Indian Journal of Science and Technology http://www.indjst.org Vol.1 No 6 (Nov. 2008)

Alternative splicing: a perspective of the nematode model-*Caenorhabditis elegans.*

Luv Kashyap
Pittsburgh Institute for Neurodegenerative Diseases, University of Pittsburgh, Pittsburgh, PA 15260.
luvkashyap@gmail.com

Abstract: The nematode *Caenorhabditis elegans* is an important, well-studied organism used in biomedical research as a model for human development, genetics, ageing, and disease. This paper reviews some basic information, recent progress and future challenges for functional genomics and bioinformatics in the area of alternative splicing. It also highlights the unique methodology adopted by us to study and to detect novel alternatively spliced transcripst from *C. elegans* genome.

Keywords: Alternative splicing, *Caenorhabditis elegans,* gene/exon prediction, genome analysis.

Introduction

The human genome project started a paradigm shift in biomedical research i.e., from an observational science to an information science. Bioinformatics, dedicated to the analysis and management of the vast amount of data became increasingly essential. The human genome project (Venter *et al.,* 2001; International Human Genome Sequencing Consortium, 2001) also brought out the relatively obscure field of alternative splicing research into the spotlight. As widely anticipated that the human genome would contain a much larger number of genes (estimates based on expressed-sequence clustering ran as high as 150,000 genes) than *Drosophila* (14,000 genes) or *Caenorhabditis elegans* (19,000 genes) the report of only 32,000 human genes thus came as a surprise (Claverie, 2001). This basic disparity indicated that the number of human expressed-sequences (mRNA) was much higher than the number of genes. How can the much greater size and complexity of humans be encoded in only twice the number of genes required by a fly?

One way to explain this paradox is to point out that the number of possible proteins from the genome can far exceed the possible number of genes (if a large percentage of the genes have the ability to encode multiple proteins). This expansion of the proteome can be accomplished through alternative precursor messenger RNA (pre mRNA) splicing, which can allow one gene to encode multiple proteins. Alternative splicing is a process by which the exons of primary transcripts (pre-mRNAs) from genes can be spliced in different arrangements to produce structurally and functionally distinct mRNA and protein variants. It is one of the most extensively used mechanisms that accounts for the greater macromolecular and cellular complexity of higher eukaryotic organisms (Black, 2003; Graveley, 2001; Matlin *et al.*, 2005).

Alternative splicing has long been regarded as a rather rare event in eukaryotic genomes. However, recent genomic and bioinformatics analysis of vast amount of transcript data in human and other organisms suggest that alternative splicing is widespread in almost all higher eukaryotic genomes (Mironov et *al.,* 1999; Croft *et al.,* 2000; Modrek *et al.,* 2001; Brett *et al.,* 2002; Xu *et al.,* 2002; Modrek & Lee, 2003; Johnson *et al.,* 2003; Sorek *et al.,* 2004). In humans, the frequency of alternative splicing has increased dramatically from a prior estimates as low as 5% (Sharp, 1994) to 35-65% of all genes and 74% of all multi-exon genes having at least one alternative splice form (Kan *et al.,* 2001; Kochiwa *et al.*, 2002; Leipzig *et al.,* 2004).

The much popular "*one gene, one protein*", a central dogma of modern molecular biology which had a profound impact on biologists' thinking, therefore need to be reconsidered as "*one gene many proteins concept*". Now it is hypothesized that more complex an organism is, more likely it can function by extracting multiple protein meanings from individual genes. As a result, our view towards many biological processes, such as protein interaction, gene expression had to be adjusted. Understanding the functional impact of alternative splicing, the regulatory mechanisms that govern RNA splicing, and the role of alternative splicing in genome evolution, has become one of the most challenging and exciting tasks for genomics and bioinformatics in the post-genomic era. The types of alternative splicing alteration that have been observed include exon skipping, intron retention, use of alternative splice donor or acceptor site and many more (Sanford & Caceres, 2004; Alekseyenko *et al.,* 2007; Singh, 2007).

