



Genetic variations and alternative splicing: the Glioma-associated oncogene 1, GLI1

Peter G. Zaphiropoulos*

Department of Biosciences and Nutrition, Karolinska Institutet, Huddinge, Sweden

Edited by:

Peng Jiang, University of Iowa, USA

Reviewed by:

Serdar Bozdag, National Institutes of Health, USA

Paola Bonizzoni, Università di Milano-Bicocca, Italy

***Correspondence:**

Peter G. Zaphiropoulos, Department of Biosciences and Nutrition, Karolinska Institutet, 141 83 Huddinge, Sweden.
e-mail: peter.zaphiropoulos@ki.se

Alternative splicing is a post-transcriptional regulatory process that is attaining stronger recognition as a modulator of gene expression. Alternative splicing occurs when the primary RNA transcript is differentially processed into more than one mature RNAs. This is the result of a variable definition/inclusion of the exons, the sequences that are excised from the primary RNA to form the mature RNAs. Consequently, RNA expression can generate a collection of differentially spliced RNAs, which may distinctly influence subsequent biological events, such as protein synthesis or other biomolecular interactions. Still the mechanisms that control exon definition and exon inclusion are not fully clarified. This mini-review highlights advances in this field as well as the impact of single nucleotide polymorphisms in affecting splicing decisions. The Glioma-associated oncogene 1, GLI1, is taken as an example in addressing the role of nucleotide substitutions for splicing regulation.

Keywords: exon skipping, splice signal, splicing enhancer, splicing silencer, cryptic splice site

REGULATION OF SPLICING – *CIS* ELEMENTS, *TRANS* FACTORS

Alternative splicing has, in recent years, been demonstrated to occur in almost all intron-containing human genes (Pan et al., 2008; Wang et al., 2008). However, the regulatory mechanisms that control splice choices are not fully deciphered. Apart from the general problem of defining exonic sequences in a vast excess of intronic DNA (Schwartz et al., 2009a), a key issue concerns the mechanisms of the differential inclusion of exons in mature mRNAs. The prevailing views favor that this is controlled by the relative abundance of a limited number of splicing factors in a particular cell type (Chen and Manley, 2009). Additionally, alternative promoter usage may affect splicing decisions, as this can influence the recruitment of splicing factors or the elongation rate of RNA polymerase II (Kornblihtt, 2005; Muñoz et al., 2009). Furthermore, histone modifications are also thought to have a role in modulating splice choices (Luco et al., 2010). Moreover, the RNA secondary structure may affect splicing decisions (McManus and Graveley, 2011), with this claim being suggested more than 25 years ago (Solnick, 1985). In fact, the expansion of splicing factors, including the SR and hnRNP proteins, in higher eukaryotes (Busch and Hertel, 2012) is in agreement with the increased alternative splicing events in more complex organisms (Kim et al., 2007).

In addition to this *trans* factor dependent regulation of splice choices, *cis* sequences are also important in influencing splicing decisions, as these are thought to act as binding sites for the splicing factors. To start with, exons are flanked by the invariant intronic dinucleotides AG and GT, with the introns also characterized by a polypyrimidine track near their 3' end and the branch point adenosine residue used for lariat formation during intron excision. However, it is obvious that additional *cis* sequence information is required in defining exons and influencing their inclusion in

mRNAs. Such auxiliary sequences, which stimulate splicing are found in both exons (exonic splicing enhancers, ESE), and introns (intronic splicing enhancers, ISE; Aznarez et al., 2008; Lomelin et al., 2010; Brooks et al., 2011). Moreover, sequences that inhibit splicing have also been characterized, as exonic and intronic splicing silencers (ESS and ISS; Yu et al., 2008; Zhang et al., 2008; Wen et al., 2010). Additionally, the sequence complexity of these splicing regulatory elements is highlighted by the *in vivo* analysis of all possible hexamers as exonic splicing regulators, which revealed that out of the 4,096 combinations more than half could act as either ESE or ESS (Ke et al., 2011).

SINGLE NUCLEOTIDE POLYMORPHISMS IN HUMAN GENOMES AND ALTERNATIVE SPLICING

The advances in sequencing technologies have made possible the complete sequence of individual human genomes at a rapid pace (Xia et al., 2012). Consequent is the expansion of the single nucleotide polymorphisms (SNP) entries in the dbSNP database, which currently reaches more than 40,000,000 validated RefSNPs¹. Thus, on the average, a SNP is present in every 75 nucleotides within the three billion human genome sequence. Moreover, since the human protein-coding gene content is estimated to be 2% of the human genome, 800,000 SNPs can potentially influence gene expression and splicing, not to mention variations outside protein-coding genes that could alter gene regulation by mechanisms that are still poorly understood (Visel et al., 2009).

