
- Tang ZC, Zhao JX, Pearson ZJ, Boskovic ZV, & Wang JX (2021, Apr). RNA-Targeting Splicing Modifiers: Drug Development and Screening Assays. *Molecules*, 26(8). 10.3390/molecules26082263
- Taniguchi N, Kawakami Y, Maruyama I, & Lotz M. (2018, Jan). HMGB proteins and arthritis. *Human Cell*, 31(1), 1–9. 10.1007/s13577-017-0182-x [PubMed: 28916968]
- Tellier M, Maudlin I, & Murphy S. (2020, Sep). Transcription and splicing: A two-way street. *Wiley Interdisciplinary Reviews-Rna*, 11(5). 10.1002/wrna.1593
- Thul PJ, Akesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, Alm T, Asplund A, Bjork L, Breckels LM, Backstrom A, Danielsson F, Fagerberg L, Fall J, Gatto L, Gnann C, Hober S, Hjelmare M, Johansson F, Lee S, Lindskog C, Mulder J, Mulvey CM, Nilsson P, Oksvold P, Rockberg J, Schutten R, Schwenk JM, Sivertsson A, Sjostedt E, Skogs M, Stadler C, Sullivan DP, Tegel H, Winsnes C, Zhang C, Zwahlen M, Mardinoglu A, Ponten F, von Feilitzen K, Lilley KS, Uhlen M, & Lundberg E. (2017, May 26). A subcellular map of the human proteome. *Science*, 356(6340). 10.1126/science.aal3321
- Trendel J, Schwarzl T, Horos R, Prakash A, Bateman A, Hentze MW, & Krijgsvelde J. (2019, Jan 10). The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. *Cell*, 176(1–2), 391–403 e319. 10.1016/j.cell.2018.11.004 [PubMed: 30528433]
- Tung CL, Hou PH, Kao YL, Huang YW, Shen CC, Cheng YH, Wu SF, Lee MS, & Li C. (2010, Mar 12). SOX2 modulates alternative splicing in transitional cell carcinoma. *Biochem Biophys Res Commun*, 393(3), 420–425. 10.1016/j.bbrc.2010.02.010 [PubMed: 20138825]
- Ueshima S, Nagata K, & Okuwaki M. (2017, Nov 15). Internal Associations of the Acidic Region of Upstream Binding Factor Control Its Nucleolar Localization. *Mol Cell Biol*, 37(22). 10.1128/MCB.00218-17
- Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szgyarto CA,

Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, von Heijne G, Nielsen J, & Ponten F. (2015, Jan 23). Proteomics. Tissue-based map of the human proteome. *Science*, 347(6220), 1260419. 10.1126/science.1260419

- UniProt C. (2021, Jan 8). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res*, 49(D1), D480–D489. 10.1093/nar/gkaa1100 [PubMed: 33237286]
- Urdaneta EC, Vieira-Vieira CH, Hick T, Wessels HH, Figini D, Moschall R, Medenbach J, Ohler U, Granneman S, Selbach M, & Beckmann BM (2019, Mar 1). Purification of cross-linked RNA-protein complexes by phenol-toluol extraction. *Nat Commun*, 10(1), 990. 10.1038/s41467-019-08942-3 [PubMed: 30824702]
- Vanzan L, Soldati H, Ythier V, Anand S, Braun SMG, Francis N, & Murr R. (2021, Jun 7). High throughput screening identifies SOX2 as a super pioneer factor that inhibits DNA methylation maintenance at its binding sites. *Nature Communications*, 12(1). 10.1038/s41467-021-23630-x
- Voong CK, Goodrich JA, & Kugel JF (2021, Oct 2). Interactions of HMGB Proteins with the Genome and the Impact on Disease. *Biomolecules*, 11(10). 10.3390/biom11101451
- Vozarikova V, Kunova N, Bauer JA, Frankovsky J, Kotrasova V, Prochazkova K, Dzugasova V, Kutejova E, Pevala V, Nosek J, & Tomaska L. (2020, Aug). Mitochondrial HMG-Box Containing Proteins: From Biochemical Properties to the Roles in Human Diseases. *Biomolecules*, 10(8). 10.3390/biom10081193
- Wang IX, Grunseich C, Fox J, Burdick J, Zhu ZW, Ravazian N, Hafner M, & Cheung VG (2018, Sep). Human proteins that interact with RNA/DNA hybrids. *Genome Research*, 28(9), 1405–1414. 10.1101/gr.237362.118 [PubMed: 30108179]
- Wang JF, Bashir M, Engelsberg BN, Witmer C, Rozmiarek H, & Billings PC (1997, Feb). High mobility group proteins 1 and 2 recognize chromium-damaged DNA. *Carcinogenesis*, 18(2), 371–375. 10.1093/carcin/18.2.371 [PubMed: 9054631]
- Wang Y, & Chen LL (2020, Dec 7). Organization and function of paraspeckles. *Essays Biochem*, 64(6), 875–882. 10.1042/EBC20200010 [PubMed: 32830222]
- Webb M, Payet D, Lee KB, Travers AA, & Thomas JO (2001, May 25). Structural requirements for cooperative binding of HMG1 to DNA minicircles. *J Mol Biol*, 309(1), 79–88. 10.1006/jmbi.2001.4667 [PubMed: 11491303]
- Webb M, & Thomas JO (1999, Nov 26). Structure-specific binding of the two tandem HMG boxes of HMG1 to four-way junction DNA is mediated by the A domain. *J Mol Biol*, 294(2), 373–387. 10.1006/jmbi.1999.3150 [PubMed: 10610765]
- Weina K, & Utikal J. (2014). SOX2 and cancer: current research and its implications in the clinic. *Clinical and Translational Medicine*, 3. 10.1186/2001-1326-3-19
- Weiss MA (2001, Mar). Floppy SOX: mutual induced fit in hmg (high-mobility group) box-DNA recognition. *Mol Endocrinol*, 15(3), 353–362. 10.1210/mend.15.3.0617 [PubMed: 11222737]
- Wheeler EC, Van Nostrand EL, & Yeo GW (2018, Jan). Advances and challenges in the detection of transcriptome-wide protein-RNA interactions. *Wiley Interdiscip Rev RNA*, 9(1). 10.1002/wrna.1436
- Williams CAC, Soufi A, & Pollard SM (2020, Dec). Post-translational modification of SOX family proteins: Key biochemical targets in cancer? *Semin Cancer Biol*, 67(Pt 1), 30–38. 10.1016/j.semcancer.2019.09.009 [PubMed: 31539559]
- Williams DC Jr., Cai M, & Clore GM (2004, Jan 9). Molecular basis for synergistic transcriptional activation by Oct1 and Sox2 revealed from the solution structure of the 42-kDa Oct1.Sox2.Hoxb1-DNA ternary transcription factor complex. *J Biol Chem*, 279(2), 1449–1457. 10.1074/jbc.M309790200 [PubMed: 14559893]
- Yamada M, Ohnishi J, Ohkawara B, Iemura S, Satoh K, Hyodo-Miura J, Kawachi K, Natsume T, & Shibuya H. (2006, Jul 28). NARF, an nemo-like kinase (NLK)-associated ring finger protein regulates the ubiquitylation and degradation of T cell factor/lymphoid enhancer factor (TCF/LEF). *J Biol Chem*, 281(30), 20749–20760. 10.1074/jbc.M602089200 [PubMed: 16714285]
- Yamanaka Y, Faghihi MA, Magistri M, Alvarez-Garcia O, Lotz M, & Wahlestedt C. (2015, May 12). Antisense RNA controls LRP1 Sense transcript expression through interaction with a chromatin-

