V23 Regular vs. alternative splicing

Regular splicing

- mechanistic steps
- recognition of splice sites

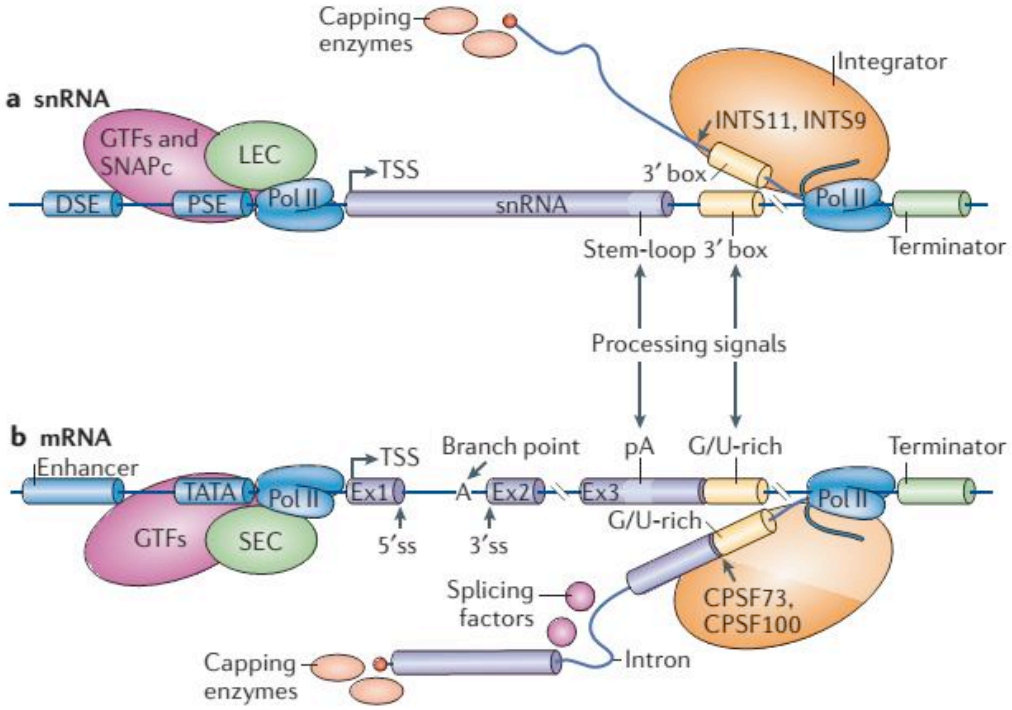
Alternative splicing

- different mechanisms
- how frequent is alternative splicing?

Effect of alternative splicing on protein – protein interactions

Interplay of alternative splicing and epigenetic modifications

Transcription + processing of snRNAs and mRNAs



Show are *cis*-acting elements and *trans*-acting factors involved in the expression of **(a)** sm-class small nuclear RNA (snRNA) genes (that are part of the spliceosome) and of **(b)** protein-coding mRNA genes.

The distal sequence element (**DSE**) and proximal sequence element (**PSE**) of snRNAs are roughly equivalent to the **enhancer** and **TATA** box elements, respectively, of mRNA genes.

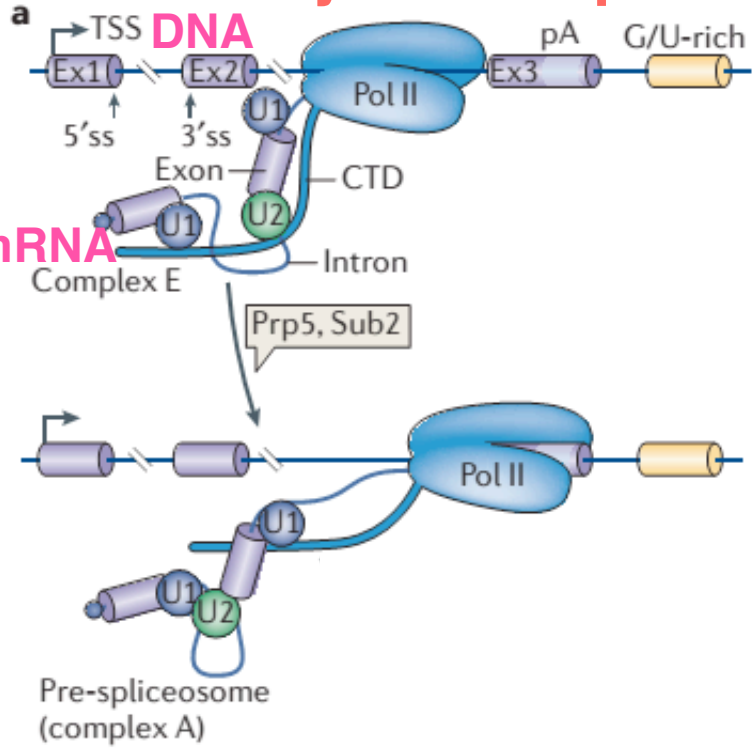
Ex, exon; pA, polyA signal; ss, splice site; TSS, transcription start site.

snRNA promoters recruit the **little elongation complex (LEC),** whereas mRNA promoters recruit the **super elongation complex (SEC)**. Initiation of snRNA transcription requires **general transcription factors (GTFs)**, as well as the **snRNA-activating protein complex (SNAPc**).

Integrator subunit 11 (INTS11) and INTS9 have sequence similarities to the mRNA 3´-processing factors **cleavage and polyadenylation specificity factor** 73 kDa subunit (CPSF73) and CPSF100, respectively.

Matera & Wang, Nature Rev Mol Cell Biol 15, 108–121 (2014)

Assembly of the splicesome + splicing steps of pre-mRNA



Spliceosome assembly takes place at sites of transcription.

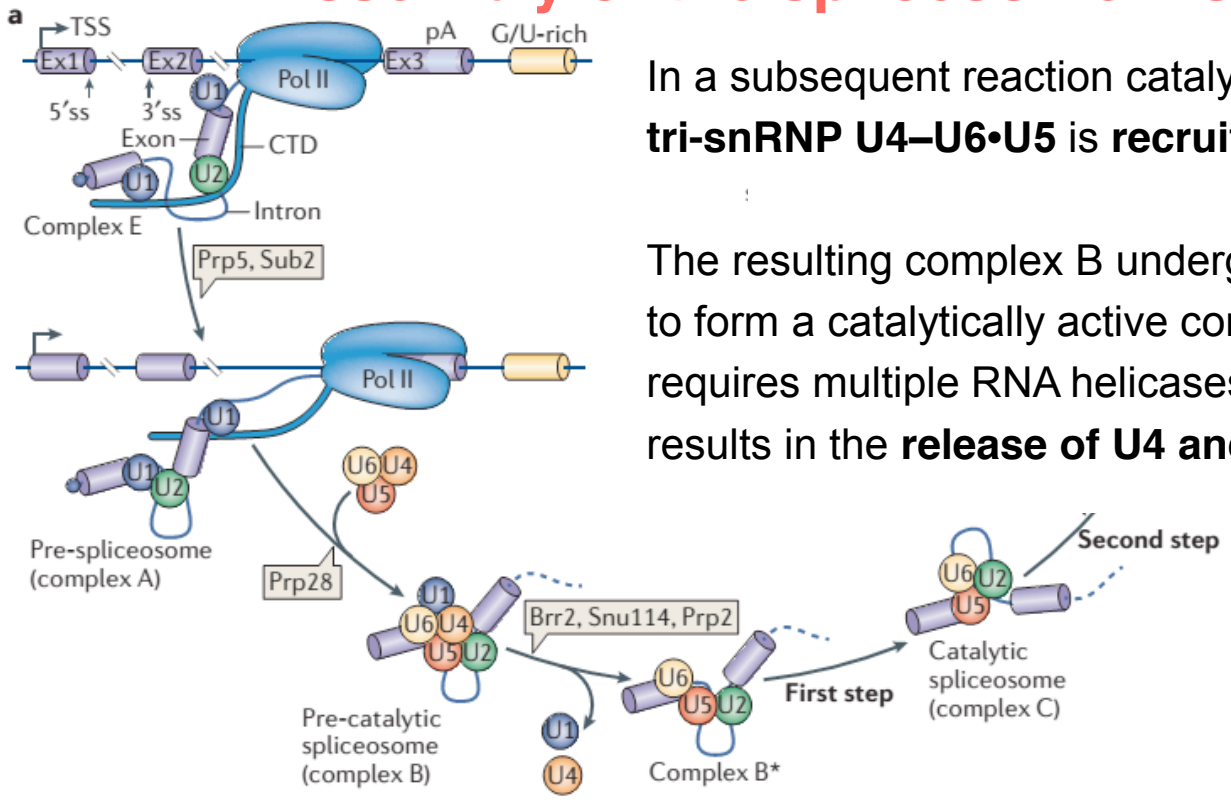
The U1 and U2 small nuclear ribonucleoproteins (**snRNPs**) assemble onto the pre-mRNA in a co-transcriptional manner through recognition of the 5´ splice site (5´ss) and 3´ss, which is mediated by the carboxy-terminal domain (CTD) of polymerase II.

The U1 and U2 snRNPs then interact with each other to form the pre-spliceosome (**complex A**).

This process is dependent on DExD/H helicases pre-mRNA-processing 5 (**Prp5**) and **Sub2**.

Matera & Wang,
Nature Rev Mol Cell Biol
15, 108–121 (2014)

Assembly of the splicesome + splicing steps



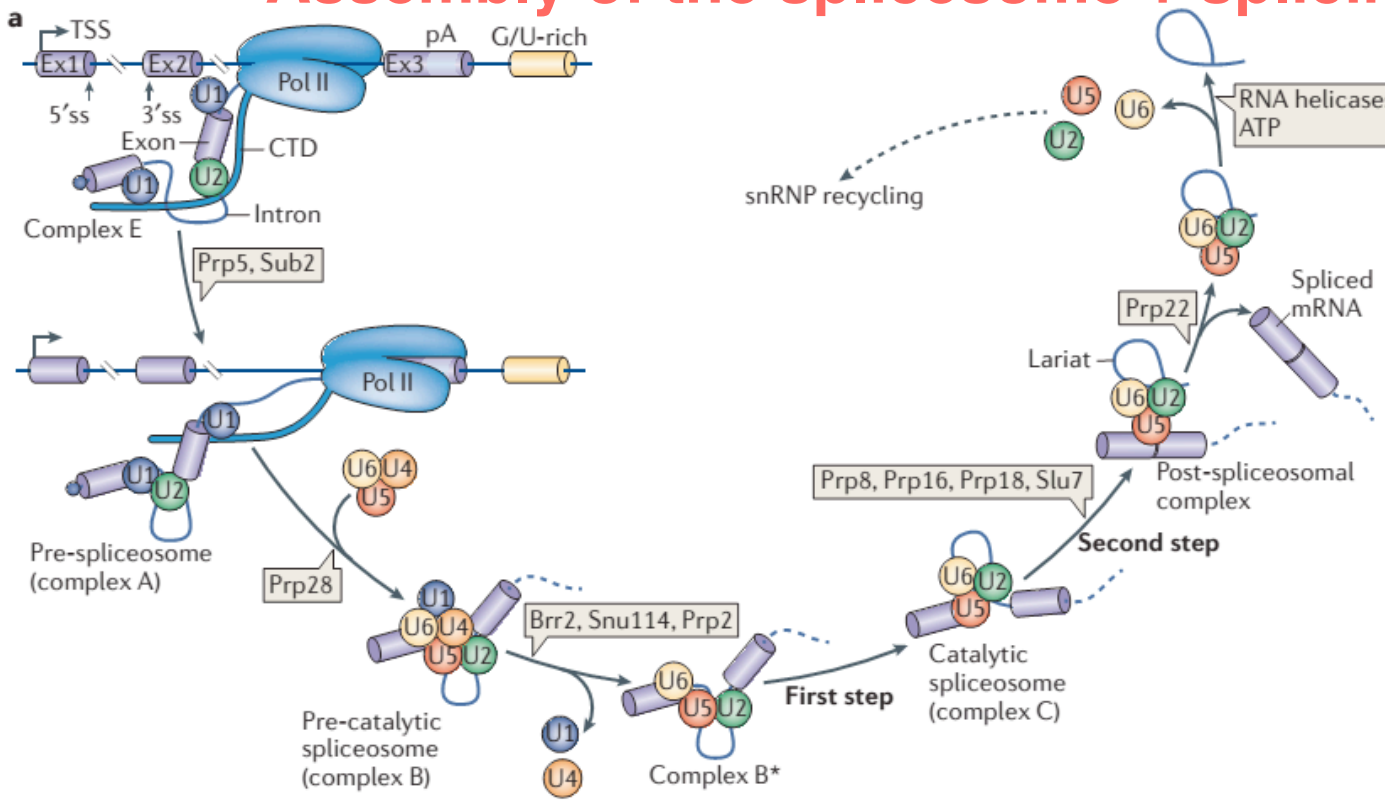
In a subsequent reaction catalysed by Prp28, the preassembled **tri-snRNP U4–U6•U5** is **recruited** to form **complex B**.

The resulting complex B undergoes a series of rearrangements to form a catalytically active complex B (**complex B***), which requires multiple RNA helicases (Brr2, Snu114, Prp2) and results in the **release of U4 and U1** snRNPs.

Complex B* then carries out the first catalytic step of splicing, generating **complex C**, which contains free exon 1 (Ex1) and the intron–exon 2 “lariat intermediate”.