Role and importance of alternative splicing

Alternative splicing is a unique mechanism through which the information stored in the genes of complex organisms can be edited in a variety of ways, making it possible for a single gene to specify two or more distinct proteins. It, allows different tissue types to perform diverse functions working from the same small gene assortment. Variation in mRNA structure can occur in many different forms (Lopez, 1998; Smith & Valcarcel, 2000). Exons can be spliced into mRNA or skipped. Introns that are normally excised can be

retained in the mRNA. The positions of either 5' or 3' splice sites can shift to make exons longer or shorter. In addition to these changes in splicing, modifications in transcriptional start site or poly-adenylation site also allow production of multiple mRNAs from a single gene. All of these changes in mRNA structure can be regulated in diverse ways, depending on sexual genotype, cellular differentiation, or the activation of particular cell signaling pathways.

Molecular analysis during the last decade demonstrates that alternative splicing plays a major role in many tissue and developmental stage specific processes, for example, functional roles in the SXL proteins related to the sex determination of Drosophila (Sosnowski *et al.,* 1989; Inoue *et al.,* 1990), antibody response in human (Lopez, 1998) and in Down's syndrome cell adhesion molecule (DSCAM) of human and Drosophila (Yamakawa *et al.,* 1998; Schmucker *et al.,* 2000). Moreover, coordinated changes in alternative splicing patterns of multiple genes pre-mRNAs are an integral component of gene expression programs like those involved in differentiation in nervous system (Grabowski, 1998), apoptotic cell death (Jiang & Wu, 1999) tumor progression (Herrlich *et al.,* 1993) etc.

Alternative splicing has also been implicated in a large number of human pathologies, including neurodegenerative, cardiovascular, respiratory and metabolic diseases, as well as cancer (Stallings-Mann *et al*., 1996; Liu *et al*., 1997; Zolezzi *et al.,* 1997; Wilson *et al.,* 1997) and Alzheimer's disease (Guo *et al.,* 2001). In tissues with complex activities, like the brain, there is a demand for a large number of proteins to perform different functions. Any defect in alternative splicing can cause severe problems for the cell due to the altered protein composition. Therefore, many neurological diseases are caused by defects in alternative splicing. Aberrant splicing has been found to be associated with various diseases, including cancer. Mutations in splicing regulatory elements within the nucleotide sequence and alterations in the cellular-splicing-regulatory machinery both result in changes in the splicing pattern of many cancer-related genes (Pajares *et al.,* 2007). Alternative splicing regulates numerous aspects of protein function, such as binding properties, intracellular localization, enzymatic activity, stability and post-translational modifications. At times, it can display strong specificity to a particular tissue or developmental stage (Joseph *et al.,* 1995; Chen *et al.,* 1999), modulating the functional characteristics of protein isoforms in specific tissues (Karpova *et al.,* 2001). Alternative splice events that affect the protein coding region of the mRNA will give rise to proteins which differ in their sequence and therefore in their activities. Alternative splicing within the non-coding regions of the RNA can result in changes in regulatory elements such as translation enhancers or RNA stability domains, which may have a dramatic effect on the level of protein expression.

Alternative splicing can have profound effects on the structure and function of the protein encoded by a gene (Lopez, 1998; Smith & Valcarcel, 2000). Many proteins are comprised of several domains, or modules, that serve a particular function. For example, one domain may help the protein bind to another protein, while another domain gives the protein enzymatic activity. By alternative splicing, exons, and, therefore, protein domains, can be mixed and matched, altering the nature of the protein. Alternative splicing not only generates segments of mRNA variability that have the potential to insert or remove amino acids, shift the reading frame or introduce a termination codon, but also affects gene expression by removing or inserting regulatory elements controlling translation, mRNA stability and localization. It also changes the structure of transcripts, which has the potential to dramatically increase the functional diversity of encoded gene products and allow mRNA isoforms to be differentially regulated in disparate biological processes (Hagiwara, 2005). Despite the importance of alternative splicing, the general mechanisms that regulate those alternative splicing are unknown so, understanding the mechanism and regulation of alternative splicing is a major goal of modern biological and biomedical research.