- associated protein, HMGB2. *Cell Rep*, 11(6), 967–976. 10.1016/j.celrep.2015.04.011 [PubMed: 25937287]
- Yan S, Zhao D, Wang C, Wang H, Guan X, Gao Y, Zhang X, Zhang N, & Chen R. (2021, Jul 11). Characterization of RNA-binding proteins in the cell nucleus and cytoplasm. *Anal Chim Acta*, 1168, 338609. 10.1016/j.aca.2021.338609
- Yanai H, Ban T, Wang Z, Choi MK, Kawamura T, Negishi H, Nakasato M, Lu Y, Hangai S, Koshiba R, Savitsky D, Ronfani L, Akira S, Bianchi ME, Honda K, Tamura T, Kodama T, & Taniguchi T. (2009, Nov 5). HMGB proteins function as universal sentinels for nucleic-acid-mediated innate immune responses. *Nature*, 462(7269), 99–103. 10.1038/nature08512 [PubMed: 19890330]
- Yanai H, Chiba S, Ban T, Nakaima Y, Onoe T, Honda K, Ohdan H, & Taniguchi T. (2011, Jul 12). Suppression of immune responses by nonimmunogenic oligodeoxynucleotides with high affinity for high-mobility group box proteins (HMGBs). *Proc Natl Acad Sci U S A*, 108(28), 11542–11547. 10.1073/pnas.1108535108 [PubMed: 21709231]
- Yoshioka K, Saito K, Tanabe T, Yamamoto A, Ando Y, Nakamura Y, Shirakawa H, & Yoshida M. (1999, Jan 12). Differences in DNA recognition and conformational change activity between boxes A and B in HMG2 protein. *Biochemistry*, 38(2), 589–595. 10.1021/bi981834i [PubMed: 9888798]
- Youn JY, Dyakov BJA, Zhang J, Knight JDR, Vernon RM, Forman-Kay JD, & Gingras AC (2019, Oct 17). Properties of Stress Granule and P-Body Proteomes. *Mol Cell*, 76(2), 286–294. 10.1016/j.molcel.2019.09.014 [PubMed: 31626750]
- Yu R, Yang D, Lei S, Wang X, Meng X, Xue B, & Zhu H. (2015, Dec 9). HMGB1 Promotes Hepatitis C Virus Replication by Interaction with Stem-Loop 4 in the Viral 5' Untranslated Region. *J Virol*, 90(5), 2332–2344. 10.1128/JVI.02795-15 [PubMed: 26656705]
- Yuan LQ, Xiao YZ, Zhou QZ, Yuan DM, Wu BP, Chen GN, & Zhou JL (2014, Jan). Proteomic analysis reveals that MAEL, a component of nuage, interacts with stress granule proteins in cancer cells. *Oncology Reports*, 31(1), 342–350. 10.3892/or.2013.2836 [PubMed: 24189637]
- Zaret KS, & Carroll JS (2011, Nov 1). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev*, 25(21), 2227–2241. 10.1101/gad.176826.111 [PubMed: 22056668]
- Zhang C, Xie L, Fu Y, Yang J, & Cui Y. (2020, Nov). lncRNA MIAT promotes esophageal squamous cell carcinoma progression by regulating miR-1301–3p/INCENP axis and interacting with SOX2. *J Cell Physiol*, 235(11), 7933–7944. 10.1002/jcp.29448 [PubMed: 31943174]
- Zhang H, Xu R, Li B, Xin Z, Ling Z, Zhu W, Li X, Zhang P, Fu Y, Chen J, Liu L, Cheng J, & Jiang H. (2021, Sep 8). LncRNA NEAT1 controls the lineage fates of BMSCs during skeletal aging by impairing mitochondrial function and pluripotency maintenance. *Cell Death Differ*. 10.1038/s41418-021-00858-0
- Zhang SZ, & Sun Y. (2020, Feb). Targeting oncogenic SOX2 in human cancer cells: therapeutic application. *Protein & Cell*, 11(2), 82–84. 10.1007/s13238-019-00673-x [PubMed: 31748974]
- Zhang Y, Gu J, & Sun Q. (2021, Aug 30). Aberrant Stress Granule Dynamics and Aggregopathy in ALS Pathogenesis. *Cells*, 10(9). 10.3390/cells10092247
- Zheng X, Wang X, He Y, & Ge H. (2022). Systematic analysis of expression profiles of HMGB family members for prognostic application in non-small cell lung cancer. *Frontiers in Molecular Biosciences*, 9, 844618. 10.3389/fmolb.2022.844618
- Zhou LL, Zhi ZK, Chen PF, Du CX, Wang BY, Fang X, Tang WB, & Li HX (2022, Mar 7). LncRNA-RMST Functions as a Transcriptional Co-regulator of SOX2 to Regulate miR-1251 in the Progression of Hirschsprung's Disease. *Frontiers in Pediatrics*, 10. 10.3389/fped.2022.749107
- Zhuang W, Ge X, Yang S, Huang M, Zhuang W, Chen P, Zhang X, Fu J, Qu J, & Li B. (2015, Jun). Upregulation of lncRNA MEG3 Promotes Osteogenic Differentiation of Mesenchymal Stem Cells From Multiple Myeloma Patients By Targeting BMP4 Transcription. *Stem Cells*, 33(6), 1985–1997. 10.1002/stem.1989 [PubMed: 25753650]
- Zraly CB, & Dingwall AK (2012, Jul). The chromatin remodeling and mRNA splicing functions of the Brahma (SWI/SNF) complex are mediated by the SNR1/SNF5 regulatory subunit. *Nucleic Acids Research*, 40(13), 5975–5987. 10.1093/nar/gks288 [PubMed: 22467207]

