News Release

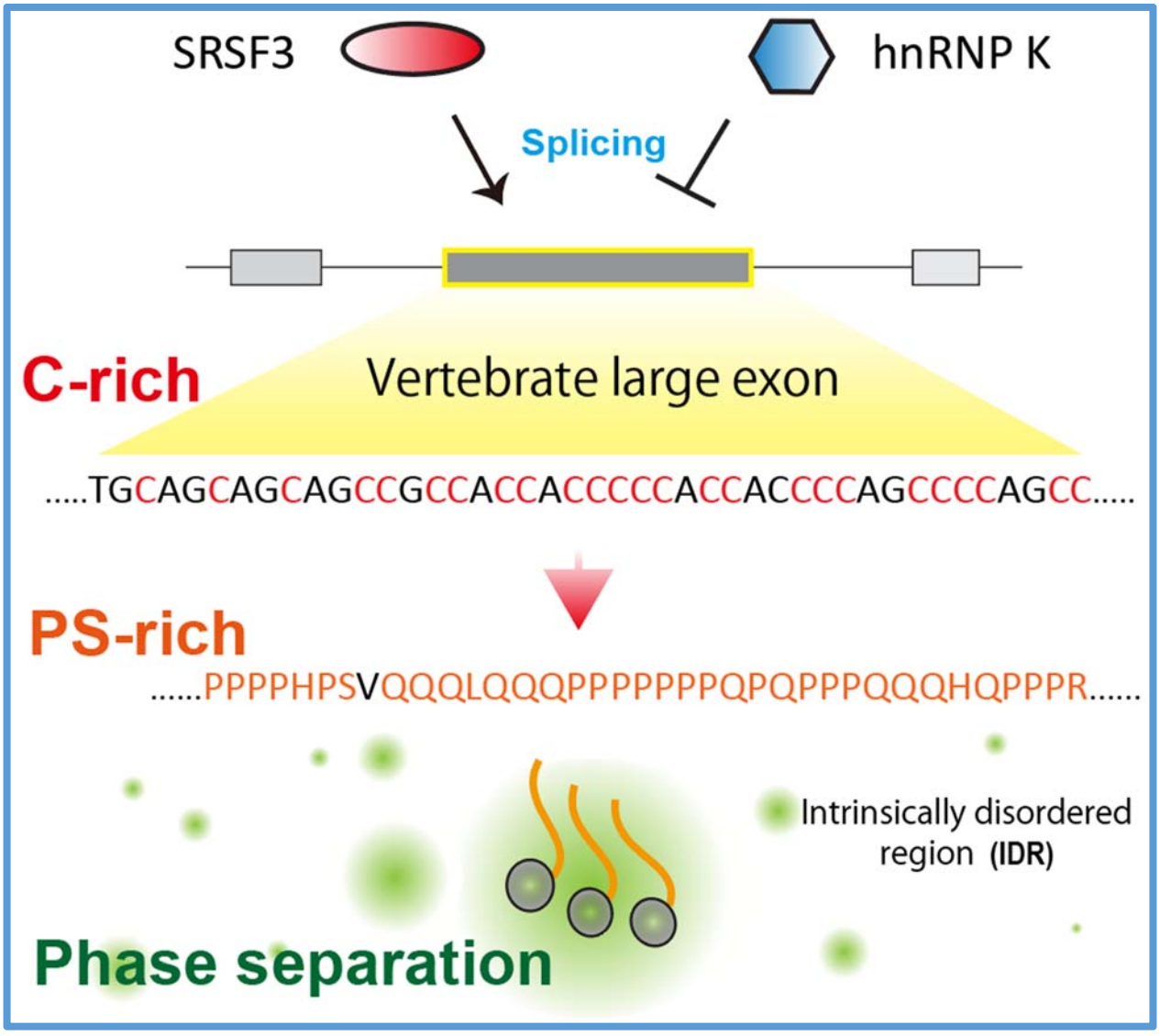
Title

Splicing regulation of large exons secures phase-separation of transcription factors in vertebrates

Key Points

We discovered that large exons extensively encode proline/serine-rich intrinsically disordered regions (IDRs) in transcription factors throughout vertebrates. These large exons are enriched with C-nucleotides to encode proline and serine, which serve as binding sites for SRSF3 and hnRNP K. hnRNP K suppresses the splicing of the large exons, which is masked by SRSF3. The absence of SRSF3 results in the loss of the IDRs of transcription factors and disrupts their assembly. The collective findings clarify the essential role of SRSF3-dependent splicing regulation in sustainable transcription.