Alternative Splicing in various organisms

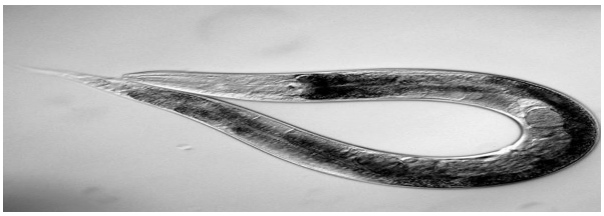
Alternative splicing is found extensively in all higher eukaryotes with most information from well-studied organisms like *Caenorhabditis elegans,* (Tabish *et al.,* 1999) *Drosophila* (Schumuker *et al.,* 2000), Mouse (Tabish & Ticku, 2004) and humans (Brett *et al.,* 2002; Bingham *et al.,* 2008), although no regulated splicing has been detected in unicellular eukaryotes such as yeasts. Stetefeld and Ruegg (2005) estimated that almost 50% of eukaryotic genes and at least one-third of the genes if less complex organisms, such as nematode or flies, are alternatively spliced. Brett *et al*. (2002), by a large-scale expressed sequence tag (EST) analysis, estimated that across a variety of distinct metazoan organisms such as nematodes, the rate of alternative splicing is similar to humans. Kan *et al*., (2002) have suggested that given a sufficient amount of EST coverage, alternative splice patterns may be observed for all genes that undergo splicing; hence, the higher the EST coverage, the higher the level of alternative

splicing we expect. Contrary to this, Kim *et al*., (2004) in a somewhat indirect method to calculate the level of alternative splicing, revealed that the level of alternative splicing varies between different organisms with a greater amount of alternative splicing in mammals compared with invertebrates. However, in reply, Harrington *et al*., (2004) found that the most of the above studies were based on the EST coverage of the organisms, which has its own inherent limitations. Recently, Kim *et al*., (2006) successfully demonstrated that vertebrates have a substantially higher percentage of alternatively spliced genes compared with other species. Their results were based on a straight forward approach based on detection of alternative splicing events in gene-oriented clusters of mRNAs and ESTs in eight eukaryotic organisms. The difference in the level of alternative splicing suggests that alternative splicing may contribute greatly to the mammal higher level of phenotypic complexity, and that accumulation of introns confers an evolutionary advantage as it allows increasing the number of alternative splicing forms. Secondly, it could also be inferred that the appearance of vertebrates, some 300 million years ago, was accompanied by an increase in the number of alternatively spliced genes. The lack of sufficiently large data sets of alternative splicing microarray data and sequenced ESTs and cDNAs has prevented reliable estimates of the proportions of genes that undergo alternative splicing in other organisms especially in *C. elegans*. Moreover, there are also insufficient data currently available to accurately assess the overall number of alternative splicing events in any one organism.

Alternative splicing in *C. elegans*

The value of *Caenorhabditis elegans* as a model organism for biomedical research is unquestioned (Fig. 1). Its short life cycle, self-fertilizing propagation, simple anatomy, and the ease of genetic and experimental manipulations made *C. elegans* an important model system in biology. It is an important, well-studied organism used in biomedical research as a model for human development, genetics, ageing, and disease. Over half of *C. elegans* genes have human orthologs, while ~42% of human disease genes have a homolog in *C. elegans* (Culetto & Sattelle, 2000; Sternberg, 2001). Deciphering the biological information from these sequenced genomes is of great importance, as the information that we get from them are applicable to higher complex organisms including humans. With its many alternatively spliced genes and evolutionarily conserved alternative splicing factors, *C. elegans* offers a remarkable model system for the study of alternative splicing.