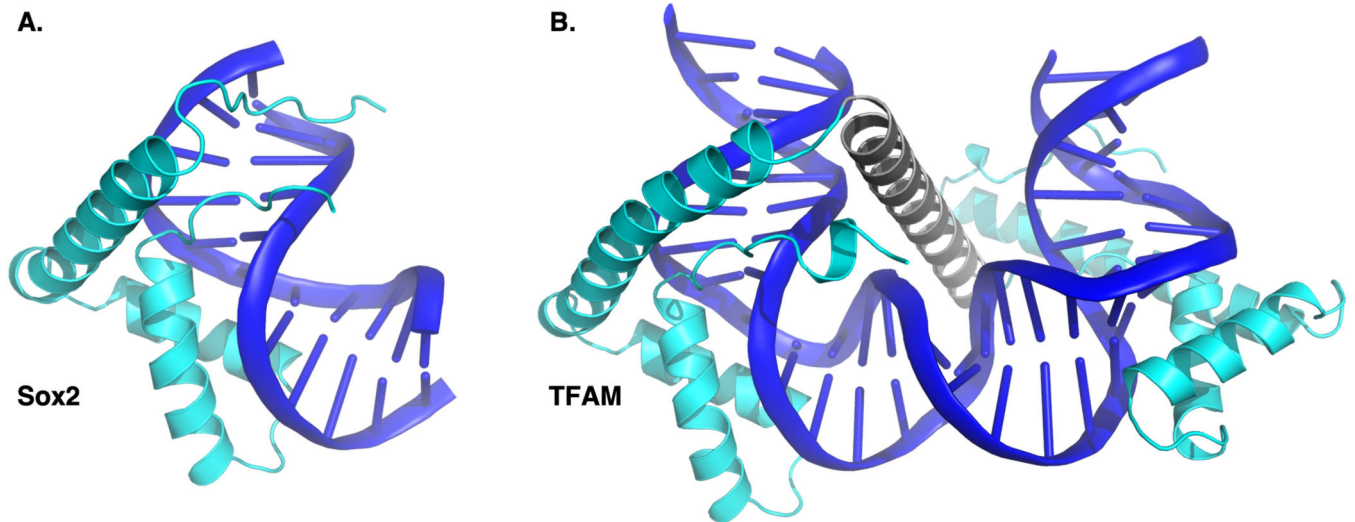


Figure 1. Structures of HMGB domains (cyan) bound to dsDNA (blue). **A.** Structure of Sox2 HMGB domain bound to consensus dsDNA, PDB: 1O4X (Williams et al., 2004). **B.** Structure of TFAM, a dual HMGB domain protein bound to dsDNA, PDB: 3TMM (Ngo et al., 2011). TFAM's grey domain is not part of the HMGB domains. These images were rendered with PyMOL.

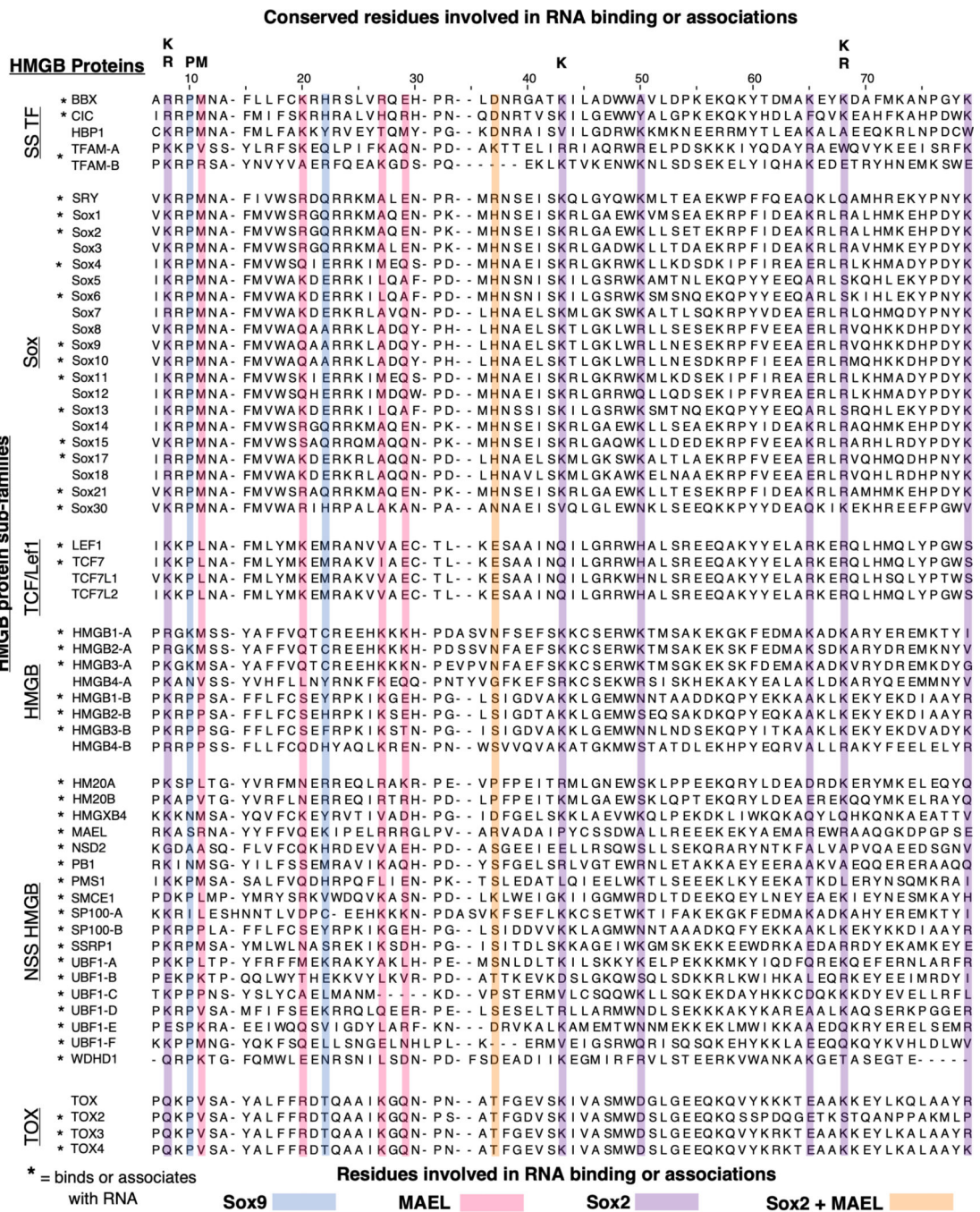


Figure 2. HMGB-DBD sequence similarities suggests RNA binding potential. Sequence alignment, using MUSCLE, (Madeira et al., 2022) of the 48 human HMGB domains with sequences obtained from UniProt (UniProt, 2021) On the left, HMGBs are split into their respective subfamilies which have higher conservation within the subfamily (Sox, TCF/Lef1, HMGB, TOX), other HMGBs do not have a defined subfamily (non-sequence specific (NSS) and sequence specific (SS)). Sox9 (blue)(Girardot et al., 2018), MAEL (magenta)(Genzor & Bortvin, 2015), Sox2 (purple)(Holmes et al., 2020), Sox2 + MAEL overlap (orange) mutated