Matera & Wang,
Nature Rev Mol Cell Biol
15, 108–121 (2014)

Assembly of the splicesome + splicing steps



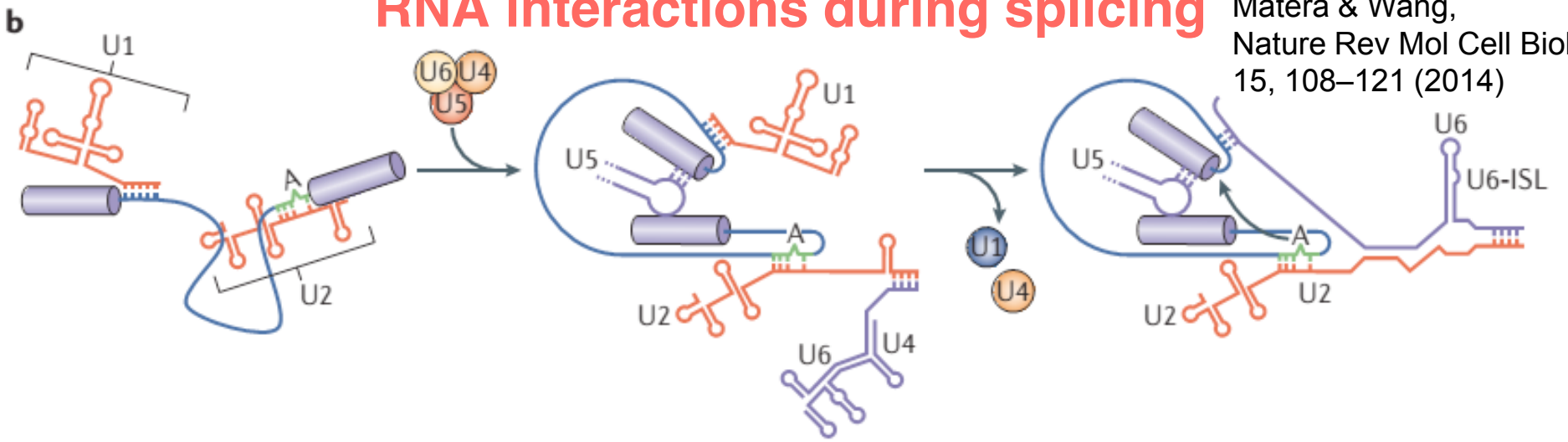
Matera & Wang,
Nature Rev Mol Cell Biol
15, 108–121 (2014)

Complex C undergoes additional rearrangements and then carries out the second catalytic step, resulting in a post-spliceosomal complex that contains the lariat intron and spliced exons.

Finally, the U2, U5 and U6 snRNPs are released from the mRNP particle and recycled for additional rounds of splicing. Release of the spliced product from the spliceosome is catalysed by the DExD/H helicase Prp22.

RNA interactions during splicing

Matera & Wang,
Nature Rev Mol Cell Biol
15, 108–121 (2014)



During splicing, RNA–RNA interactions are rearranged in a stepwise manner to create the catalytic centre of the spliceosome. Initially, U1 and U2 small nuclear RNA (snRNA) pair with the 5´ss and the branch point sequence within **complex A** (the branch point adenosine is indicated by the letter A).

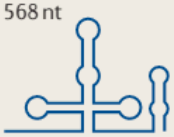
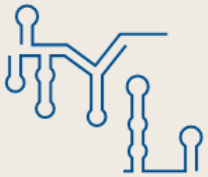
Subsequently, complex A associates with the U4–U6•U5 tri-snRNP, leading to new base pairs between U2 and U6 snRNA and between U5 snRNA and exonic sequences near the 5´ss.

The U4 snRNA is disassociated from U6 to expose the 5´ end of U6, which then base pairs with the 5´ss to displace U1 snRNA.

In the end, an extensive network of base-pairing interactions is formed between U6 and U2, juxtaposing the 5´ss and branch-point adenosine for the first catalytic step of splicing. The central region of U6 snRNA forms an intramolecular stem-loop (the U6-ISL), which is essential for splicing catalysis. pA, polyA signal.

Composition of spliceosomal snRNPs

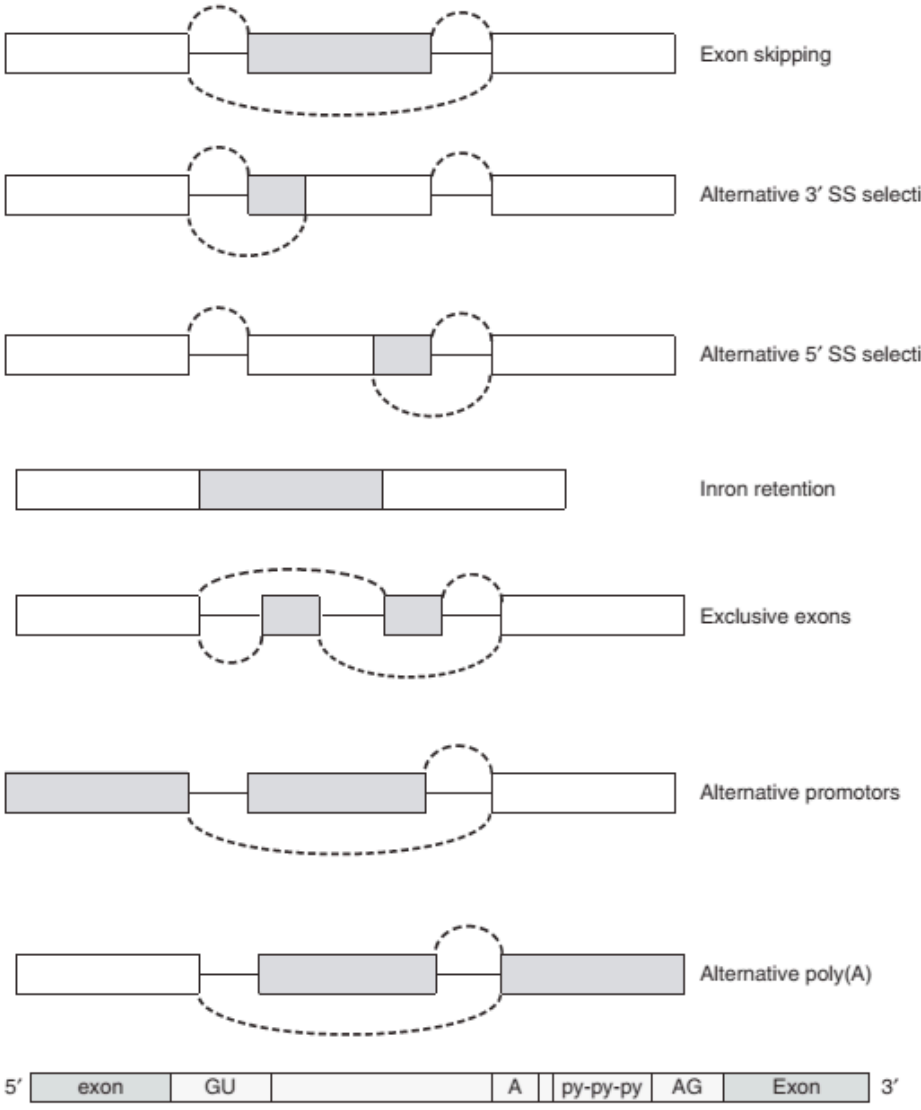
Table 1 | **Composition of major spliceosomal snRNPs***

snRNP	RNA secondary structure‡	Sm proteins	Other core proteins associated with snRNA	Associated proteins
U1	568 nt 	B, D3, G, E, F, D2 and D1	Snp1 (U1-70K), Mud1 (U1A) and Yhc (U1C)	Prp39, Prp40, Prp42, Snu71, Nam8, Snu56 and Urn1
U2	1,175 nt 	B, D3, G, E, F, D2 and D1	Lea1 (U2A′), Msl1 (U2B″), Prp9 (SF3B60), Prp11 (SF3A66), Prp21 (SF3A120), Rds3 (SF3B14B), Snu17 (SF3B14A; also known as p14), Hsh155 (SF3B155), Cus1 (SF3B145), Rse1 (SF3B130), Hsh49 (SF3B49) and Ysf3 (SF3B10)	U2AF35, Mud2 (U2AF65) and Msl5 (SF1; also known as BBP)
U4–U6	160 nt and 112 nt, respectively 	U4: B, D3, G, E, F, D2 and D1; U6: Lsm2–8	Prp3, Prp31, Prp4 and Snu13	
U5	179 nt for short form; 214 nt for long form 	B, D3, G, E, F, D2 and D1	Prp8, Prp6, Prp28, Brr2, Snu114, U5-40K and Dib1	Snu23, Prp38, Prp2, Spp2, Yju2 and Cbc2 (52K)
U4–U6•U5		U4 and U5: B, D3, G, E, F, D2 and D1 (i.e. two sets); U6: Lsm2–8	Prp3, Prp31, Prp4, Snu13/15.5K, Prp8, Prp6, Prp28, Brr2, Snu114, U5–40K, snRNP27 and Dib1	Snu23, Prp38, Prp2, Spp2, Yju2, Snu66 and Sad1

Cus1, cold-sensitive U2 snRNA suppressor 1; Lea1, looks exceptionally like U2A 1; Lsm, Sm-like; Msl, Male-specific lethal homologue; nt, nucleotides; Prp, pre-mRNA-processing; Rds3, regulator of drug sensitivity 3; Rse1, RNA splicing and ER to Golgi transport factor 1; Sad1, snRNP assembly-defective 1; SF, splicing factor; Spp2, suppressor of PRP; snRNA, small nuclear RNA; snRNP, small nuclear ribonucleoprotein; Snu, U5 small nuclear ribonucleoprotein component; U2AF, U2 auxiliary factor; Urn1, U2-U5-U6 snRNP, RES complex and NTC-interacting pre-mRNA-splicing factor 1. *The protein composition is primarily based on a mass spectrometric analysis of the yeast spliceosome[203]; certain regulatory factors that are closely associated with the core spliceosome (such as SR proteins) are not included. Proteins are listed using the budding yeast nomenclature unless there is no known yeast homologue. In certain cases, the common name of a metazoan homologue is also included in brackets. ‡The snRNA lengths are based on yeast transcripts.

Matera & Wang,
Nature Rev Mol Cell Biol
15, 108–121 (2014)

Mechanisms of alternative splicing



Gray: exons – white: introns
The (gray) protein coding regions are excluded/included in different transcripts.

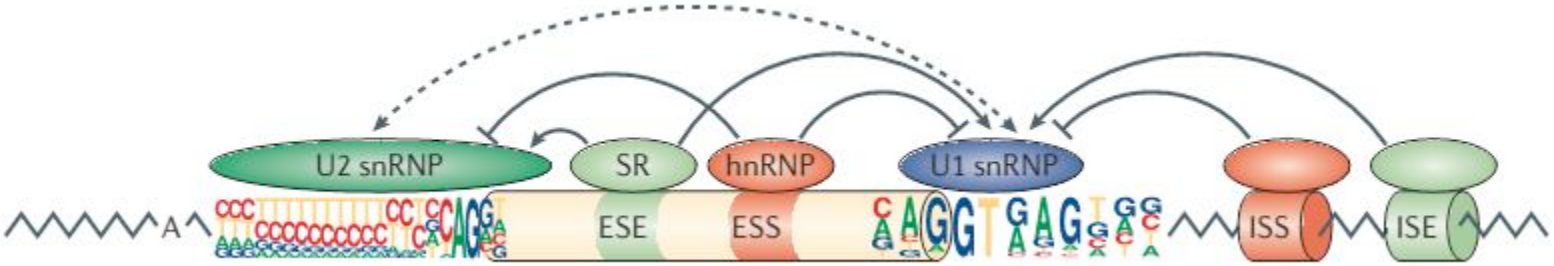
Bottom: most important sequence patterns related to a splicing.

The splicing starts with an AG site and is preceded with a non-AG (pyrimidine rich) region preceded by the branch point that includes an Adenosine residue.

The 5’ end of the intro contains an almost invariant GU sequence.

Light & Elofsson Curr Opin Struct Biol (2013) 23: 451-458

Regulation of alternative splicing

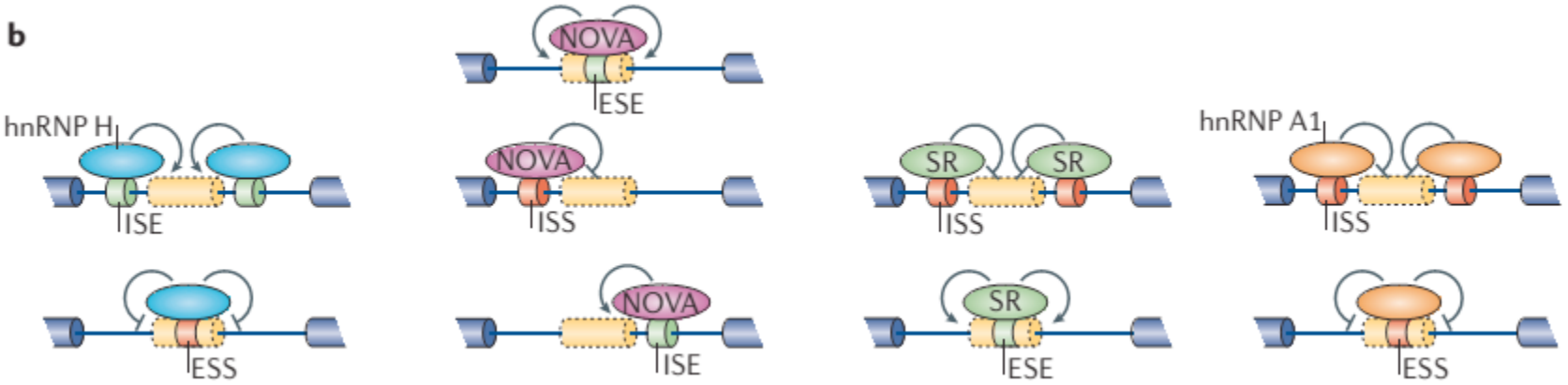


Splice site choice is regulated through *cis*-acting splicing regulatory elements (SREs) and *trans*-acting splicing factors. On the basis of their relative locations and activities, SREs are classified as exonic splicing enhancers (**ESEs**), intronic splicing enhancers (**ISEs**), exonic splicing silencers (**ESSs**) or intronic splicing silencers (**ISSs**). These SREs specifically recruit splicing factors to promote or inhibit recognition of nearby splice sites.
Common splicing factors include SR proteins, which recognize ESEs to promote splicing, as well as various heterogeneous nuclear ribonucleoproteins (hnRNPs), which typically recognize ESSs to inhibit splicing.