Fig. 1. A wild type Adult Hermaphrodite Caenorhabditis elegans



The number of alternatively-spliced forms per gene tends to be relatively less in *C. elegans*. Over 90% of the genes have either one or two alternatively-spliced forms. Out of the total spliced transcripts, only 7,833 (28.5%) genes have been fully confirmed by the available cDNA and EST sequences. This is likely an underestimate of the total percentage of *C. elegans* genes that are alternatively spliced as these values are dependent on EST and cDNA coverage of the genome and at the current time 28.5% of the total annotated *C. elegans* coding sequences have confirmed cDNA or EST support (Wormbase WS160 release notes). For these reasons, it is difficult to estimate the percentage of alternatively spliced genes in the nematode, all one can say is that it is certainly higher than current estimates, and probably lower than in mammals. So, there is a large probability and chances that further analysis using different gene/exon finders may possibly identify large number of alternatively spliced transcripts not available in the database.

Computational Methods for Identifying alternative spliced transcripts

Several different strategies have been applied to alternative splicing analysis, including (i) *EST mapping against mRNA* (Brett *et al.,* 2000) (ii) *mRNA/EST/protein mapping to the genome* (Gupta *et al.,* 2004) (iii) *splicing microarray analysis* (Johnson *et al.,* 2003) (iv) ab initio machine learning approaches (Cawley *et al.,* 2003; Dror *et al.,* 2005; Ratsch *et al.,* 2005) and *(v) Various other approaches* published in recent years involved identifying and exploiting local sequence features for prediction. For instance, Dror *et al*., (2005) used features like exon length, its divisibility by three, the length of flanking introns and the intensity of the polypyrimidine tract were utilized to identify possible spliced transcripts. Kan *et al.,* (2004) showed that EST and genome data for related organisms such as human and mouse can be combined to obtain predictions of alternative splicing. Yeo *et al.,* (2005) described an approach ACESCAN that is able to identify conserved exon skipping events in both human and mouse. This approach also uses exonic and intronic conservation as well as splice site scores, exon

and intron lengths, and oligonucleotide composition. Ohler *et al.,* (2005) demonstrated that even alternative exons that are completely missed in current gene annotations can be discovered by applying a pair hidden Markov model algorithm to orthologous human-mouse introns. Finally, Ratsch *et al.,* (2005) used a support vector machine to predict alternative exons. These studies demonstrate that a classifier based on characteristic genomic features can reliably predict exon skipping events *ab initio*.

Inspite of a large pool of methods available, none of the above approaches have been fully successful in delineating all possible alternative splice transcripts of a gene because of their inherent limitations or technical flaws or appliacbilty to a particular specie(discussed in detail in reviews published before) (Christopher & Qi, 2005; Sylvain & Thomas, 2005; Srinivasan *et al.,* 2005; Florea, 2006; Agrawal & Stormo, 2006; Bonizzoni *et al.,* 2006; Talavera *et al.,* 2007; Leparc & Mitra, 2007). Moreover, most of the work and studies have been limited to Humans and Mouse, not much emphasis has been given to study the alternatively spliced transcripts from *C. elegans* genome. Threfore, a strong need to develop a method capable of delineating the full spectrum of all possible spliced transcripts of a gene is warranted.