- SRSF3 and hnRNP K bind to C-rich elements accumulated in large exons.
- hnRNP K suppresses splicing of the large exons, which is totally overridden by SRSF3 to achieve constitutive splicing of the large exons.
- The SRSF3-dependnet large exons extensively encode IDRs in hundreds of transcription factors and coactivators to form phase-separated assemblies.
- IDR-bearing large exons were evolutionarily acquired in vertebrates.



Summary

Large exons cannot be readily recognized by the spliceosome. Nevertheless, many large exons are evolutionarily conserved and constitutively spliced. Furthermore, the proteomic significance of large exons remains elusive. In this study, we identified a set of ~3,000 SRSF3-dependent large constitutive exons (S3-LCEs). The enriched C-nucleotides in S3-LCEs recruit two splicing factors, hnRNP K and SRSF3. hnRNP K induces the splicing suppression of S3-LCEs, which is mitigated by SRSF3 to achieve constitutive splicing of S3-LCEs. SRSF3 depletion deletes intrinsically disordered regions (IDRs) of transcription factors by skipping S3-LCEs and disrupts their phase-separated assemblies. Enrichment of C-nucleotides in large exons to code for proline and serine in intrinsically disordered regions of transcription factors was evolutionarily acquired in vertebrates. Layered splicing regulation by hnRNP K and SRSF3 secures their proper phase separation of transcription factors in vertebrates.

Research Background

In vertebrates, the lengths of exons have decreased during evolution. In the human genome, the median length of internal exons is 122 nt, and only approximately 5% of the exons exceed 300 nt. Large exonic sizes perturb recognition by spliceosome complexes. Nevertheless, a substantial number of large exons are evolutionarily conserved and constitutively spliced in mammals. Splicing-enhancing mechanism(s) that enable the recognition of large constitutive exons remain unidentified.

Research Results

In this study, we first analyzed over 800 RNA-seq data derived from RBP-depletion experiments and found that SRSF3 affected the splicing of the largest exons among other RBPs (Fig. 1). We named nearly 3,000 such exons as SRSF3-dependent large constitutive exons (S3-LCEs). S3-LCEs were enriched with C-nucleotides, which serve as

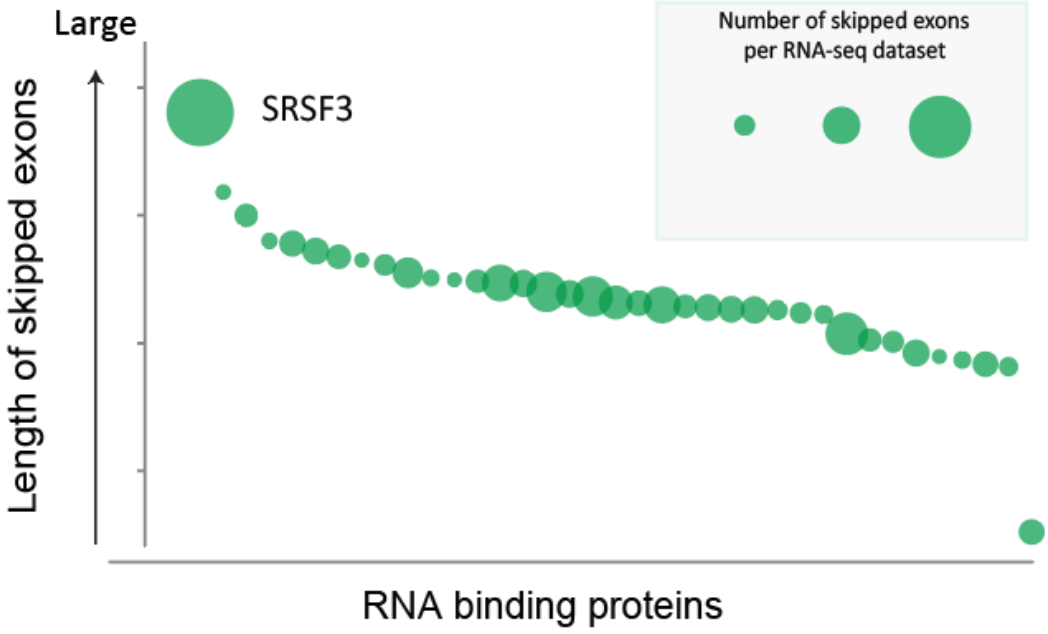


Fig. 1. SRSF3 regulates splicing of large exons.
Median length and number of exons skipped by each RBP depletion. Bubble sizes indicate the average number of skipped exons per dataset.