Show sequence motifs are the consensus motifs of splice sites. The height of each letter represents the nucleotide frequency in each position. The dashed arrow represents the formation of the exon definition complex.

Matera & Wang,
Nature Rev Mol Cell Biol
15, 108–121 (2014)

Activity of splicing factors and SREs

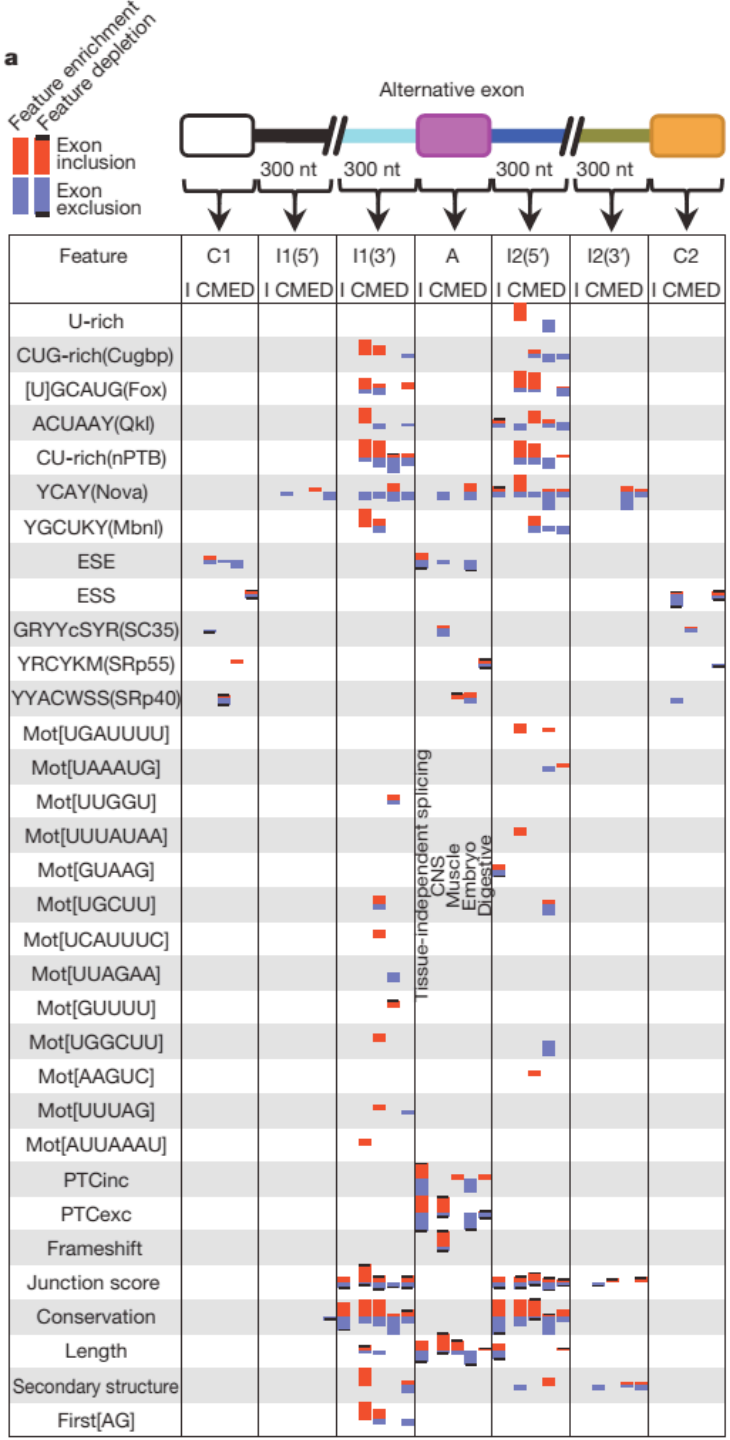


The activity of splicing factors and *cis*-acting SREs is context-dependent.
Shown are 4 well-characterized examples.
Oligo-G tracts, recognized by hnRNP H, function as ISEs to promote splicing when they are located inside an intron (top), and as ESSs when located within exons (bottom).
YCAY motifs, recognized by neuro-oncological ventral antigen (NOVA), act as ESEs when located inside an exon (top), as ISSs when located in the upstream intron of an alternative exon (middle) and as ISEs when located in the downstream intron (bottom).
Binding sites for SR proteins and hnRNP A1 also have distinct activities when located at different regions on the pre-mRNA.

Matera & Wang,
Nature Rev Mol Cell Biol
15, 108–121 (2014)

The splicing code

Barash et al. Nature 465, 53- (2010)



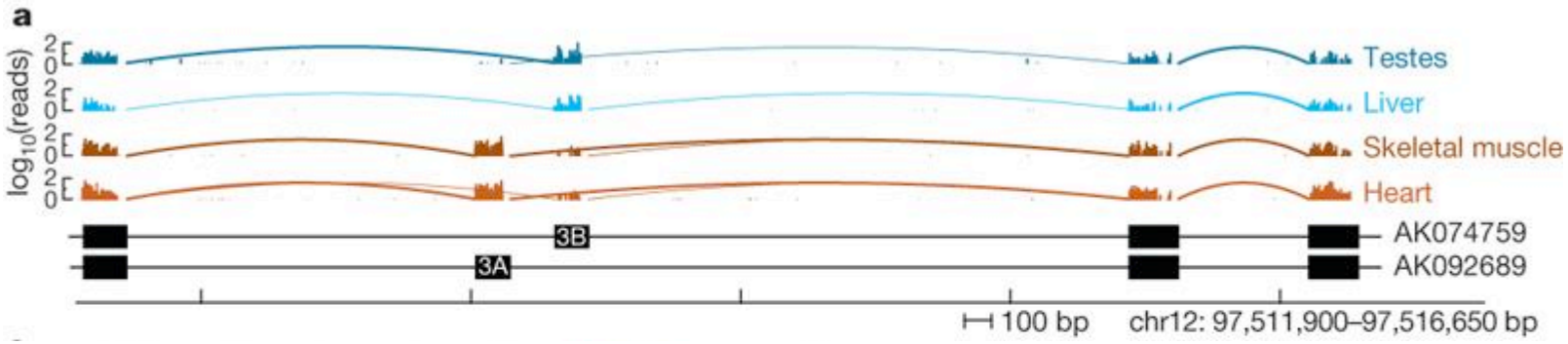
Top panel: defines 3 regions before „alternative exon“ (**A**) and behind.

Each table cell shows the region-specific activity of each feature in increased exon inclusion (red bar) or exclusion (blue bar) in 5 different mouse tissues: CNS (C), muscle (M), embryo (E) and digestive (D) tissues, plus a tissue-independent mixture (I).

Bar size conveys enrichment P-value;
$P < 0.005$ in all cases.

Potential feature binding proteins (Cugbp, Fox, Okl, PTB, Nova, Mbnl, SC35, SRp55, SRp40) are shown in parentheses.

Tissue-specific alternative splicing



mRNA-Seq reads mapping to a portion of the *SLC25A3* gene locus.
SLC25a3 is a mitochondrial phosphate carrier.

The number of mapped reads starting at each nucleotide position is displayed ($\log_{10}$) for the tissues listed at the right.

Arcs represent junctions detected by splice junction reads.

Bottom: exon/intron structures of representative transcripts containing mutually exclusive exons 3A and 3B (GenBank accession numbers on the right).

Wang et al. Nature (2008) 456: 470-6

tissue-specific regulation of alternative mRNA isoforms

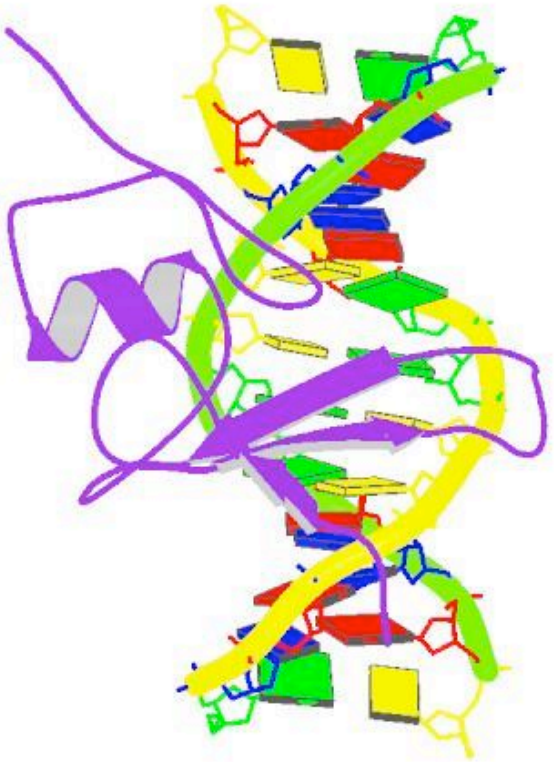
Alternative transcript events	Total events ($\times 10^3$)	Number detected ($\times 10^3$)	Both isoforms detected	Number tissue-regulated
Skipped exon	37	35	10,436	6,822
Retained intron	1	1	167	96
Alternative 5′ splice site (A5SS)	15	15	2,168	1,386
Alternative 3′ splice site (A3SS)	17	16	4,181	2,655
Mutually exclusive exon (MXE)	4	4	167	95
Alternative first exon (AFE)	14	13	10,281	5,311
Alternative last exon (ALE)	9	8	5,246	2,491
Tandem 3′ UTRs	7	7	5,136	3,801
Total	105	100	37,782	22,657

Blue, **red**, grey: mapped reads supporting expression of upper isoform, lower isoform or both isoforms.

Columns 1–4 show the numbers of events of each type:
(1) supported by cDNA and/or EST data;
(2) with 1 isoform supported by mRNA-Seq reads;
(3) with both isoforms supported by reads;
(4) events detected as tissue-regulated (Fisher's exact test).

Wang et al. Nature (2008) 456: 470-6

MBD2 recognizes methylated cytosines



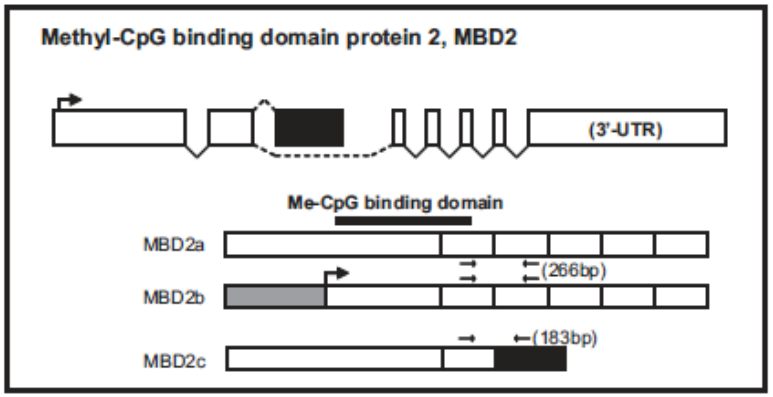
Ohki et al. (2001) Cell **105**: 487-497

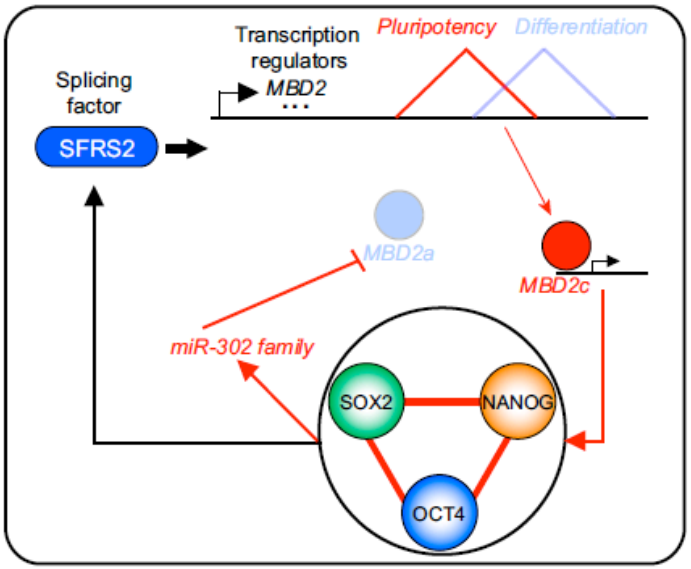
MBD2 is alternatively spliced and then plays a role for maintenance of pluripotency