Bioinformatics strategy adopted by us for detection of spliced transcripts

Recently, a novel bioinformatics approach used by us has led to a successfull identification and verification of the existence of several previusly un-detected spliced transcripts of various *C. elegans* genes (Kashyap *et al.,* 2007a, Kashyap *et al.,* 2007b, Kashyap *et al.,* 2007c). Our research work comprised of complete analysis of the un-annotated unusually large intronic, 5' and 3' untranslated (UTR) genomic regions of chromosome 1 of *C. elegans.* The major thrust was on finding new exons and genes encoded by chromosome 1 by using a wide array of bioinformatics tools like gene/exon prediction programmes, ORF finders, Blast analysis tools, alignment programmes etc. to analyze the genes/exons encoded by chromosome one of *C. elegans* with special reference to alternatively spliced transcripts arsing from respective gene. When large genomic sequences from large UTR and intronic gap regions were fed into this array of tools in a specific and pre decided manner, it was predicted that several exons possibly capable of replacing the existing exon(s) and thus creating alternatively spliced variant of the gene. From the several exons predicted above, we selected only the "common exons" capable of replacing the existing exon(s) and thus creating spliced transcript of the gene without affecting the reading frame of the protein. Further, the possibility of occurrence of that spliced exon/transcript was analyzed by a comparative analysis between the original protein and the new protein formed by addition/exclusion of alternative exons using various alignment tools. Lastly, several other parameters like percent-amino acid replacement, codon usage, sense nature i.e. whether from positive or negative strand, the probability score of occurrence of that exon etc. were also checked to ensure almost accuracy of predicted spliced transcript of the gene.

Thus, by using the above methodology, we identified around 186 new alternatively spliced transcripts arising from unusually large untranslated and intronic regions of various genes from chromosome 1 of *C. elegans*. Following the computational predictions of these novel spliced transcripts, the Yuji Khoara's and NCBI dbEST *C. elegans* EST database and various other relevant databases were searched to look for putative EST/cDNA support for possible occurrence of these new exons/transcripts. Briefly, using various alignment tools, we looked for insertions or deletions in ESTs relative to a set of known mRNAs or by aligning the ESTs exactly to their identified genomic sequence in the draft genome to identified potential alternative splices. Secondly, as intronic sequences at splice junctions are highly conserved (99.24% of introns have a GT-AG at their 5' and 3' ends, respectively); hence these splices were identified and intronic splice junction donor and acceptor sites were checked using various splice site predicting tools. Finally, NCBI BLAST search was done to look for homology or prospective similarity with other polypeptides of these new spliced transcripts. Further, to experimentally validate our findings, we did RT-PCR using gene specific primers and RNA isolated from mixed population of *C. elegans* for few of the predicted spliced transcripts of the genes, like the *lfe-2* encoding C46H11.4 gene (Kashyap *et al.,* 2007a), RhoGEF domain encoding Y95B8A.12 gene (Kashyap *et al.,* 2007b) and classical cadherin encoding W02B9.1 gene (Kashyap *et al.,* 2007c).

Our results demonstrate that, we are still far from completely deciphering these hidden transcripts from the genome of sequenced organism and most of the studies have probably underestimated the extent of alternative splicing. Although we were successful in identifying potentially new spliced transcripts and alternative exons from chromosome 1 of *C. elegans* but our findings indicate that there could be approximately one

thousand or more alternatively spliced transcripts expressed from the genome of *C. elegans* which have not been annotated or identified earlier. This could be the reason why number of gene products is suspected to be underestimated. With this experience, we propose that the genome data may be analyzed using combination of bioinformatics tools and programmes to predict all possible gene products. These new coding sequences, not annotated or identified earlier will not only enhance the available splice data base of *C. elegans* but will also improve our knowledge about understanding of the genome structure and evolution of higher eukaryotes specially in context to humans.

The above findings could be very useful to biologists in several ways: Firstly our data not only increases the available database for alternatively spliced genes in *C. elegans* but also point towards the complex mechanism of alternative splicing in *C. elegans* genes and their role in downstream regulatory steps. Secondly, our results not only emphasize the needs to develop more efficient algorithms and methods capable of identifying alternative spliced transcripts but also toward analyzing genome data using a combination of gene/exon finders to delineate all possible gene products and to decipher the true extent of alternative splicing in *C. elegans* genes. Moreover, similar exhaustive studies could be taken up in several other finished genomes especially of humans with whom *C. elegans* share a close gene homology. Lastly, due to limited domain of our work, further studies using more advanced techniques like the RNA interference (RNAi) could be taken up which would enhance our knowledge about the biological and functional significance of these spliced transcripts and their possible role in *C. elegans* gene functioning and regulation.