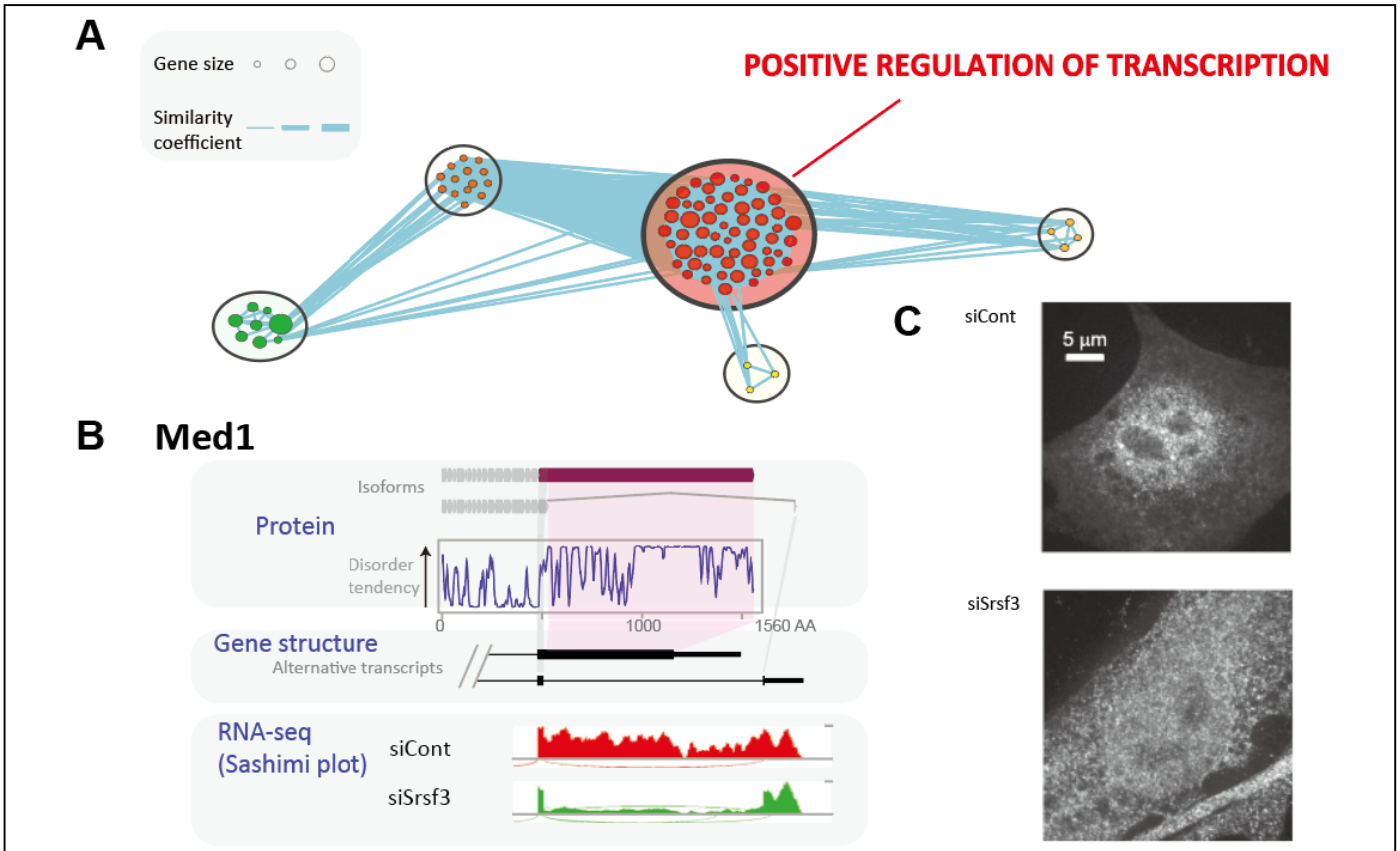


Fig. 2. *Srsf3* silencing disrupts transcription assemblies
(A) Clustering of GO terms enriched in SRSF3-dependent large constitutive exons.
(B) SRSF3-responsive exons of the *Med1* gene. (Top panel) Disorder analysis for MED1.
(C) Immunofluorescence images of MED1 in C2C12 cells treated with siCont or siSrsf3.