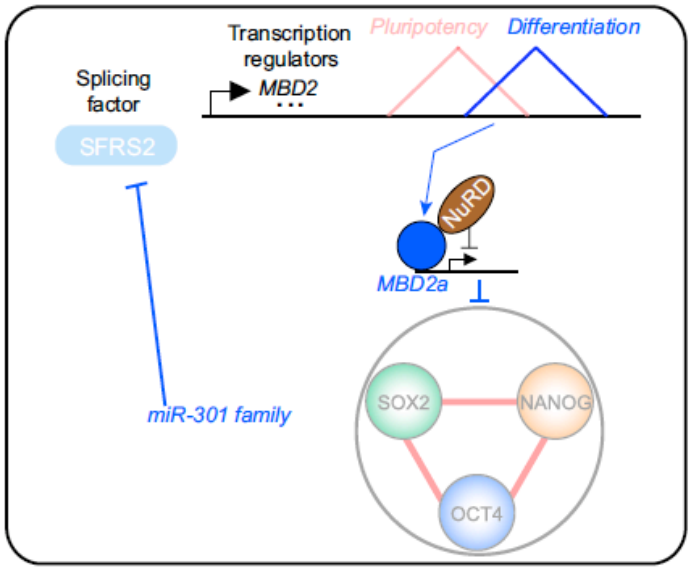
Cell Stem Cell
Short Article

Alternative Splicing of MBD2 Supports Self-Renewal in Human Pluripotent Stem Cells

Yu Lu,[1,2,3,13] Yuin-Han Loh,[2,4,5,13] Hu Li,[6,7,8,13] Marcella Cesana,[2,4] Scott B. Ficarro,[1,2,9] Jignesh R. Parikh,[9,10] Nathan Salomonis,[11] Cheng-Xu Delon Toh,[5] Stelios T. Andreadis,[12] C. John Luckey,[3,14] James J. Collins,[6,7,14] George Q. Daley,[2,4,14,*] and Jarrod A. Marto[1,2,9,14,*]
[1]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, MA 02115, USA
[2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA







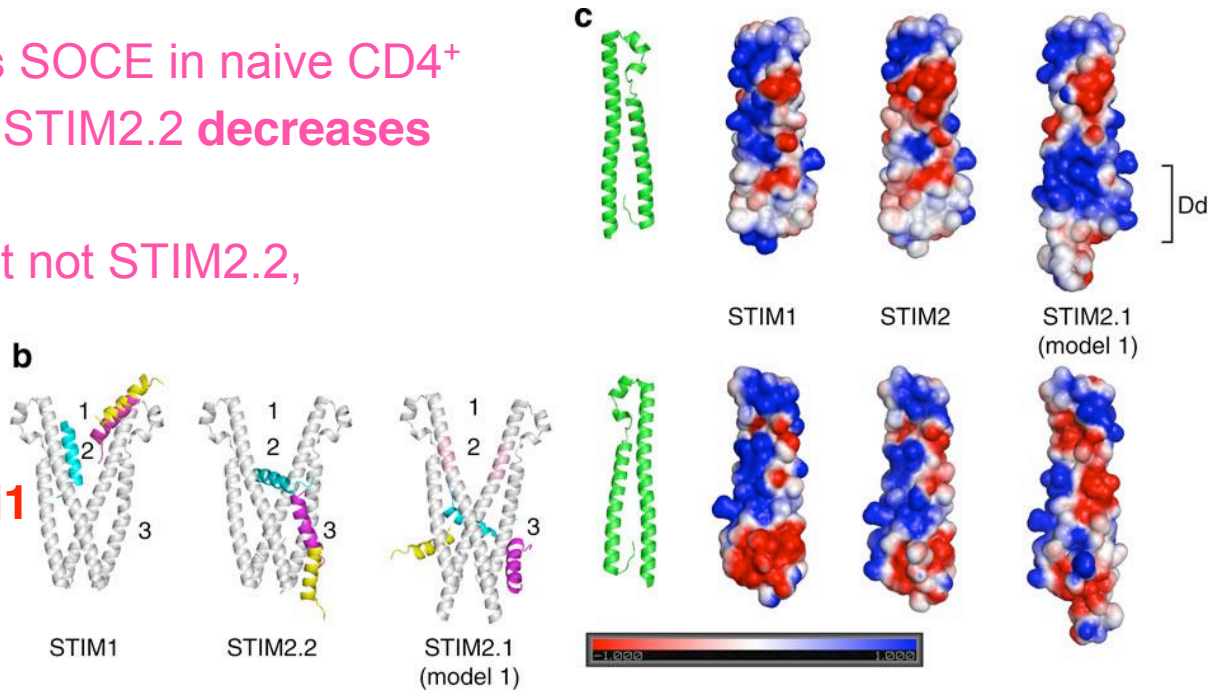
Alternative splicing may affect PP interactions: STIM2 splice variant

STIM proteins regulate **store-operated calcium entry** (SOCE) by sensing Ca^{2+} concentration in the ER and forming oligomers to trigger Ca^{2+} entry through plasma membrane-localized Orai1 channels.

Niemeyer and co-workers characterized a *STIM2* splice variant which retains an additional 8-amino acid exon in the region encoding the channel-activating domain.

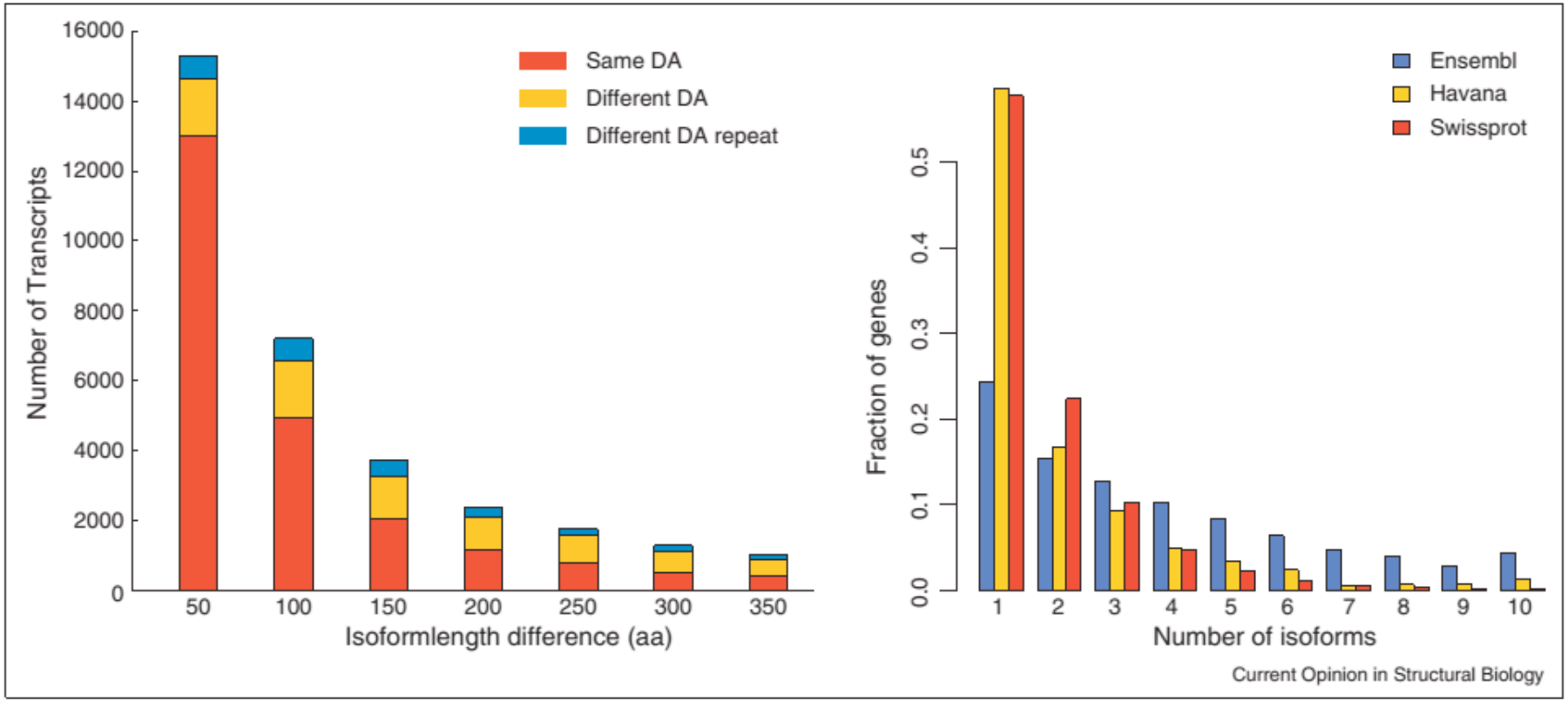
STIM2.1 knockdown **increases** SOCE in naive $CD4^+$ T cells, whereas knockdown of STIM2.2 **decreases** SOCE.
Overexpression of STIM2.1, but not STIM2.2, decreases SOCE.

STIM2.1 interaction with Orai1 is impaired and prevents Orai1 activation.



Miederer, ..., Lee, ..., Helms, Barbara Niemeyer
Nature Commun 6, 6899 (2015)

Effect of AS on protein domain architecture



(left) fraction of proteins where the domain architecture (DA) is altered as a result of splicing (based on Swissprot transcripts)

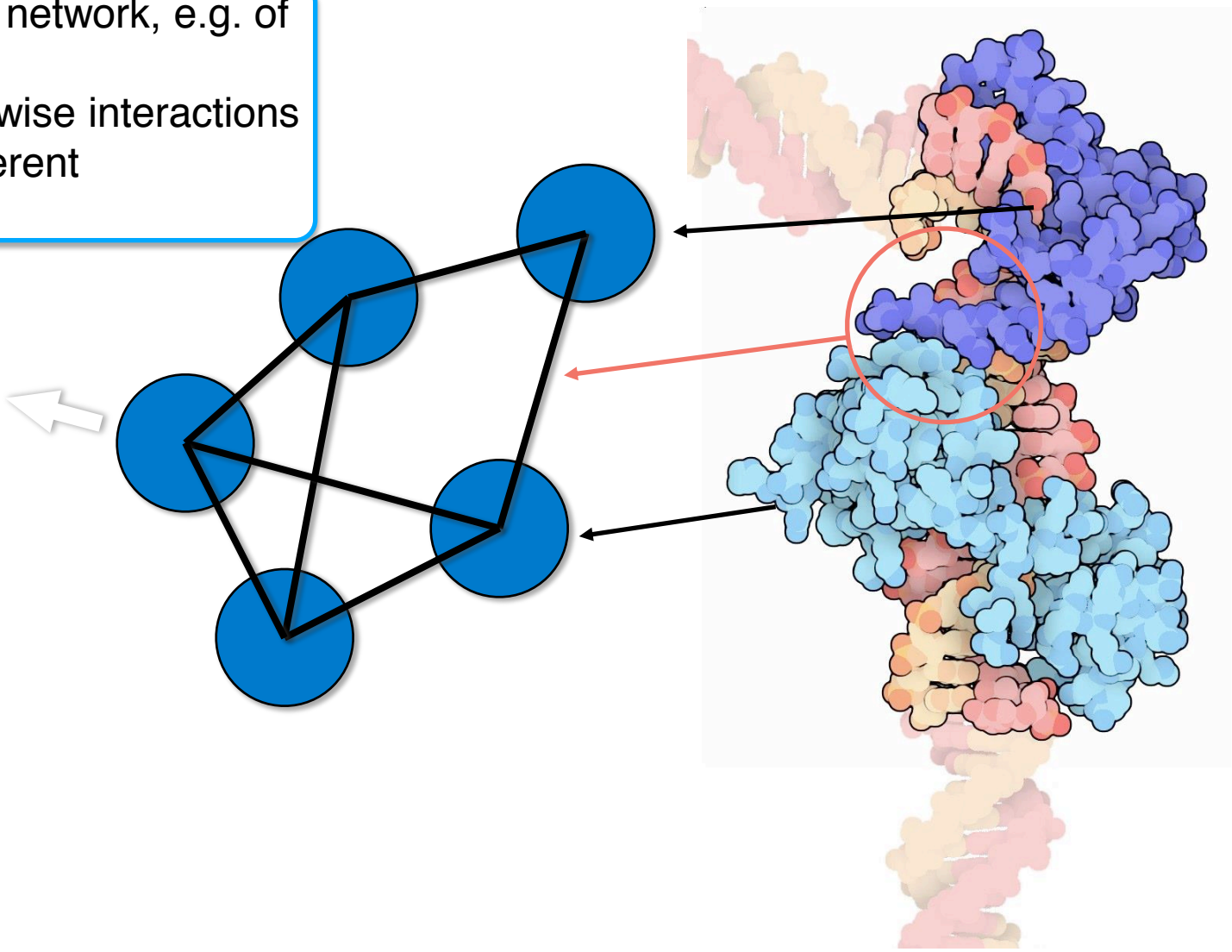
(right) number of isoforms for 3 databases; Ensembl, Vega/ Havana and Swissprot.