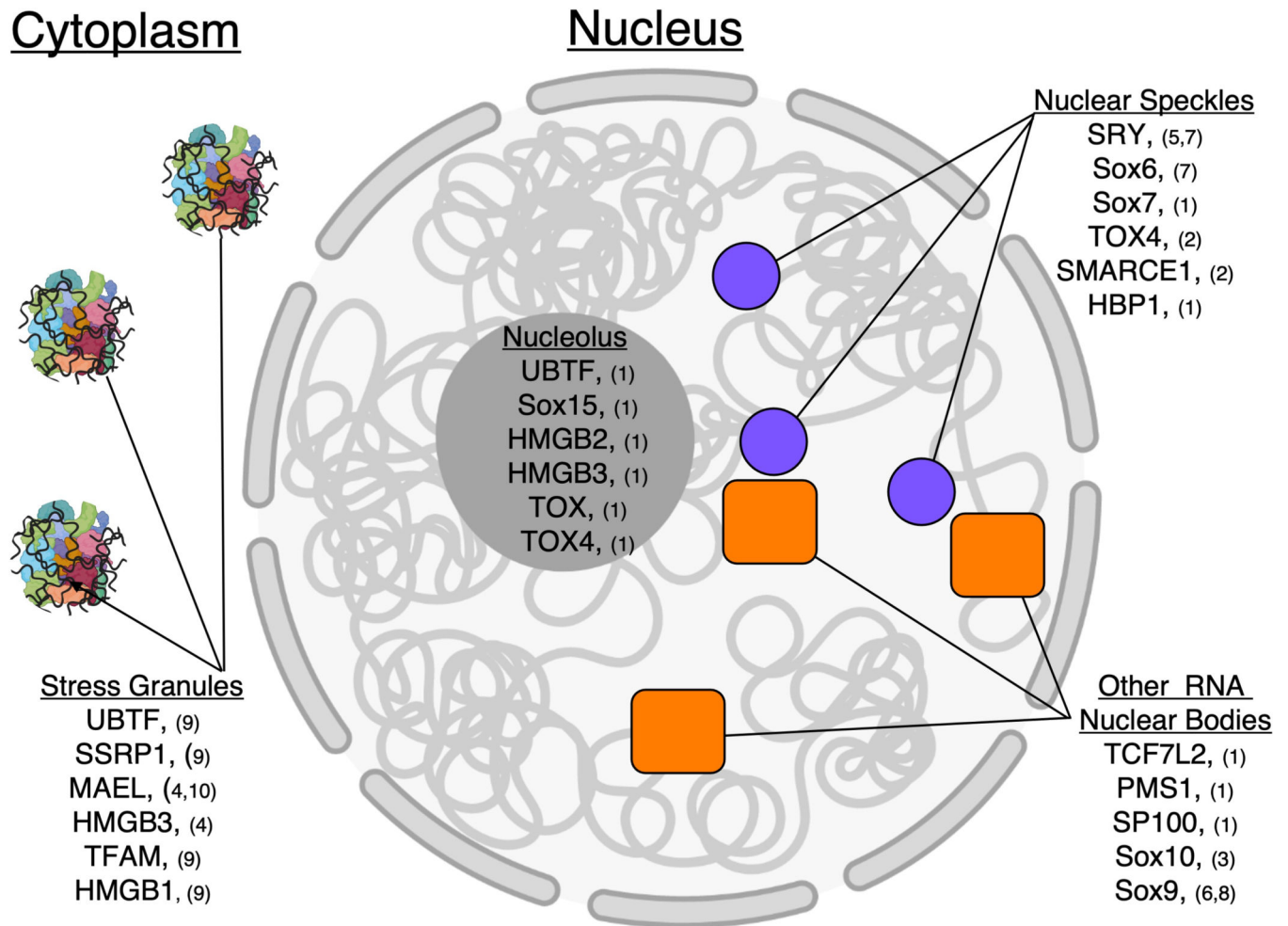
residues are involved in RNA binding or associations. Some of these residues are highly conserved throughout the HMGB family (top). * = HMGB with evidence for binding or associating with RNA.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Figure 3.**

List of HMGBs with evidence for localization within stress granules, nucleolus, nuclear speckles, or other RNP granules. HMGBs localized to **stress granules**: UBTF (Markmiller et al., 2018), SSRP1 (Markmiller et al., 2018), MAEL (Jain et al., 2016; Yuan et al., 2014), HMGB3 (Jain et al., 2016), TFAM (Markmiller et al., 2018), HMGB1 (Markmiller et al., 2018). HMGBs localized to the **nucleolus**: UBTF, Sox15, HMGB2, HMGB3, TOX, TOX4 (Thul et al., 2017). HMGBs localized to **nuclear speckles**: SRY (Sato et al., 2011; Ohe et al., 2002), Sox6 (Ohe et al., 2002), Sox7 (Thul et al., 2017), TOX4 (Thul et al., 2017), SMARCE1 (Saitoh et al., 2004), HBP1 (Thul et al., 2017). HMGBs localized to **other RNP granules**: TCF7L2 (Thul et al., 2017), PMS1 (Thul et al., 2017), SP100 (Thul et al., 2017), Sox10 (Chaoui et al., 2015), Sox9 (Hata et al., 2008; Penrad-Mobayed et al., 2018).