binding sites for SRSF3 and hnRNP K. SRSF3 depletion caused the skipping of S3-LCEs, which was rescued by hnRNP depletion. hnRNP K depletion alone had no substantial effect on the splicing of S3-LCEs. Furthermore, S3-LCEs encode the IDRs of transcription factors. Loss of an IDR in these molecules by SRSF3 depletion disrupted the assembly of transcription factors (Fig. 2), including the mediator complex, which leads to cell death. Evolutionary analyses revealed that large exons are extensively enriched in C-nucleotides encoding proline (P) and serine (S), which constitute the IDRs of transcription factors in vertebrates (Fig. 3). These large exons have emerged in vertebrates and rapidly evolved in higher species.

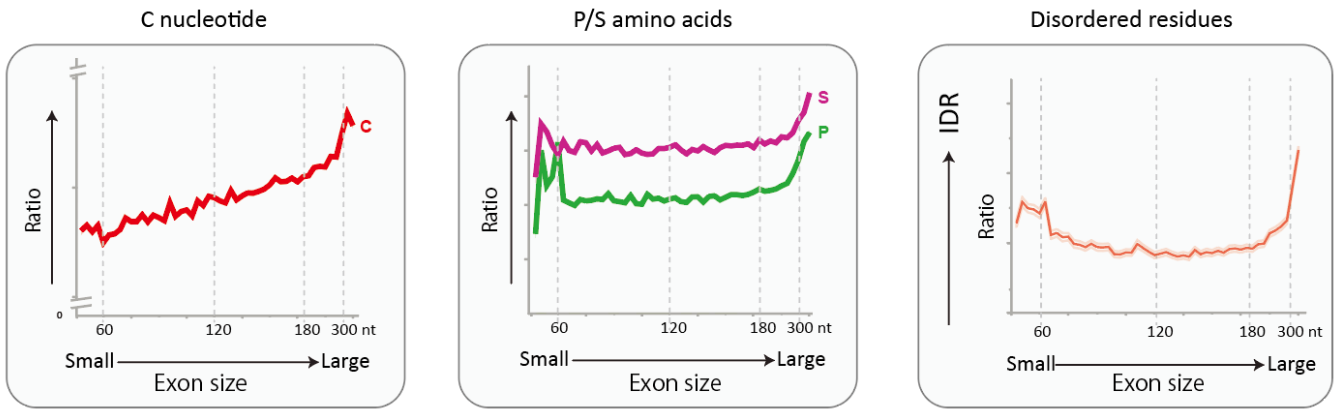


Fig. 3. Ratios of C nucleotide, P/S amino acids, and disordered residues in the human internal exons.

Research Summary and Future Perspective

Phase separation has emerged as a key mechanism for the organization of highly dynamic, membrane-less compartments regulating diverse cellular processes. IDRs play a key role to drive phase separation. Our findings indicate that large exons are extensively enriched with C-nucleotides to encode proline/serine-rich IDRs in transcription factors throughout vertebrates. We demonstrated that the enrichment of C-nucleotides might have unexpectedly recruited hnRNP K, which is extensively overridden by SRSF3. The collective findings clarify the essential role of SRSF3-dependent large exons in sustainable transcription.

During evolution, various sets of IDR-bearing exons might have distinctly developed, having different amino acid preferences and distinct functions. Further analyses will disclose diverse sets of IDR-bearing exons and their roles in evolution and cell homeostasis, in the future.

Publication

Title: Regulated splicing of large exons is linked to phase separation of vertebrate transcription factors

Authors: Toshihiko Kawachi[1,2], Akio Masuda[1,2,*], Yoshihiro Yamashita[1], Jun-ichi Takeda[1], Bisei Ohkawara[1], Mikako Ito[1], Kinji Ohno[1]

Affiliation: [1]Division of Neurogenetics, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine, Nagoya, Japan.

[2]These authors contributed equally.

[*]Corresponding author

Journal：*The EMBO Journal*

DOI: https://www.embopress.org/doi/10.15252/embj.2020107485

Japanese ver. https://www.med.nagoya-u.ac.jp/medical_J/research/pdf/EMBO_211004.pdf