Light & Elofsson Curr Opin Struct Biol (2013) 23: 451-458

Toward condition-specific protein interaction networks

Full interaction PP network, e.g. of human
= collection of pairwise interactions compiled from different experiments

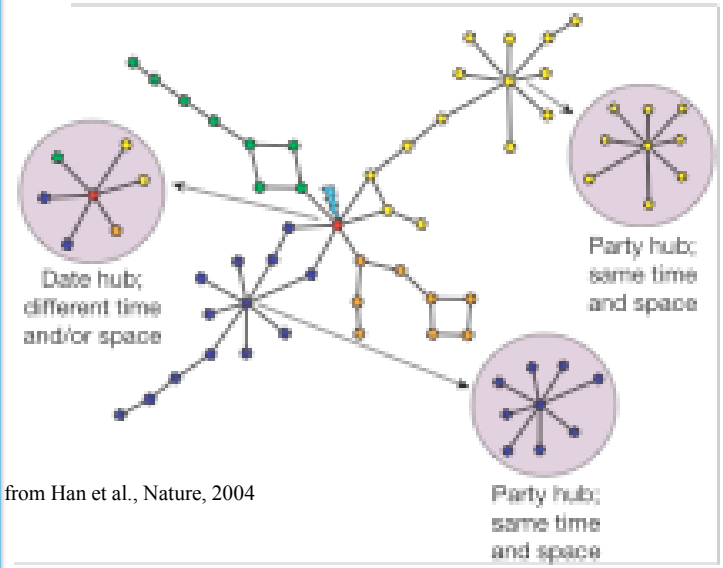
broad range of applications



Oct1/Sox2 from RCSB Protein Data Bank, 2013

But protein interactions can be …

dynamic in time and space

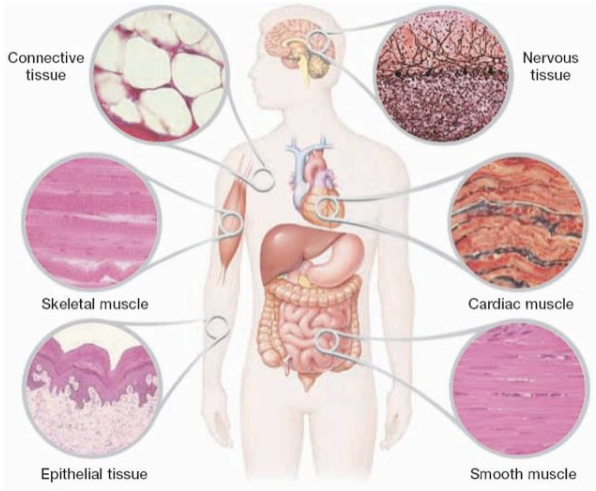


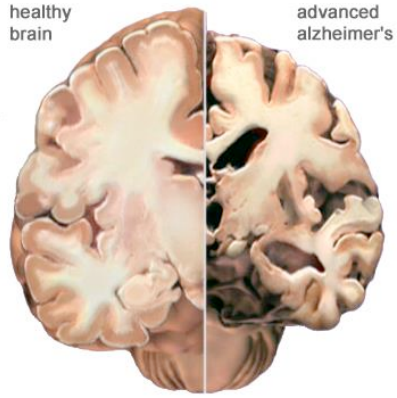
from Han et al., Nature, 2004

same color = similar expression profiles

interaction data itself generally **static**

condition-specific protein composition

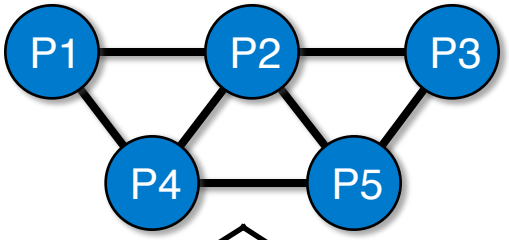




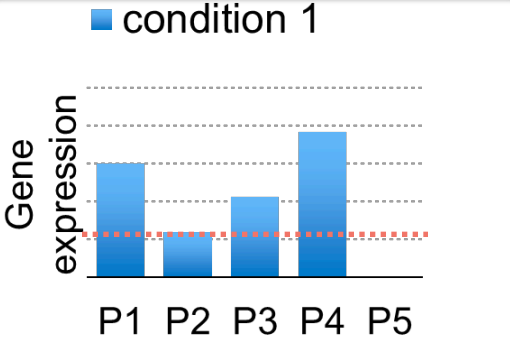
Human tissues from www.pharmaworld.pk
Alzheimer from www.alz.org

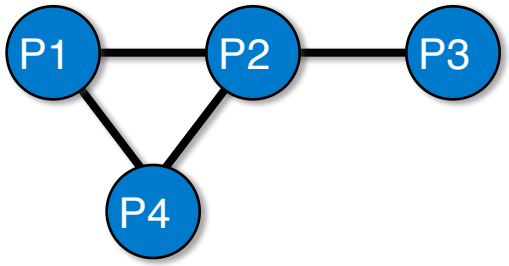
Simple condition-specific PPI networks

database(s) ⋯



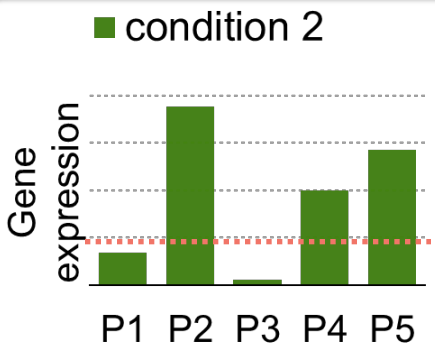
complete protein interaction network

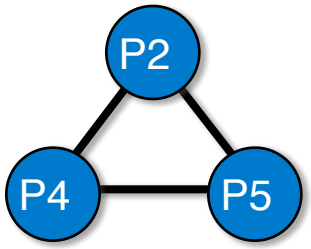




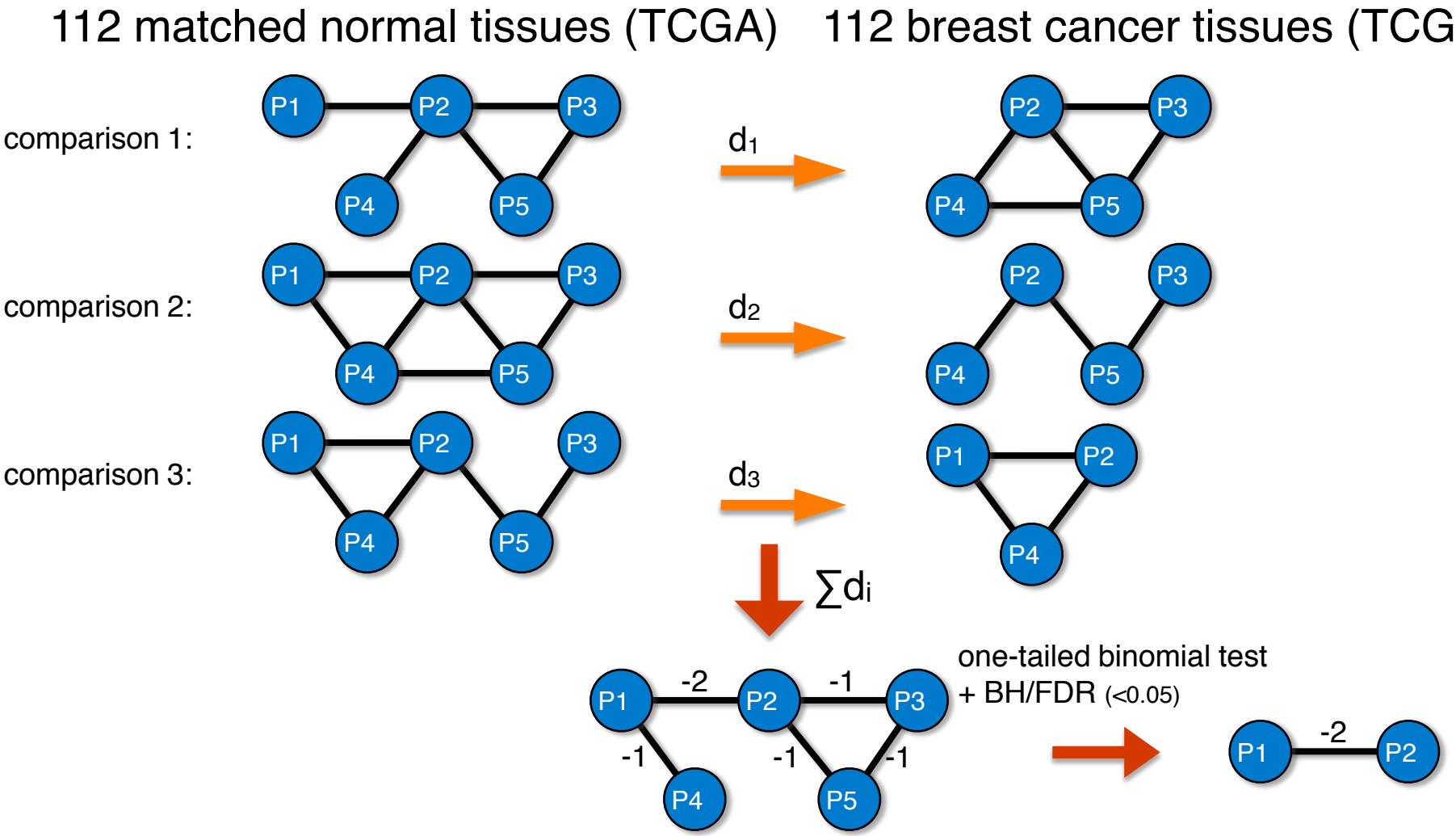
idea:
prune to subset of expressed genes

e.g.:
Bossi and Lehner, Mol. Syst. Bio., 2009
Lopes et al., Bioinformatics, 2011
Barshir et al., PLoS CB, 2014





Differential PPI wiring analysis



Check whether rewiring of a particular PP interaction occurs in a significantly large number of patients compared to what is expected by chance rewiring events.

Coverage of PPIs with domain information

Standard deviations reflect differences betwen patients.

About 10.000 out of 133.000 protein-protein interactions are significantly rewired between
normal and cancer samples.

	GENE
avg. number of proteins (normal)	$12,678 \pm 223$
avg. number of proteins (tumor)	$12,528 \pm 206$
avg. number of interactions (normal)	$134,348 \pm 2,387$
avg. number of interactions (tumor)	$133,128 \pm 2,144$
P_{rew}	0.067 ± 0.016
significantly rewired interactions	9,754

Table S7: *Results obtained using the BioGRID interaction data and using either gene- or various transcript-based network construction approaches. The given numbers denote the sizes of the constructed networks. For all deterministic approaches the standard deviation across all 112 matched samples is shown, for the randomized approach the deviation shown is the average of standard deviations per run. A part of the results for P_{rew} and significantly rewired interactions are also shown in the upper half of Table 3 in the main text.*
Both net loss of proteins and interactions from normal to tumor were significant according to a two-sided Wilcoxon signed-rank test applied to the matched pairs of samples. For the

Will, Helms, Bioinformatics, 47, 219 (2015)
doi: 10.1093/bioinformatics/btv620

Rewired PPIs are associated with hallmarks

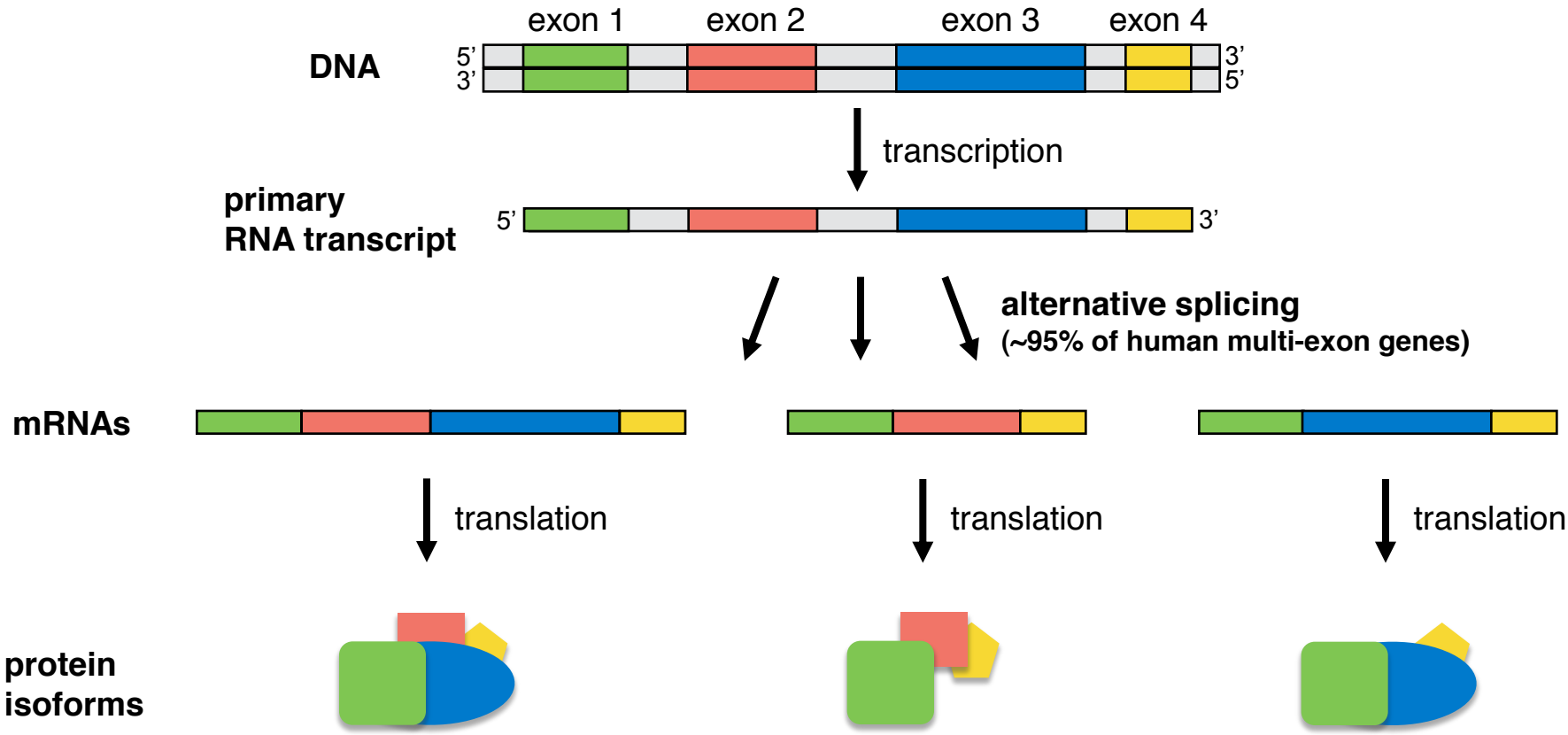
	GENE
rewired interactions	9,754
participation in any hallmark term	7,028
fraction in any hallmark term	0.721
Resisting Cell Death	4,064 (0.417)
Activating Invasion and Metastasis	2,244 (0.230)
Sustaining Proliferative Signaling	3,964 (0.406)
Inducing Angiogenesis	169 (0.017)
Tumor-Promoting Inflammation	516 (0.053)
Genome Instability and Mutation	1,362 (0.140)
Enabling Replicative Immortality	232 (0.024)
Evading Growth Suppressors	3,362 (0.345)
Avoiding Immune Destruction	752 (0.077)
Deregulating Cellular Energetics	821 (0.084)
avg.	1,749 (0.179)

A large fraction (72%) of the rewired interactions affects genes that are associated with „hallmark of cancer“ terms.