What characterizes the SNPs that modulate splicing? Firstly, exons that are known to be alternatively spliced, in contrast to constitutive exons, appear to have a higher proportion of SNPs that affect exonic splicing regulators. Additionally, ESSs are more frequently modified by SNPs than ESEs (de Souza et al., 2011).

¹http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi

A systematic analysis of the SNPs that are known to affect splicing revealed a number of characteristics, including a predominance of ESS gains relative to ESE losses (Woolfe et al., 2010). These features have been introduced in a web-based tool allowing the prediction of the impact of SNPs on alternative splicing regulation, the Skippy software². In this context it is worth noting that SNPs that affect the strength the 5' splice site, but not the conserved GT dinucleotide, are compensated by other regulatory signals, resulting in minimal changes in splicing, arguing in favor of a certain robustness in maintaining splicing patterns (Lu et al., 2011).

MUTATIONS AND GENE FUNCTION

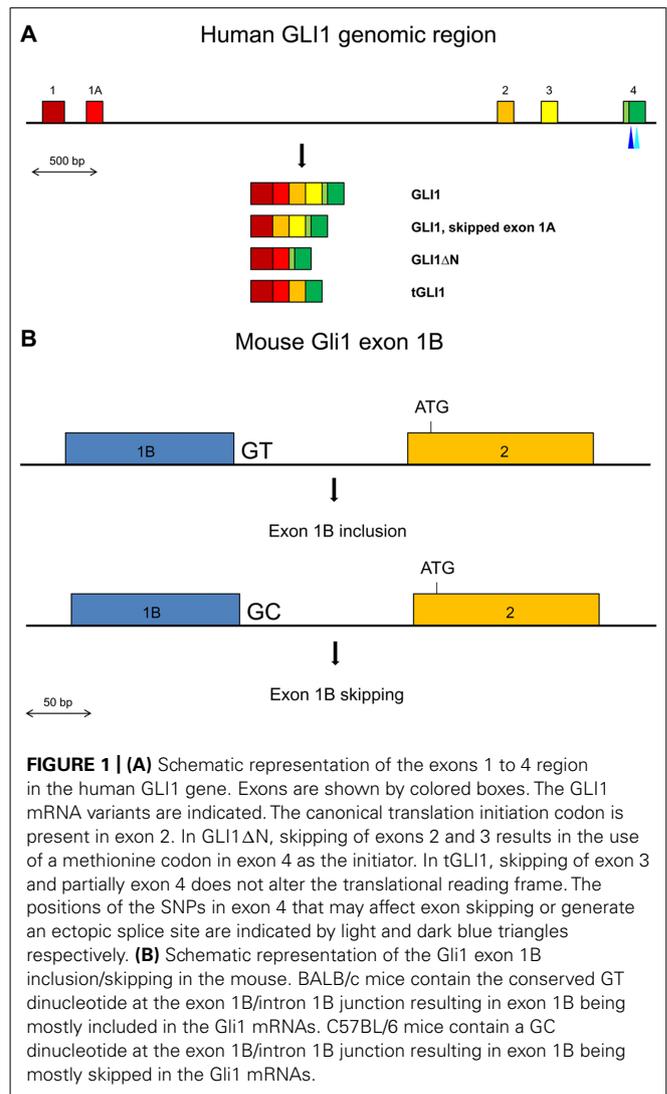
It has generally been assumed that exonic mutations (the term mutation refers to a polymorphism occurring in less than 1% of the population) affect the protein product of the corresponding gene by introducing a stop codon (non-sense mutation), by altering an amino acid (missense mutation) or by changing the translation reading frame (frameshift mutation). However, this postulate had been challenged by the finding that nucleotide substitutions that do not change the encoded amino acid (synonymous mutation) may also alter protein structure, through an indirect way, namely by modulating splicing patterns (Chamary et al., 2006). This can be rationalized by either an activation of a cryptic splice site, as in the case of the LMNA gene (Eriksson et al., 2003), or the presence of ESE and ESS motifs, which are affected by the introduction of a synonymous nucleotide substitution and consequently may alter the canonical splicing pattern. In fact, it has been convincingly argued that the genetic code can effectively carry additional information, such as splicing signals, besides its capacity to encode amino acids (Itzkovitz and Alon, 2007). Moreover, an analysis of 27,681 non-sense and missense disease-causing mutations revealed that about 25% of these actually change exonic splicing regulators, providing evidence that alterations of splicing patterns is not a sole characteristic of synonymous substitutions (Sterne-Weiler et al., 2011).