A. HMGB protein - protein interactors with an RNA centric Gene Ontology Assignment
 HMGB protein - protein interactors without an RNA centric Gene Ontology Assignment

HMGBs	# of APID protein interactors	Gene Ontology Aspects												
		Molecular Function		Biological Process					Cellular Component					
		% with DNA binding	% with RNA binding	Splicing	RNA transport	RNA stability	RNA processing	Other RNA metabolism	RNP complex	Spliceosome complex	Nuclear speckle	Cajal body	Nucleolus	PML body
BBX	78	22	35
CIC	37	22	16
HBP1	119	18	46
LEF1	66	35	6
Sox1	6	83	0
Sox2	580	21	31
Sox3	17	53	24
Sox4	50	20	12
Sox5	52	27	10
Sox6	18	39	17
Sox7	9	33	22
Sox8	19	79	11
Sox9	26	42	15
Sox10	45	53	18
Sox11	4	100	0
Sox12	4	50	0
Sox13	24	38	17
Sox14	19	32	11
Sox15	3	100	0
Sox17	7	29	29
Sox18	5	60	40
Sox21	8	50	38
Sox30	19	11	21
SRY	23	65	0
TCF7	17	24	12
TCF7L1	9	44	0
TCF7L2	197	19	26
TFAM	320	6	25
HMG20A	159	19	9
HMG20B	57	18	11
HMG21	386	19	39
HMG22	116	25	28
HMG23	41	22	29
HMG24	4	0	0
HMG25	63	21	22
MAEL	16	25	81
NSD2	214	22	48
PBRM1	101	40	14
PMS1	80	21	19
SMARCE1	207	25	10
SP100	211	17	31
SSRP1	228	27	25
TOX	9	22	0
TOX2	31	16	6
TOX3	18	17	0
TOX4	101	31	21
UBTF	152	20	20
WDHD1	85	15	21

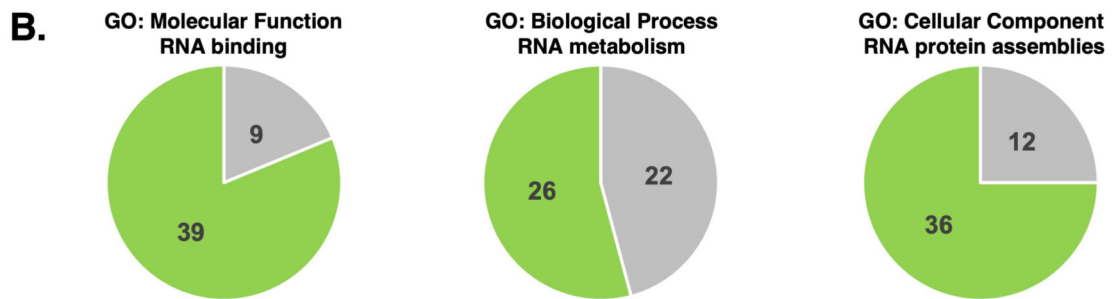


Figure 4. HMGB protein interactome is enriched with biologically productive RNA binding proteins. APID (Alonso-Lopez et al., 2019; Alonso-Lopez et al., 2016) coupled with a Gene Ontology analysis (Ashburner et al., 2000; Gene Ontology, 2021) identified HMGBs that interact with proteins involved in the following: Molecular Function, binds DNA or RNA; Biological Process, RNA metabolism; and Cellular Component, found within RNP assemblies. **A.** If at least one HMGB protein interactor had the Gene Ontology assignment it is colored green and if no evidence was found it is colored grey. # of HMGB protein interactors from

APID, the % of interactors with DNA or RNA binding assignments, and interactors with or without evidence for each Gene Ontology assignment. **B.** Total number of HMGBs from **A.** compiled into pie charts for each Gene Ontology assignment.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