Table S10: *Results for the rewiring analysis of the BioGRID network in terms of rewired interactions that affect proteins associated with hallmarks of cancer as defined by [1]. A protein interaction was considered relevant regarding a hallmark term if at least one of its associated proteins was part of the corresponding set of hallmark proteins. The results for individual hallmark terms are reported as the absolute quantity of matches (left number) and as fraction of the total number of rewired interactions listed in the first row (in brackets).*

Will, Helms, Bioinformatics, 47, 219 (2015)
doi: 10.1093/bioinformatics/btv620

Not considered yet: alternative splicing

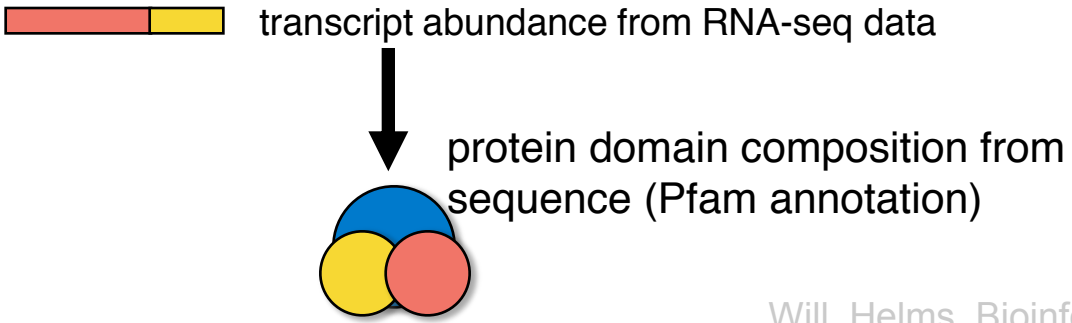


AS affects ability of proteins to interact with other proteins

PPIXpress uses domain information

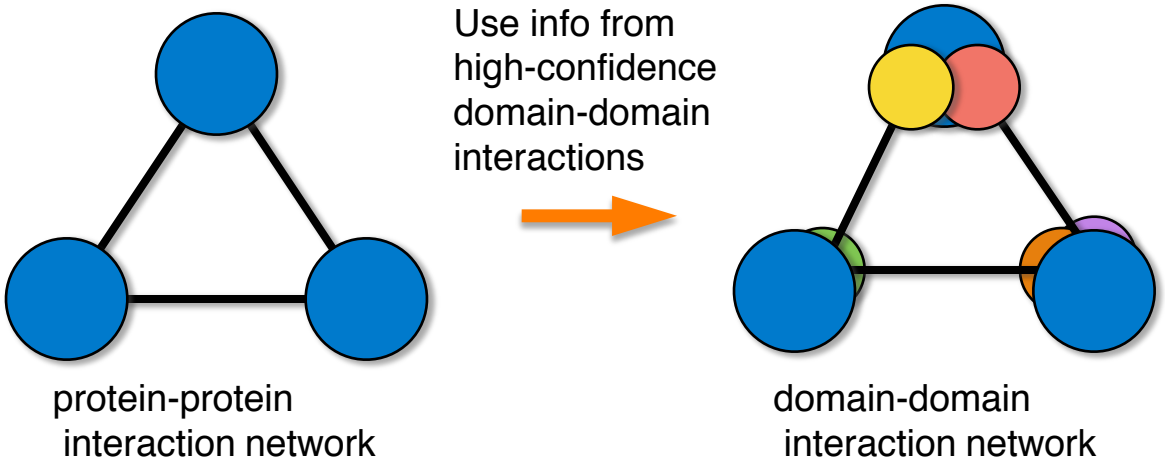
see http://sourceforge.net/projects/ppixpress

I. Determine “building blocks“ for all proteins

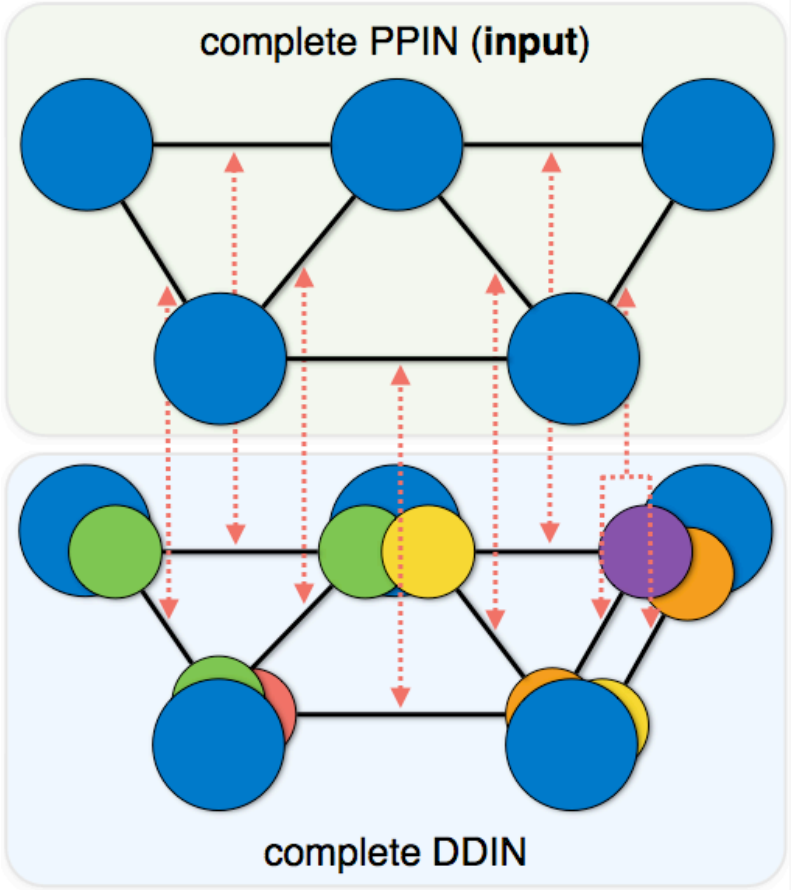


Will, Helms, Bioinformatics, 47, 219 (2015)
doi: 10.1093/bioinformatics/btv620

II. Connect them on the domain-level



PPIXpress method



mapping:

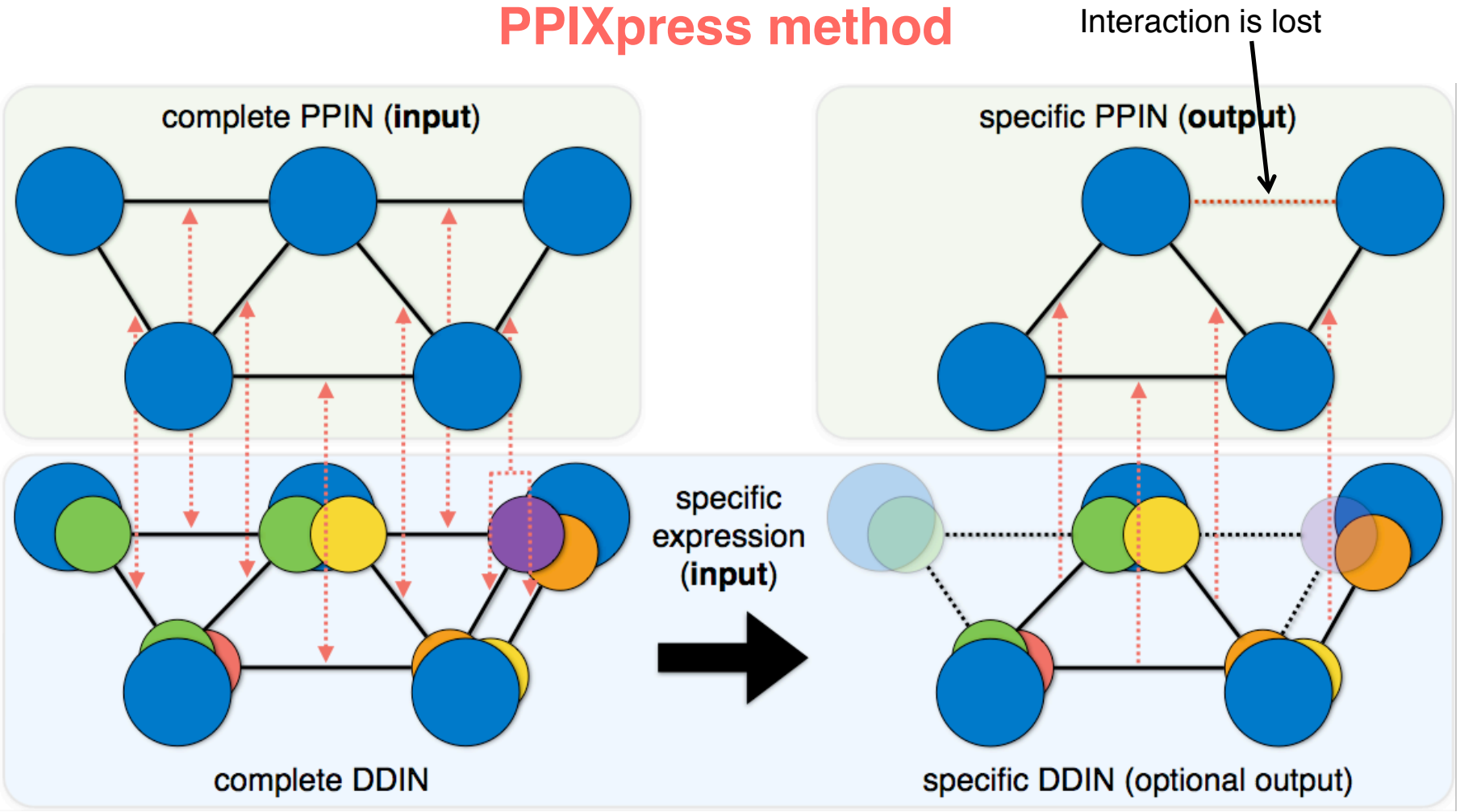
protein-protein interaction

establish
one-to-at-least-one
relationship

domain-domain interaction

reference: principal protein isoforms = longest coding transcript

PPIXpress method



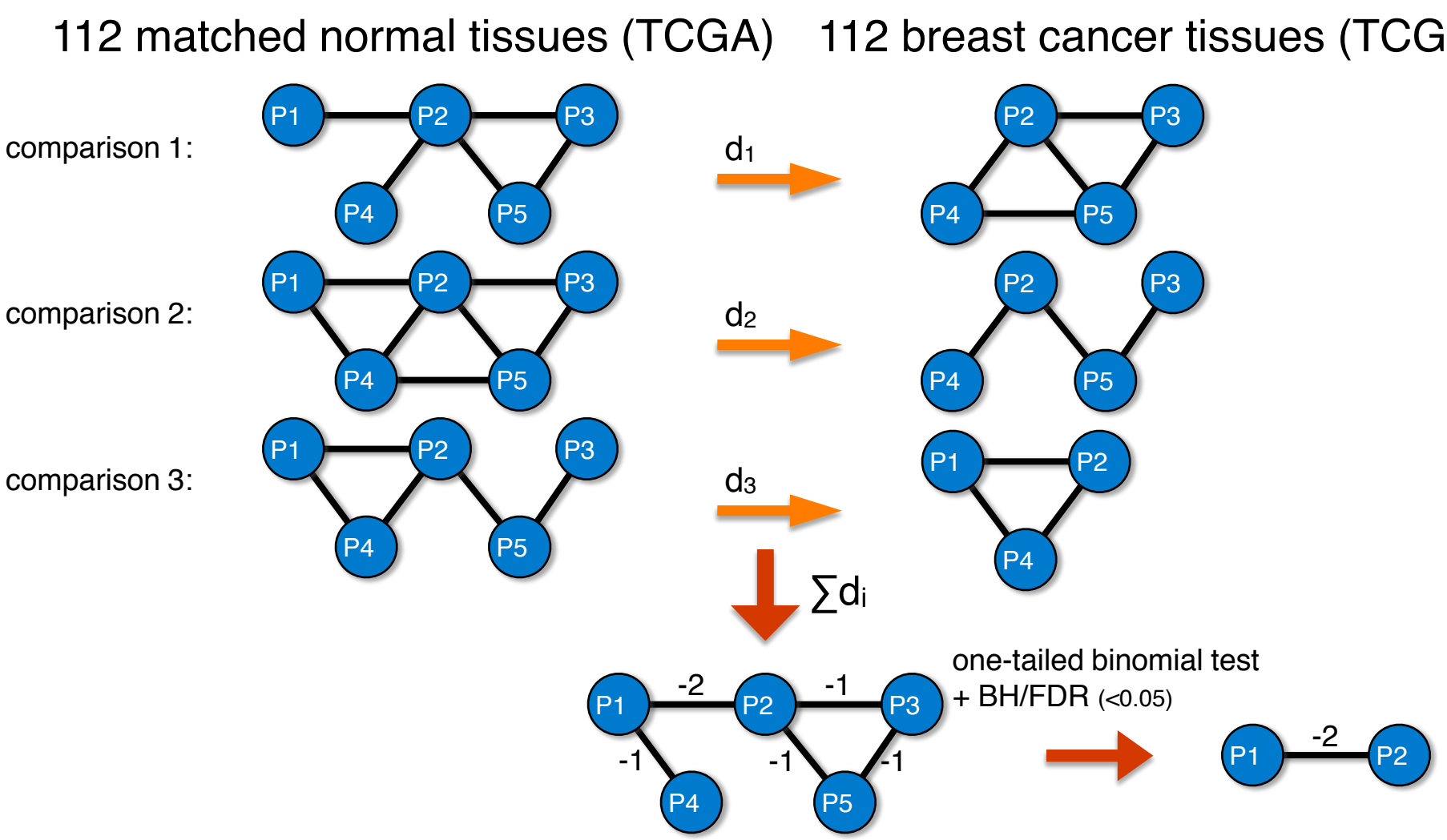
reference: principal protein isoforms

built using most abundant protein isoforms

I. mapping

II. instantiation

Differential PPI wiring analysis at domain level



Coverage of PPIs with domain information

protein set	size of set	fraction of matched PPIs	fraction of contributing proteins
complete network*	15086	0.264	0.517
all HM	4407	0.280	0.684
non HM	10679	0.227	0.449

Domain information is currently available for 51.7% of the proteins of the PP interaction network.