THE GLIOMA-ASSOCIATED ONCOGENE 1, GLI1 – ALTERNATIVE SPLICING

The GLI1 oncogene encompasses about 12 kb of genomic sequences in chromosome 12q13.2–q13.3 and is composed of two 5' non-coding exons, exons 1 and 1A, and 11 coding exons, exons 2 to 12. Two splice variants have been identified in human tissues that affect coding exons, GLI1 Δ N, which lacks exons 2 and 3 (Shimokawa et al., 2008), and tGLI1, which lacks exon 3 and part of exon 4 (Lo et al., 2009). However, variants that do or do not include the non-coding exon 1A have also been described (Wang and Rothnagel, 2001; **Figure 1A**).

In the mouse, the Gli1 gene contains a third non-coding exon, exon 1B (**Figure 1B**). Interestingly, the 5' splice site of exon 1B in the BALB/c mouse strain contains the canonical GT dinucleotide, while in the C57BL/6 mouse strain this is substituted with GC. The splice signal alteration confers a preferential inclusion of exon 1B in BALB/c mice resulting in a longer 5' UTR region. Moreover, this increased 5' UTR length was found to have a negative impact on the translation efficiency of the Gli1 mRNA, with a

²<http://research.nhgri.nih.gov/skippy>



consequent reduced amount of synthesized protein (Palaniswamy et al., 2010). These observations argue that SNPs that affect splicing may not necessarily alter the primary structure of the protein product but can also have an impact on the amount of produced protein, as a result of changes in the length/structure of the 5' UTR region.

THE GLIOMA-ASSOCIATED ONCOGENE 1, GLI1 – SNPs

The dbSNP database revealed that there are 249 variants in the human GLI1 gene sequence (May 2012). From these 127 are intronic, 32 are synonymous, 59 are non-synonymous, and 14 are in the UTR regions.

By centering on the genomic region that engages in alternative splicing of protein-coding exons, namely exons 2 to 4 and the 100 bases flanking sequence, 27 SNPs, with four being synonymous and four non-synonymous, were identified. Moreover, the use of the Skippy software revealed that the highest score for exon skipping of these eight exonic SNPs is for the C to T missense rs143548857 in exon 4. Direct experimentation will be required

to examine whether this predicted impact on exon 4 splicing is indeed taking place, as this will introduce a translational frameshift.

From the 36 synonymous and non-synonymous SNPs in the remaining protein-coding internal exons, namely exons 5 to 11, only the C to T synonymous rs7973381 in exon 8 has an equally high score for exon exclusion as rs143548857. Interestingly, the size of exon 8 is 150 nucleotides, highlighting that potential exon skipping will not initiate non-sense mediated decay (Durand and Lykke-Andersen, 2011).

Since in the natural tGLI1 variant an alternative 3' splice site is used in exon 4 (Figure 1A), SNPs that could potentially mediate activation of ectopic splice sites in exon 4 were examined using the Skippy software. Interestingly, the G to A synonymous rs139988338 in exon 4 creates a 3' ectopic splice site with a high score. Moreover, none of the SNPs in the remaining internal exons of GLI1 create a 3' ectopic splice site that has a higher score. Worth noting is that the strength of the rs139988338-mediated 3' ectopic splice site in exon 4 is, in fact, higher than the natural tGLI1 alternative 3' splice site, according to the MaxEnt splice site scoring program³.

Additionally, in the non-coding but alternatively spliced exon 1A three SNPs are identified, rs10783826, rs10783827, and rs118093490. These SNPs cannot effectively be evaluated by the Skippy software, as the program is not optimized for non-coding exons (Laura Elnitski, personal communication). Moreover, exon 1A is not fully annotated in the human Genome Build 37, which automatically excludes it from the software analysis, since only annotated exons can be used. On the other hand, manual inspection of the elicited base substitution by the three SNPs and taking advantage of the published strength of all hexamers as ESEs and ESSs (Ke et al., 2011) revealed that for rs10783826 and rs118093490 the nucleotide change increased the impact of the corresponding hexamer as a silencer, while for rs10783827 the change was in favor of an enhancer. These findings were observed irrespective of

whether the hexamers were shifted by one base so the SNP was placed at either the third or the fourth nucleotide position within the hexamer. Moreover, the SROOGLE software⁴ (Schwartz et al., 2009b), which uses additional programs to identify ESEs and ESSs, revealed a similar trend. Thus, it is possible that the differential inclusion of exon 1A may be dependent of these SNPs, and the resulting variable length of the 5' UTR region could have an impact on the amount of the translated protein product, as previously seen with the mouse Gli1.