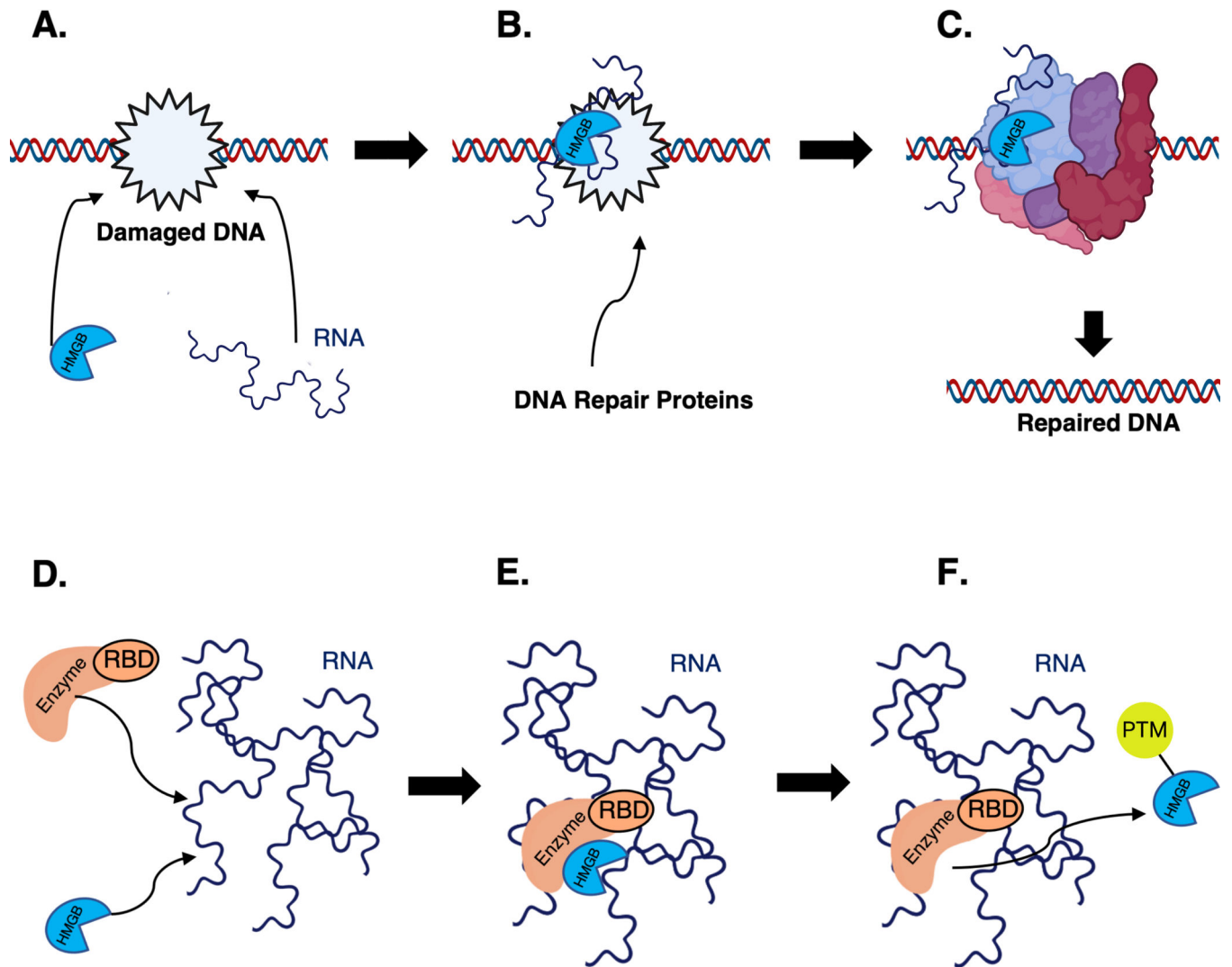


Figure 5. Cartoon of mechanism for how HMGBs interact with RNAs that act as scaffolds. **A.** HMGBs recognize damaged DNA and recruit or interact with RNA at sites of DNA damage. **B.** HMGBs and RNA recruit DNA repair proteins **C.** DNA repair complex repairs damaged DNA. **D.** HMGBs and enzymes with RNA binding domains (RBD) bind to the same RNA. **E.** This brings RNA binding enzymes and their HMGB substrates together to expedite post translation modifications (PTM). **F.** HMGB is released following PTM.

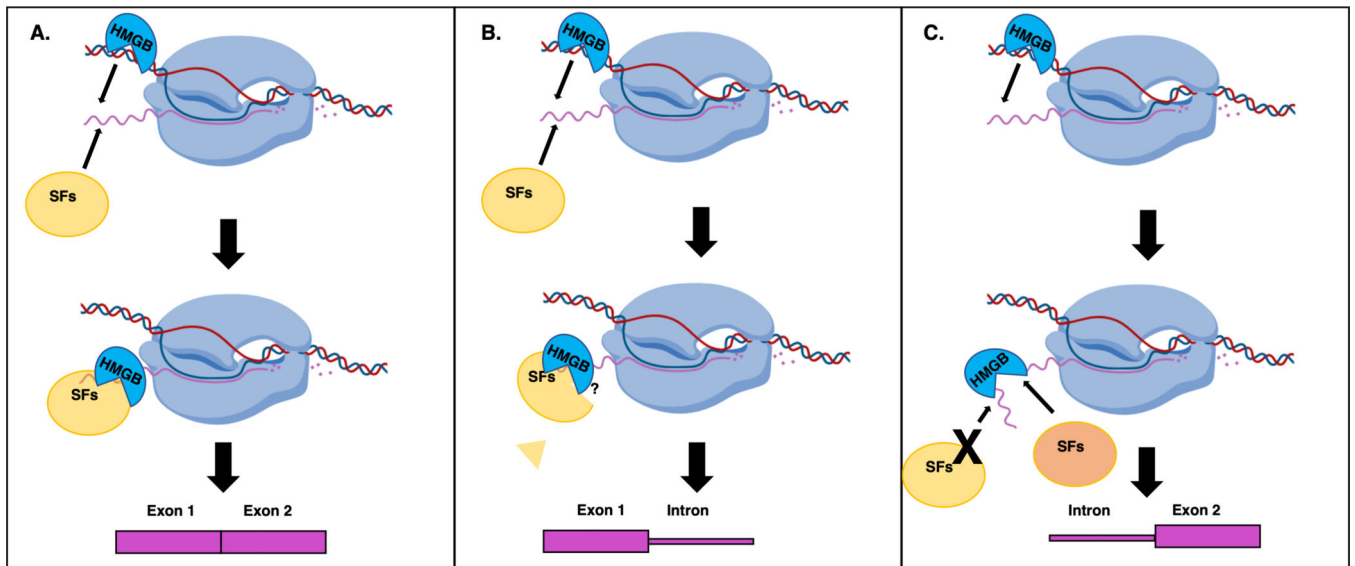


Figure 6.

Cartoon of possible mechanisms for how HMGBs impact splicing outcomes by binding RNA. **A.** HMGBs and splicing factors (SFs) form RNA protein complexes upon nascent RNA, impacting splicing outcomes. **B.** HMGBs interaction with nascent RNA may block SFs access to RNA or alter splicing complex composition, resulting in altered splicing. **C.** HMGBs may bind and alter the structure of the RNA to recruit specific SFs or to prevent binding of SFs.