This means that domain information supports about one quarter (26.7%) of all PPIs.

All other PPIs are connected by artificially added domains (1 protein = 1 domain).

Will, Helms, Bioinformatics, 47, 219 (2015)
doi: 10.1093/bioinformatics/btv620

Coverage of PPIs with domain information

	GENE	ALL-DDI
avg. number of proteins (normal)	$12,678 \pm 223$	$12,660 \pm 224$
avg. number of proteins (tumor)	$12,528 \pm 206$	$12,512 \pm 208$
avg. number of interactions (normal)	$134,348 \pm 2,387$	$133,240 \pm 2,451$
avg. number of interactions (tumor)	$133,128 \pm 2,144$	$132,057 \pm 2,196$
P_{rew}	0.067 ± 0.016	0.069 ± 0.017
significantly rewired interactions	9,754	10,111

At domain-level, slightly more (10.111 vs. 9.754) PPIs out of 133.000 PPIs are significantly rewired between normal and cancer samples.

Table S7: *Results obtained using the BioGRID interaction data and using either gene- or various transcript-based network construction approaches. The given numbers denote the sizes of the constructed networks. For all deterministic approaches the standard deviation across all 112 matched samples is shown, for the randomized approach the deviation shown is the average of standard deviations per run. A part of the results for P_{rew} and significantly rewired interactions are also shown in the upper half of Table 3 in the main text.*
Both net loss of proteins and interactions from normal to tumor were significant according to a two-sided Wilcoxon signed-rank test applied to the matched pairs of samples. For the ALL-DDI construction in BioGRID, for example, there were less proteins in the tumor PPINs with $p < 5.9 * 10^{-8}$ (GENE: $p < 3.6 * 10^{-8}$) and less interactions with $p < 3.8 * 10^{-6}$ (GENE: $p < 3.9 * 10^{-6}$), respectively.

Will, Helms, Bioinformatics, 47, 219 (2015)
doi: 10.1093/bioinformatics/btv620

Rewired PPIs are associated with hallmarks

	GENE	ALL-DDI
rewired interactions	9,754	10,111
participation in any hallmark term	7,028	7,343
fraction in any hallmark term	0.721	0.726
Resisting Cell Death	4,064 (0.417)	4,316 (0.427)
Activating Invasion and Metastasis	2,244 (0.230)	2,285 (0.226)
Sustaining Proliferative Signaling	3,964 (0.406)	4,142 (0.410)
Inducing Angiogenesis	169 (0.017)	172 (0.017)
Tumor-Promoting Inflammation	516 (0.053)	537 (0.053)
Genome Instability and Mutation	1,362 (0.140)	1,419 (0.140)
Enabling Replicative Immortality	232 (0.024)	360 (0.036)
Evading Growth Suppressors	3,362 (0.345)	3,557 (0.352)
Avoiding Immune Destruction	752 (0.077)	772 (0.076)
Deregulating Cellular Energetics	821 (0.084)	850 (0.084)
avg.	1,749 (0.179)	1,841 (0.182)

Table S10: *Results for the rewiring analysis of the BioGRID network in terms of rewired interactions that affect proteins associated with hallmarks of cancer as defined by [1]. A protein interaction was considered relevant regarding a hallmark term if at least one of its associated proteins was part of the corresponding set of hallmark proteins. The results for individual hallmark terms are reported as the absolute quantity of matches (left number) and as fraction of the total number of rewired interactions listed in the first row (in brackets).*

The construction at transcript-level also found a larger fraction (72.6 vs 72.1%) of differential interactions that can be associated with hallmark terms than the gene-level based approach.

Will, Helms, Bioinformatics, 47, 219 (2015)
doi: 10.1093/bioinformatics/btv620

Enriched KEGG and GO-BP terms in gene-level \ transcript-level set

	GENE		ALL-DDI	
	term	p	term	p
KEGG	hsa04012:ErbB signaling pathway	0.0013	hsa05200:Pathways in cancer	$1.5 * 10^{-17}$
	hsa05212:Pancreatic cancer	0.0491	hsa04110:Cell cycle	$1.8 * 10^{-15}$
			hsa05220:Chronic myeloid leukemia	$3.5 * 10^{-15}$
			hsa05212:Pancreatic cancer	$1.4 * 10^{-8}$
			hsa05223:Non-small cell lung cancer	$4.3 * 10^{-8}$
GO BP	GO:0007242 intracellular signaling cascade	$6.9 * 10^{-5}$	GO:0010604 positive regulation of macromolecule metabolic process	$4.3 * 10^{-16}$
	GO:0043065 positive regulation of apoptosis	0.0252	GO:0042981 regulation of apoptosis	$3.6 * 10^{-15}$
	GO:0043068 positive regulation of programmed cell death	0.0272	GO:0043067 regulation of programmed cell death	$6.1 * 10^{-15}$
	GO:0010942 positive regulation of cell death	0.0287	GO:0010941 regulation of cell death	$7.7 * 10^{-15}$
	GO:0051329 interphase of mitotic cell cycle	0.0409	GO:0007049 cell cycle	$1.7 * 10^{-14}$

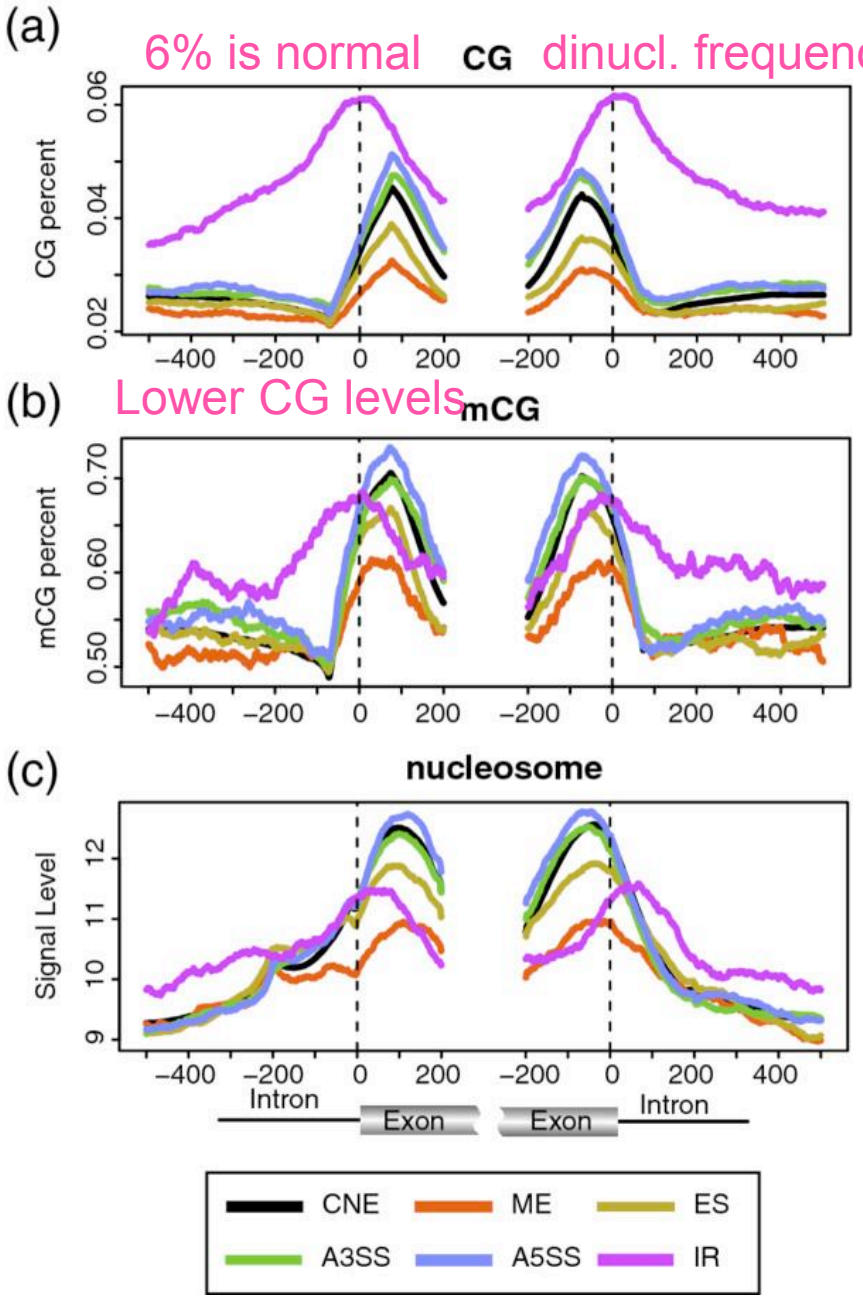
Table S16: *Comparison of rewiring results between the gene-based construction and a transcript-based construction method for the BioGRID network. Here, the top five enriched terms and their p-values are shown for the proteins affected by interactions exclusively found by the transcript-based method using the ALL-DDI dataset or the gene-based approach, respectively. Enrichment in KEGG pathways and GO biological processes was determined using DAVID [2] where we used the proteins included in the corresponding input network as the background. Enrichment was defined as $p < 0.05$ (Bonferroni-adjusted).*

The enriched terms that are exclusively found by the transcript-level method (right) are closely linked to carcinogenetic processes.

Hardly any significant terms are exclusively found at the gene level (left).

Will, Helms, Bioinformatics, 47, 219 (2015)
doi: 10.1093/bioinformatics/btv620

Coupling AS ↔ epigenetic modifications



The association of DNA methylation and nucleosome occupancy with AS.

(a) Distribution of genomic CpG levels around the splice sites of different types of AS events.

(b) Distribution of DNA methylation level (mCG) around the splice sites of different types of AS events.

Both **a** and **b** use a sliding window of 147 bp

(**c**) The distribution of nucleosome occupancy around the splice sites of different types of AS events (ChIP signal, no sliding window)

CNE: constitutively spliced exon (no AS)

ES: exon skipping

ME : mutually exclusive exon

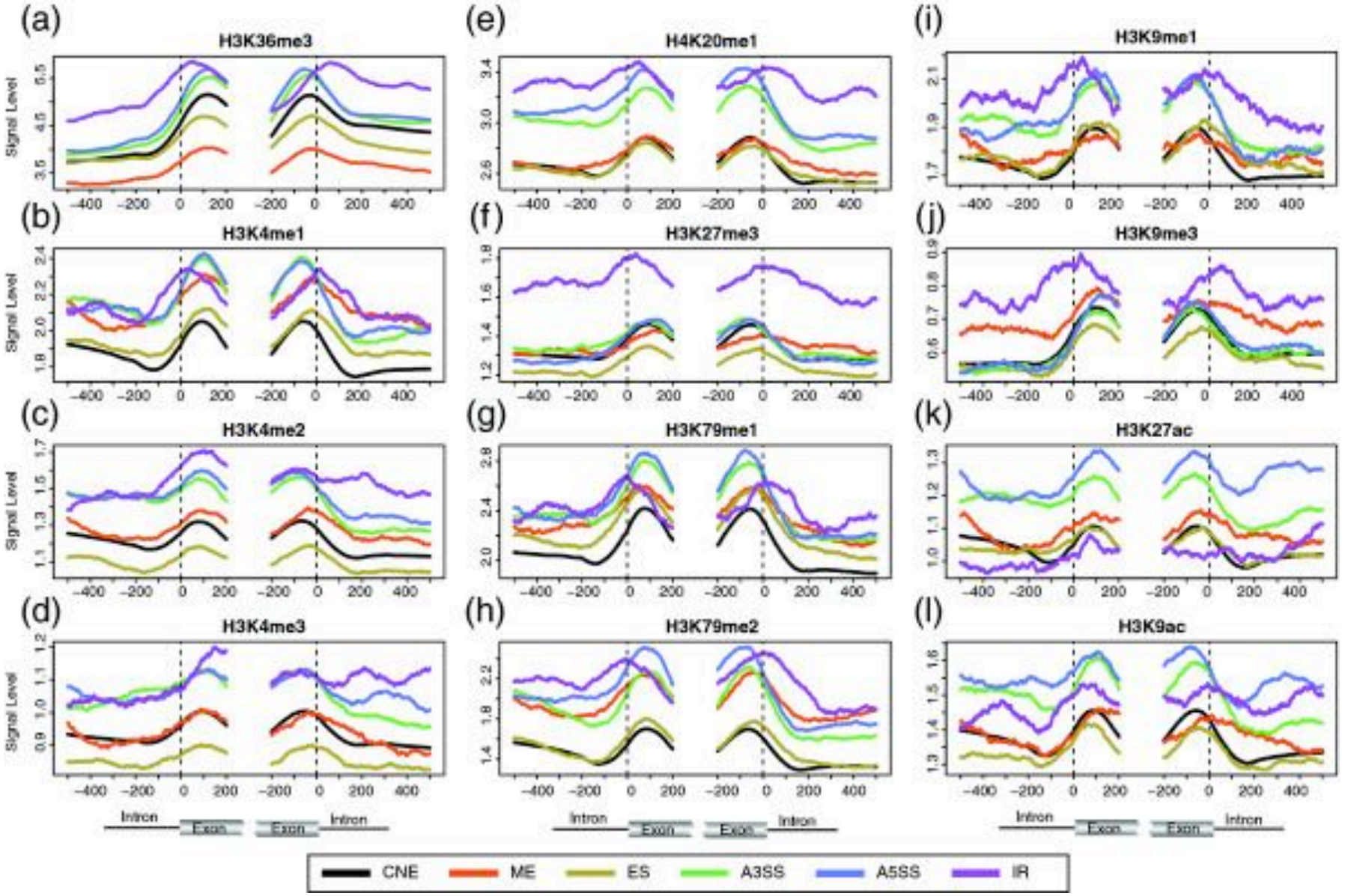
A5SS : alternative 5' splice site selection

A3SS : alternative 3' splice site selection

IR : intron retention.