CONCLUSION

Alternative splicing is a major post-transcriptional mechanism that diversifies gene expression. In addition to *trans* acting factors, *cis* regulatory sequences have also an impact in controlling splice choices, highlighting the potential significance of SNPs in modulating splicing decisions. SNPs that affect splicing may be present in both intronic and exonic regions, mapping at regulatory elements such as splicing enhancers and silencers. Moreover, for exonic SNPs, not only synonymous but also missense and non-sense SNPs can influence splice choices, providing evidence that the major biological impact of some of these exon variants may be on mRNA splicing rather than on protein synthesis.

Using the Glioma-associated oncogene 1, GLI1, as an example, SNPs are predicted to have a role in controlling splicing patterns. Interestingly, GLI1 exons that are known to engage in alternative splicing appear to preferentially encompass SNPs that may affect splice choices. Finally, although user-friendly prediction programs simplify the analysis of the role of SNPs in splicing regulation, the requirement of direct experimentation has not been fully substituted.

ACKNOWLEDGMENTS

This work has been supported by the Swedish Childhood Cancer Foundation, the Swedish Cancer Fund, and the Swedish Research Council.

³http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq_acc.html

⁴<http://sroogle.tau.ac.il>

REFERENCES

- Aznarez, I., Barash, Y., Shai, O., He, D., Zielenski, J., Tsui, L. C., Parkinson, J., Frey, B. J., Rommens, J. M., and Blencowe, B. J. (2008). A systematic analysis of intronic sequences downstream of 5' splice sites reveals a widespread role for U-rich motifs and TIA1/TIAL1 proteins in alternative splicing regulation. *Genome Res.* 18, 1247–1258.
- Brooks, A. N., Aspden, J. L., Podgoraia, A. I., Rio, D. C., and Brenner, S. E. (2011). Identification and experimental validation of splicing regulatory elements in *Drosophila melanogaster* reveals functionally conserved splicing enhancers in metazoans. *RNA* 17, 1884–1894.
- Busch, A., and Hertel, K. J. (2012). Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdiscip. Rev. RNA* 3, 1–12.
- Chamary, J. V., Parmley, J. L., and Hurst, L. D. (2006). Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nat. Rev. Genet.* 7, 98–108.
- Chen, M., and Manley, J. L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* 10, 741–754.
- de Souza, J. E., Ramalho, R. F., Galante, P. A., Meyer, D., and de Souza, S. J. (2011). Alternative splicing and genetic diversity: silencers are more frequently modified by SNVs associated with alternative exon/intron borders. *Nucleic Acids Res.* 39, 4942–4948.
- Durand, S., and Lykke-Andersen, J. (2011). SnapShot: nonsense-mediated mRNA decay. *Cell* 145, 324–324.e2.
- Eriksson, M., Brown, W. T., Gordon, L. B., Glynn, M. W., Singer, J., Scott, L., Erdos, M. R., Robbins, C. M., Moses, T. Y., Berglund, P., Dutra, A., Pak, E., Durkin, S., Csoka, A. B., Boehnke, M., Glover, T. W., and Collins, F. S. (2003). Recurrent de novo point mutations in lamin A cause Hutchinson–Gilford progeria syndrome. *Nature* 423, 293–298.
- Iitzkovitz, S., and Alon, U. (2007). The genetic code is nearly optimal for allowing additional information within protein-coding sequences. *Genome Res.* 17, 405–412.
- Ke, S., Shang, S., Kalachikov, S. M., Morozova, I., Yu, L., Russo, J. J., Ju, J., and Chasin, L. A. (2011). Quantitative evaluation of all hexamers as exonic splicing elements. *Genome Res.* 21, 1360–1374.
- Kim, E., Magen, A., and Ast, G. (2007). Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res.* 35, 125–131.
- Kornblihtt, A. R. (2005). Promoter usage and alternative splicing. *Curr. Opin. Cell Biol.* 17, 262–268.
- Lo, H. W., Zhu, H., Cao, X., Aldrich, A., and Ali-Osman, F. (2009). A novel splice variant of GLI1 that promotes glioblastoma cell migration and invasion. *Cancer Res.* 69, 6790–6798.
- Lomelin, D., Jorgenson, E., and Risch, N. (2010). Human genetic variation recognizes functional elements in noncoding sequence. *Genome Res.* 20, 311–319.