Future directions

Although many model organisms have now been completely sequenced, we are still very far from understanding cellular function from genome sequence. Interpreting the raw sequence data into useful biological information, also known as genome annotation, is a complex process that requires the efficient integration of computational analyses, auxiliary biological data, and biological expertise. One complicating factor in this uphill task is the expression of multiple alternative mRNA transcripts from a single gene called “alternative splicing”. Alternative splicing is without doubt one of the most important gene regulatory mechanisms, and one that has reemerged as a central concept in the post-genome sequencing era. Determining the extent and importance of alternative splicing required the confluence of critical advances in data acquisition, improved understanding of biological processes and the development of fast and accurate computational analysis tools. Hence, the major goal of alternative splicing annotation project in various organisms should be aimed at not only developing better and more efficient methods to identify the entire spectrum of alternative splice forms but also at developing an exhaustive pool of alternative transcripts of an organism to fully understand the complexity of eukaryotes.

References

1. Agrawal R and Stormo GD (2006) Using mRNAs lengths to accurately predict the alternatively spliced gene products in *Caenorhabditis elegans*. *Bioinformatics*. 22, 1239-1244.
2. Alekseyenko AV, Kim N and Lee CJ (2007) Global analysis of exon creation versus loss and the role of alternative splicing in 17 vertebrate genomes.*RNA*. 13, 661-670.
3. Bingham JL, Carrigan PE, Miller LJ, Srinivasan S.(2008) Extent and diversity of human alternative splicing established by complementary database annotation and microarray analysis. OMICS. 12, 83-92.
4. Black DL (2003) Mechanisms of alternative pre-messenger RNA splicing. *Annu. Rev. Biochem.* 72, 291-336.
5. Bonizzoni P, Rizzi R and Pesole G (2006) Computational methods for alternative splicing prediction. *Brief Funct. Genomic Proteomic*. 5, 46-51.
6. Brett D, Hanke J, Lehmann G, Haase S, Delbrück S, Krueger S, Reich J, Bork P (2000) EST comparison indicates 38% of human mRNAs contain possible alternative splice forms. *FEBS Lett*. 474, 83-86.
7. Brett D, Pospisil H, Valcarcel J, Reich J and Bork P (2002) Alternative splicing and genome complexity. *Nature Genet*. 30, 29-30.
8. Cawley S and Pachter L (2003) HMM sampling and applications to gene finding and AS. *Bioinformatics.* 19, 36-41.
9. Chen CD, Kobayashi R and Helfman DM (1999) Binding of hnRNP H to an exonic splicing silencer is involved in the regulation of alternative splicing of the rat beta-tropomyosin gene. *Genes Dev.* 13, 593-606.
10. Christopher L and Qi W (2005) Bioinformatics analysis of alternative splicing. *Briefings in Bioinformatics*. 6(1), 23-33.
11. Claverie JM (2001) Gene number. What if there are only 30,000 human genes? *Science.* 291, 1255-1257.
12. Croft L, Schandorff S, Clark F, Burrage K, Arctander P, and Mattick JS (2000) ISIS, the intron information system, reveals the high