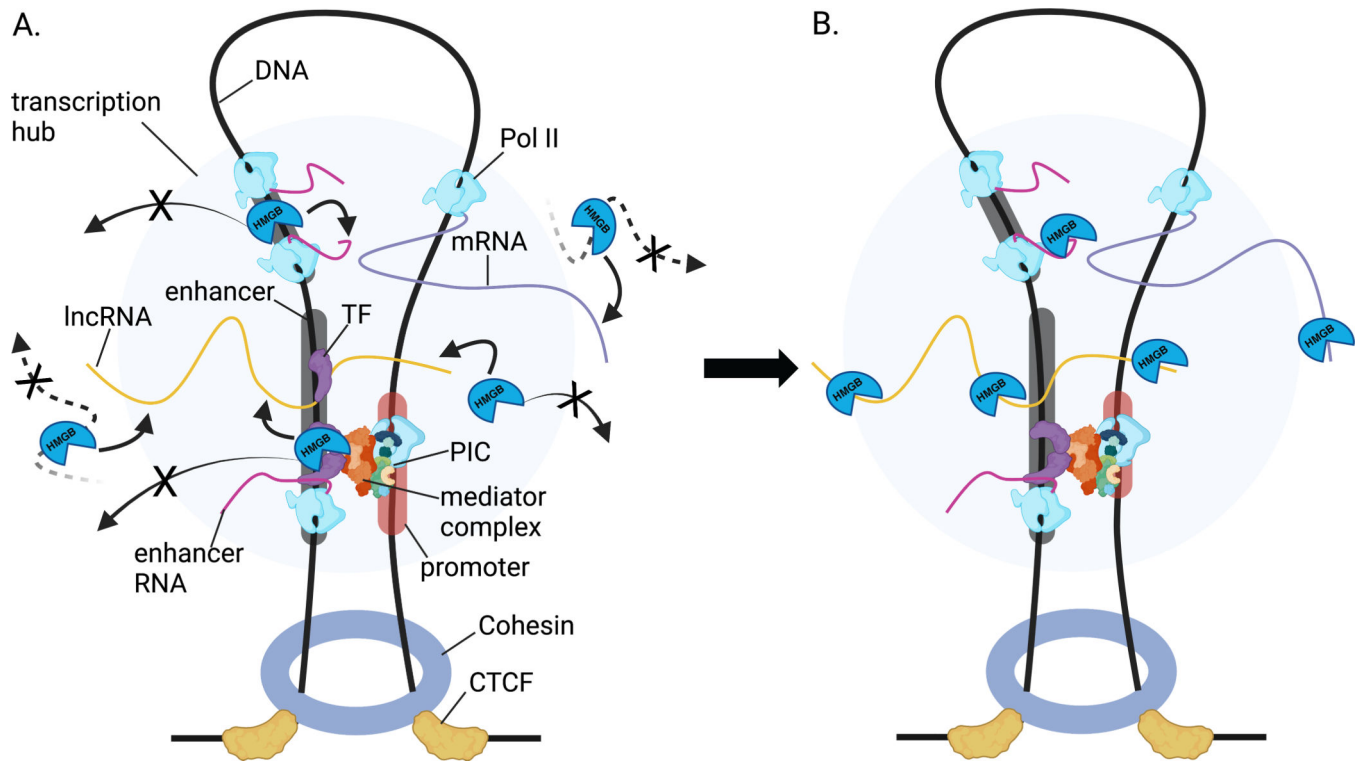


Figure 7. Cartoon depiction of HMGB-RNA interactions impacting HMGB genomic localization **A.** HMGBs may assist in the formation and maintenance of transcription hubs. Rather than HMGBs diffusing away from their genomic bound sites, nascent or stable, processed RNAs interact with HMGBs. HMGBs not within the transcription hub but diffusing nearby may be captured by RNAs on the periphery of the transcription hub. **B.** HMGB interactions with RNA at transcription hubs maintain or increase HMGBs concentration within.

Table 1.

Summary of literature evidence for HMGB family interactions with RNA. List of HMGBs that directly interact with RNA are labeled as binds. HMGBs that interact with RNA, but it is unknown whether the interaction is direct or indirect are assigned as associates. *HMGBs that do not yet have direct RNA binding or association evidence are listed at the bottom. Unknown = UK.

HMGB protein	Binds to RNA, (reference)	Associates with RNA, (reference)
BBX	(Yan et al., 2021)	(Caudron-Herger et al., 2019)
CIC	(Queiroz et al., 2019)	UK
LEF1	(Hamilton et al., 2022)	(B. Li et al., 2021)
Sox1	(Hou et al., 2020)	UK
Sox2	(Hamilton et al., 2022; Holmes et al., 2020), (Cajigas et al., 2021; Hou et al., 2020; Tung et al., 2010)	(Bai et al., 2022; Du et al., 2021; Guo et al., 2018; Jiang et al., 2018; Jing et al., 2020; Q. Li et al., 2021; Ng et al., 2013; Ng et al., 2012; Samudiyata et al., 2019; Stamou et al., 2020; Zhang et al., 2020; H. Zhang et al., 2021; Zhou et al., 2022; Zhuang et al., 2015)
Sox4	UK	(Chen et al., 2020)
Sox6	(Hamilton et al., 2022)	UK
Sox9	(Hamilton et al., 2022)	(Girardot et al., 2018; Pei et al., 2021; Penrad-Mobayed et al., 2018)
Sox10	UK	(Stamou et al., 2020)
Sox11	(Hamilton et al., 2022)	UK
Sox13	(Yan et al., 2021)	UK
Sox15	(Hamilton et al., 2022)	UK
Sox17	(Hamilton et al., 2022)	UK
Sox21	(Hamilton et al., 2022)	UK
Sox30	(Hamilton et al., 2022)	UK
SRY	(Hamilton et al., 2022)	UK
TCF7	(Park et al., 2005; Lee & Jeong, 2004)	UK
TFAM	(Castello et al., 2016; Trendel et al., 2019; Baltz et al., 2012; Brown et al., 2015; Castello et al., 2012; Dvir et al., 2021; Lyonais et al., 2017; Sysoev et al., 2016; Urdaneta et al., 2019; Yan et al., 2021)	(Caudron-Herger et al., 2019)
HM20A	UK	(Mallam et al., 2019)
HM20B	UK	(Caudron-Herger et al., 2019; Mallam et al., 2019)
HMGB1	(Baltz et al., 2012; Beckmann et al., 2015; Bell et al., 2008; Brieno-Enriquez et al., 2019; Castello et al., 2012; Castello et al., 2016; Choi et al., 2020; Dvir et al., 2021; He et al., 2016; Huang et al., 2018; Ignarski et al., 2019; Lee et al., 2012; Oh & Lee, 2014; Queiroz et al., 2019; Sofiadis et al., 2021; Trendel et al., 2019; Yan et al., 2021; Yanai et al., 2009; Yu et al., 2015)	(Gao et al., 2017; Lou et al., 2021; Scherer et al., 2020; Yanai et al., 2011; Yamanaka et al., 2015)
HMGB2	(Baltz et al., 2012; Beckmann et al., 2015; Castello et al., 2012; Castello et al., 2016; He et al., 2016; Iadevaia et al., 2020; Ignarski et al., 2019; Kwon et al., 2013; Queiroz et al., 2019; Trendel et al., 2019; Yamanaka et al., 2015; Yan et al., 2021; Yanai et al., 2009)	(Han et al., 2019; Yamanaka et al., 2015)
HMGB3	(Baltz et al., 2012; Beckmann et al., 2015; Castello et al., 2016; Khoury et al., 2021; Queiroz et al., 2019; Trendel et al., 2019; Yan et al., 2021; Yanai et al., 2009)	UK
HMGXB4	(He et al., 2016; Urdaneta et al., 2019)	UK
MAEL	(Chen et al., 2015; Genzor & Bortvin, 2015)	UK

HMGB protein	Binds to RNA, (reference)	Associates with RNA, (reference)
NSD2	(Sysoev et al., 2016; Yan et al., 2021)	(Caudron-Herger et al., 2019)
PBRM1	(He et al., 2016; Ignarski et al., 2019), (Bao et al., 2018; Urdaneta et al., 2019), (Yan et al., 2021)	(Caudron-Herger et al., 2019)
PMS1	(Castello et al., 2016),	UK
SMARCE1	(He et al., 2016; Urdaneta et al., 2019)	(Caudron-Herger et al., 2019)
SP100	(Queiroz et al., 2019; Yan et al., 2021)	UK
SSRP1	(Baltz et al., 2012; Bao et al., 2018; Beckmann et al., 2015; Castello et al., 2012; He et al., 2016; Huang et al., 2018; Ignarski et al., 2019; Perez-Perri et al., 2018; Queiroz et al., 2019; Urdaneta et al., 2019; Yan et al., 2021)	(Caudron-Herger et al., 2019; Mallam et al., 2019; Wang et al., 2018)
TOX2	UK	(Caudron-Herger et al., 2019)
TOX3	UK	(Caudron-Herger et al., 2019)
TOX4	(He et al., 2016; Urdaneta et al., 2019; Yan et al., 2021)	(Caudron-Herger et al., 2019)
UBTF	(He et al., 2016; Trendel et al., 2019), (Baltz et al., 2012; Ignarski et al., 2019), (Bao et al., 2018; Castello et al., 2012; Perez-Perri et al., 2018; Queiroz et al., 2019; Yan et al., 2021)	(Caudron-Herger et al., 2019; Scherer et al., 2020)
WDHD1	(Bao et al., 2018; Queiroz et al., 2019; Urdaneta et al., 2019; Yan et al., 2021)	UK

* **HMGBs with unknown RNA binding or association experimental evidence:** HBP1, Sox3, Sox5, Sox7, Sox8, Sox12, Sox14, Sox18, TCF7L1, TCF7L2, HMGB4, TOX.