- Lu, Z. X., Jiang, P., Cai, J. J., and Xing, Y. (2011). Context-dependent robustness to 5' splice site polymorphisms in human populations. *Hum. Mol. Genet.* 20, 1084–1096.
- Luco, R. F., Pan, Q., Tominaga, K., Blencowe, B. J., Pereira-Smith, O. M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science* 327, 996–1000.
- McManus, C. J., and Graveley, B. R. (2011). RNA structure and the mechanisms of alternative splicing. *Curr. Opin. Genet. Dev.* 21, 373–379.
- Muñoz, M. J., Pérez Santangelo, M. S., Paronetto, M. P., de la Mata, M., Pelisch, F., Boireau, S., Glover-Cutter, K., Ben-Dov, C., Blaustein, M., Lozano, J. J., Bird, G., Bentley, D., Bertrand, E., and Kornblihtt, A. R. (2009). DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation. *Cell* 137, 708–720.
- Palaniswamy, R., Teglund, S., Lauth, M., Zaphiropoulos, P. G., and Shimokawa, T. (2010). Genetic variations regulate alternative splicing in the 5' untranslated regions of the mouse glioma-associated oncogene 1, Gli1. *BMC Mol. Biol.* 11, 32. doi: 10.1186/1471-2199-11-32
- Pan, Q., Shai, O., Lee, L. J., Frey, B. J., and Blencowe, B. J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* 40, 1413–1415.
- Schwartz, S., Hall, E., and Ast, G. (2009b). SROOGLE: webserver for integrative, user-friendly visualization of splicing signals. *Nucleic Acids Res.* 37, W189–W192.
- Schwartz, S., Meshorer, E., and Ast, G. (2009a). Chromatin organization marks exon–intron structure. *Nat. Struct. Mol. Biol.* 16, 990–995.
- Shimokawa, T., Tostar, U., Lauth, M., Palaniswamy, R., Kasper, M., Toftgård, R., and Zaphiropoulos, P. G. (2008). Novel human glioma-associated oncogene 1 (GLI1) splice variants reveal distinct mechanisms in the terminal transduction of the hedgehog signal. *J. Biol. Chem.* 283, 14345–14354.
- Solnick, D. (1985). Alternative splicing caused by RNA secondary structure. *Cell* 43, 667–676.
- Sterne-Weiler, T., Howard, J., Mort, M., Cooper, D. N., and Sanford, J. R. (2011). Loss of exon identity is a common mechanism of human inherited disease. *Genome Res.* 21, 1563–1571.
- Visel, A., Rubin, E. M., and Pennacchio, L. A. (2009). Genomic views of distant-acting enhancers. *Nature* 461, 199–205.
- Wang, E. T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S. E., Schroth, G. P., and Burge, C. B. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476.
- Wang, X. Q., and Rothnagel, J. A. (2001). Post-transcriptional regulation of the gli1 oncogene by the expression of alternative 5' untranslated regions. *J. Biol. Chem.* 276, 1311–1316.
- Wen, J., Chiba, A., and Cai, X. (2010). Computational identification of tissue-specific alternative splicing elements in mouse genes from RNA-Seq. *Nucleic Acids Res.* 38, 7895–7907.
- Woolfe, A., Mullikin, J. C., and Elnitski, L. (2010). Genomic features defining exonic variants that modulate splicing. *Genome Biol.* 11, R20.
- Xia, J., Wang, Q., Jia, P., Wang, B., Pao, W., and Zhao, Z. (2012). NGS catalog: a database of next generation sequencing studies in humans. *Hum. Mutat.* doi: 10.1002/humu.22096 [Epub ahead of print].
- Yu, Y., Maroney, P. A., Denker, J. A., Zhang, X. H., Dybkov, O., Lührmann, R., Jankowsky, E., Chasin, L. A., and Nilsen, T. W. (2008). Dynamic regulation of alternative splicing by silencers that modulate 5' splice site competition. *Cell* 135, 1224–1236.
- Zhang, C., Li, W. H., Krainer, A. R., and Zhang, M. Q. (2008). RNA landscape of evolution for optimal exon and intron discrimination. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5797–5802.

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 06 May 2012; paper pending published: 29 May 2012; accepted: 13 June 2012; published online: 06 July 2012.

Citation: Zaphiropoulos PG (2012) Genetic variations and alternative splicing: the Glioma-associated oncogene 1, GLI1. *Front. Genet.* 3:119. doi: 10.3389/fgene.2012.00119

This article was submitted to *Frontiers in Bioinformatics and Computational Biology*, a specialty of *Frontiers in Genetics*. Copyright © 2012 Zaphiropoulos. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.