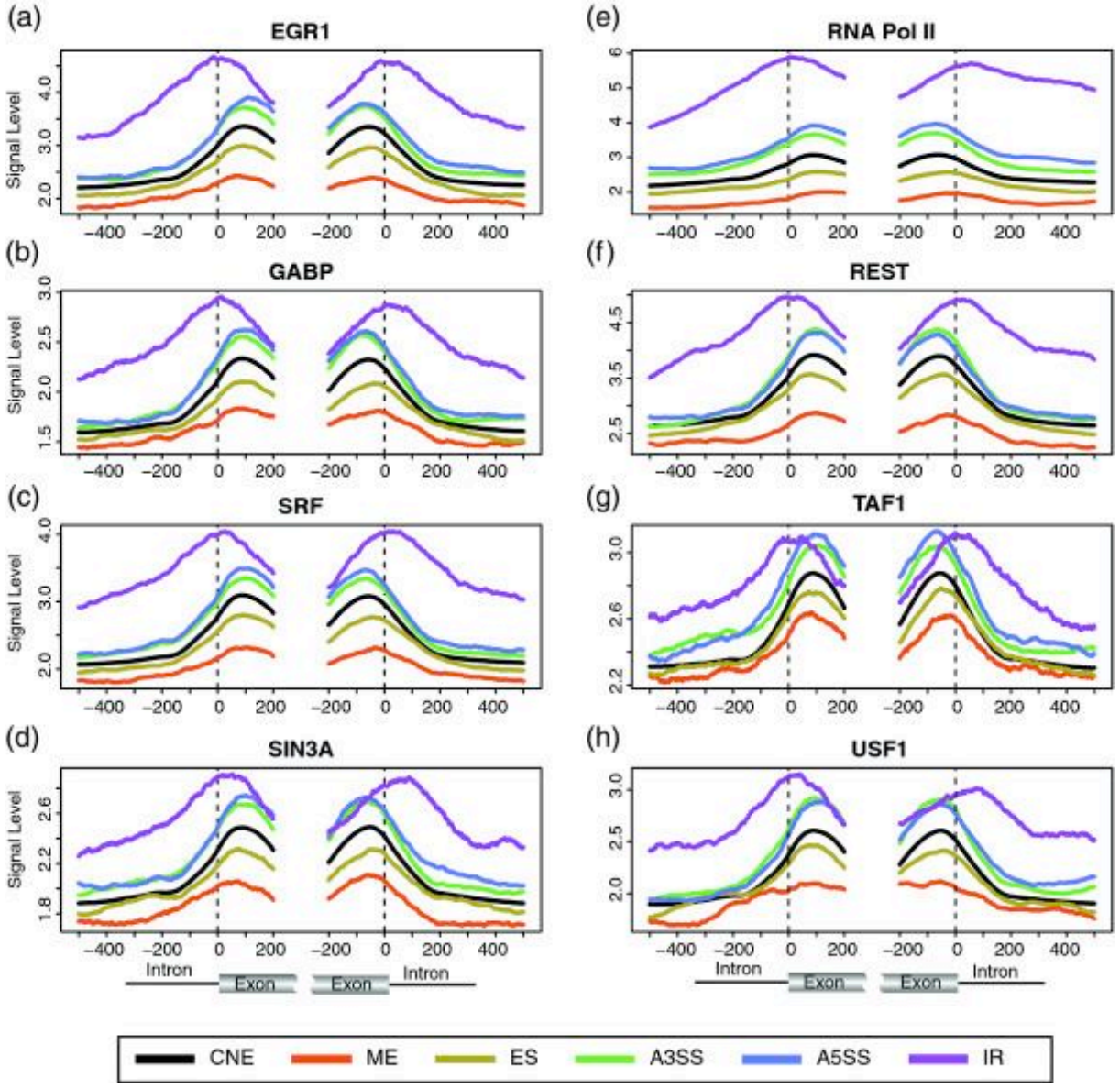
Zhou et al. BMC Genomics (2012) 13:123

Association of histone modification with AS



Zhou et al. BMC Genomics (2012) 13:123

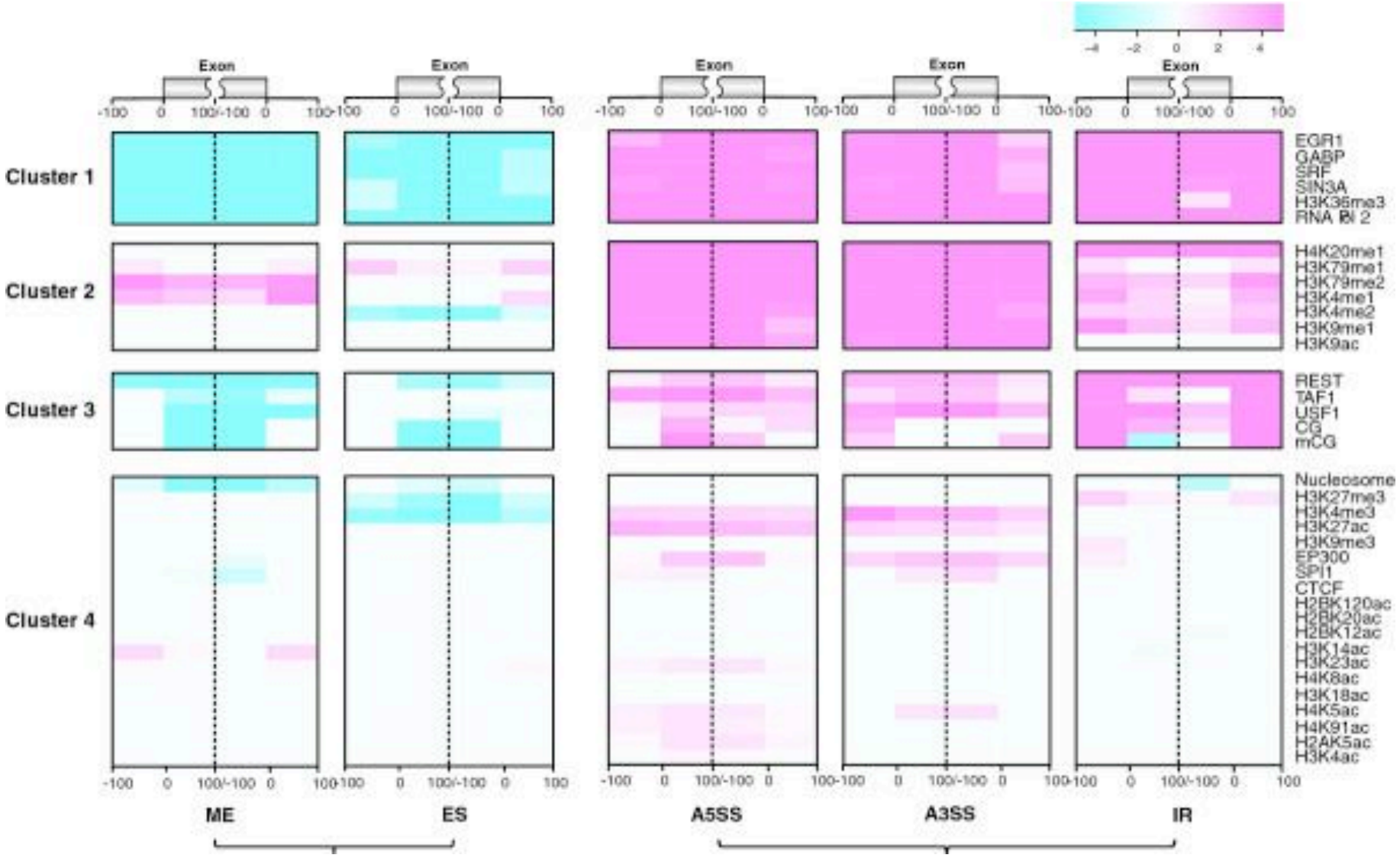
Association of protein features with AS



ChIP-seq data for TF binding.

Zhou et al. BMC Genomics (2012) 13:123

Clustering of associations



Zhou et al. BMC Genomics (2012) 13:123

epigenetic modifications that are associated with AS

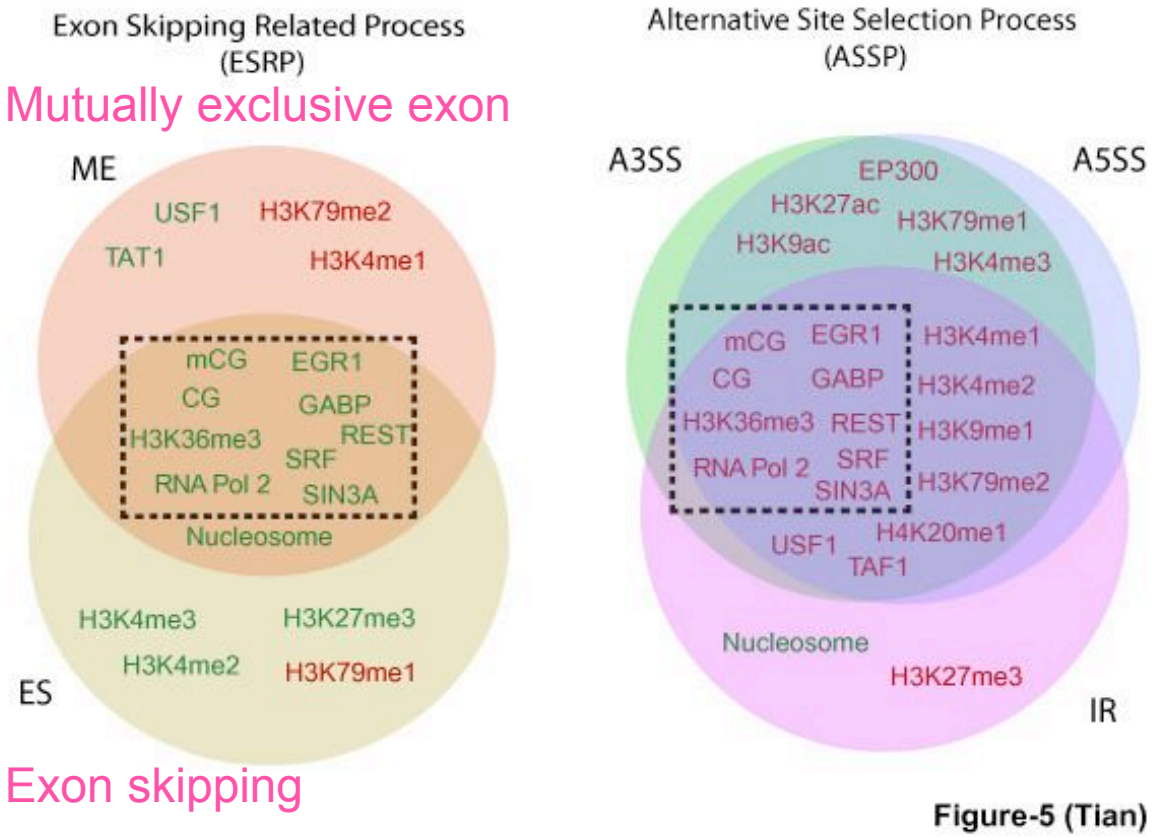


Figure-5 (Tian)

The Epigenetic features strongly associated with different types of AS.

The features showing higher level and lower level in AS events than in CNE are colored in red and green, respectively. The features inside the dashed black box are those common in both ESRP and ASSP; note their association patterns are very different in between ESRP and ASSP.

Zhou et al. BMC Genomics (2012) 13:123

Coupling AS ↔ epigenetic modifications

Epigenetic features are strongly associated with AS.

This suggests that epigenetic regulation may be involved in AS.

Clustering yielded 4 tight clusters of epigenetic features were identified that are associated with AS.

The AS events are grouped to 2 classes: the exon skipping related process (ESRP) (including ME and ES) and the alternative splice site selection process (ASSP) (including A3SS, A5SS and IR) on the basis of their association patterns with epigenetic features.

This indicates that these 2 processes may involve different mechanisms of epigenetic regulation.

Zhou et al. BMC Genomics (2012) 13:123