Table 2.

List of HMGBs that directly interact with RNA *in vitro*. If evidence was found, RNA and DNA binding preferences with affinities are listed followed by references. Unknown = UK.

HMGB protein (species)	HMGBs Bind RNA <i>in vitro</i>		DNA Binding <i>in vitro</i>
	Preferred RNA Ligands, [binding affinity] (reference)	Less preferred RNA ligands, [binding affinity] (reference)	DNA Ligands, [binding affinity] (reference)
Sox2, (human)	dsRNA elements including: stem loops with internal bulges, [~10–80 nM]; tRNA ^{Leu} , [27 nM]; riboswitch, [26 nM], EVF2 lncRNA (Cajigas et al., 2021; Holmes et al., 2020; Hou et al., 2020; Tung et al., 2010)	ssRNA, [500 nM] (Holmes et al., 2020)	consensus dsDNA, [0.4–15 nM]; nucleosomal DNA, [0.3–1.4 nM]; non-consensus dsDNA, [550 nM] (Holmes et al., 2020; Moosa et al., 2018; Palasingam et al., 2009; Scaffidi & Bianchi, 2001; Soufi et al., 2015)
Sox Family HMGB domains, (human)	RNA 4way-junction, [1.6–96 nM]; Stem loop w/ internal bulge, [6.9–360 nM]; ES2 lncRNA fragment, [1.5–100 nM]; TERRA G-quadruplex, [0.2–130 nM] (Hamilton et al., 2022)	UK	consensus dsDNA, [0.5–31 nM]; non-consensus dsDNA, [310–1200 nM] (Connor et al., 1994; Hamilton et al., 2022; Mertin et al., 1999; Palasingam et al., 2009)
Sox30, (human)	RNA 4way-junction, [990 nM]; ES2 lncRNA fragment, [910 nM]; TERRA G-quadruplex, [460 nM] (Hamilton et al., 2022)	Stem loop w/ internal bulge, [3100 nM] (Hamilton et al., 2022)	consensus dsDNA, [72 nM]; non-consensus dsDNA, [5000 nM] (Hamilton et al., 2022)
LEF1, (human)	RNA 4way-junction, [89 nM], ES2 lncRNA fragment, [61 nM], TERRA G-quadruplex, [4.1 nM] (Hamilton et al., 2022)	Stem loop w/ internal bulge, [630 nM] (Hamilton et al., 2022)	consensus dsDNA, [1–1.6 nM]; non-consensus dsDNA, [1000 nM] (Giese et al., 1991; Hamilton et al., 2022)
TCF7, (human)	dsRNA elements containing stem loop with internal bulge, [100–500 nM] (Lee & Jeong, 2004; Park et al., 2005)	dsRNA elements with stem loop with internal bulges (Lee et al., 2005)	UK
TFAM, (human)	RNA:DNA 4WJ, [300 nM]; RNA 4WJ, [279 and 99 nM]; tRNA, [71 nM]; Hybrid DNA:RNA and RNA G-quadruplexes, [1–2 nM] (Brown et al., 2015; Lyonais et al., 2017)	ssRNA (poly AC or U), dsRNA, dsRNA hairpin loops, linear dsRNA/DNA hybrids (Brown et al., 2015)	consensus dsDNA, [0.3–9 nM]; non-consensus dsDNA [100 nM]; DNA 4WJ, [63 nM] (Brown et al., 2015; Lyonais et al., 2017)
HMGD, (fly)	TAR RNA, [29 nM]; dsRNA stem loop, [36 nM]; linear dsRNA, [70 nM]; bulged dsRNA, [53 nM] (Arimondo et al., 2000)	ssRNA (Arimondo et al., 2000)	dsDNA and bulged DNA, [~85 nM]; DNA microcircle, [~1 nM] (Payet & Travers, 1997)
HMGB1, (human)	dsRNA elements including stem loops and internal bulges, siRNA, poly (I:C) (Bell et al., 2008; Choi et al., 2020; Lee et al., 2012; Oh & Lee, 2014; Yanai et al., 2009; Yu et al., 2015)	<i>E. coli</i> tRNA (Yu et al., 2015)	4WJ DNA; DNA minicircles, [0.01–5 nM]; semicatenated DNA, [~0.16 pm]; ssDNA; cruciform DNA; supercoiled DNA; Z-DNA (Gaillard & Strauss, 2000; Jaouen et al., 2005; Webb et al., 2001; Webb & Thomas, 1999)
HMGB2, (human)	LRP1-AS ncRNA (Yamanaka et al., 2015)	LRP1-mRNA (Yamanaka et al., 2015)	chromium damaged DNA, [1 nM]; dsDNA, [610 nM] semicatenated DNA, (Gaillard & Strauss, 2000; Wang et al., 1997; Yoshioka et al., 1999)
HMGB3, (human)	HIV-1 Tat mRNA containing dsRNA elements, including stem loops and internal bulges (Khoury et al., 2021)	UK	UK
MAEL, (mouse, bombyx)	RNA 4WJ, [640 nM]; ssRNA, [130–220] (Chen et al., 2015; Genzor & Bortvin, 2015)	dsRNA, RNA hairpins, ssRNA (Genzor & Bortvin, 2015)	DNA 4wj, [14 nM] (Genzor & Bortvin, 2015)
CsHMG, (plant)	poly (U) and poly (G) RNA (de Souza et al., 2012)	poly (AU) RNA (de Souza et al., 2012)	dsDNA (de Souza et al., 2012)
DssRP, (fly)	poly (U) and poly (G) RNA (Hsu et al., 1993)	poly (A) or poly (C) RNA (Hsu et al., 1993)	ssDNA (Hsu et al., 1993)