frequency of alternative splicing in the human genome. *Nat. Genet.* 24, 340-341.
13. Culetto E and Sattelle DB (2000) A role for *Caenorhabditis elegans* in understanding the function and interactions of human disease genes. *Human Molecular Genetics.* 9, 869-877.
14. Dror G, Soreck R and Shamir R (2005) Accurate identification of alternatively spliced exons using support vector machine. *Bioinformatics.* 21, 897-901.
15. Florea L (2006) Bioinformatics of alternative splicing and its regulation. *Brief Bioinform*. 7, 55-69.
16. Grabowski PJ (1998) Splicing regulation in neurons: tinkering with cell-specific control. *Cell*. 92, 709-712.
17. Graveley BR (2001) Alternative splicing: Increasing diversity in the proteomic world. *Trends Genet.* 17, 100-107.
18. Guo HA, Lutsch JK, Leheste G, *et al.,* (2001) A: Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. *Cell.* 106, 319-329.
19. Gupta S, Zink D, Korn B, Vingron M and Haas SA (2004) Genome wide identification and classification of alternative splicing based on EST data. *Bioinformatics*. 20 (16), 2579-2585.
20. Hagiwara M (2005) Alternative splicing: a new drug target of the post-genome era *Biochim. Biophys. Acta*., 1754, 324-331.
21. Hammes A, Guo JK, Lutsch G, Leheste JR, Landrock D, Ziegler U, Gubler MC and Schedl A (2001) Two splice variants of the Wilms' tumor 1 gene have distinct functions during sex determination and nephron formation. *Cell* .106, 319-329.
22. Harrington ED, Buoe S, Valcarcel J, Reich JG, Bork P (2004) Estimating rates of alternative splicing in mammals and invertebrates. *Nat Genet.* 36, 916-917.
23. Herrlich P, Zoller M, Ponta H (1993) CD44 splice variants: Metastases meets lymphocytes. Immunol Today 14,395-399,
24. Inoue K, Hoshijima K, Sakamoto H, Shimura Y (1990) Binding of the Drosophila sex-lethal gene product to the alternative splice site of transformer primary transcript. *Nature*. 29, 344,461-463.
25. International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature.* 409, 860-921.
26. Jiang ZH and Wu JY (1999) Alternative splicing and programmed cell death. *Proc. Soc. Exp. Biol. Med*. 220, 64-72.
27. Johnson JM, Castle J, Garrett-Engele P, Kan Z, Loerch PM, Armour CD, Santos R, Schadt EE, Stoughton R and Shoemaker DD (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science.* 302, 2141-2144.
28. Joseph R, Dou D and Tsang W (1995) Neuronatin mRNA: alternatively spliced forms of a novel brain-specific mammalian developmental gene. *Brain Res.* 690, 92-98.
29. Kan Z, Rouchka EC, Gish WR and States DJ (2001) Gene structure prediction and alternative splicing analysis using genomically aligned ESTs. *Genome Res*. 11, 889-900.
30. Kan Z, Startes D and Gish W (2002) Selecting for functional alternative splices in ESTs. *Genome Res.* 12, 1837-1845.
31. Karpova AY, Ronco LV and Howley PM (2001) Functional characterization of interferon regulatory factor 3a (IRF-3a), an alternative splice isoform of IRF-3. *Mol. Cell. Biol.* 21, 4169-4176.
32. Kashyap L and Tabish M (2007c) Alternatively spliced isoforms encoded by cadherin genes from *C. elegans* genome. *Bioinformation*. 2, 50-56.
33. Kashyap L, Tabish M, Ganesh S *et al.,* (2007a) Computational and molecular characterization of multiple isoforms of *lfe-2* gene in nematode *C. Elegans. Bioinformation*. 2, 17-21.
34. Kashyap L, Tabish M, Ganesh S *et al.,* (2007b) Identification and comparative analysis of novel alternatively spliced transcripts of RhoGEF domain encoding gene in *C. elegans* and *C. briggsae*, *Bioinformation*. 2, 43-49.
35. Kim E, Magen A, Ast G (2007) Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res*. 35, 125-131.
36. Kim H, Klein R, Majewski J and Ott J (2004) Estimating rates of alternative splicing in mammals and invertebrates. *Nat. Genet*. 36, 915-916.
37. Kochiwa H, Suzuki R, Washio T, Saito R, Bono H, Carninci P, Okazaki Y, Miki R, Hayashizaki Y, and Tomita M (2002) Inferring alternative splicing patterns in mouse from a full-length cDNA library and microarray data. *Genome Res*. 12, 1286-1293.
38. Leipzig J, Pevzner P and Heber S (2004) The Alternative Splicing Gallery (ASG): bridging the gap between genome and transcriptome. *Nucleic Acids Res.* 32, 3977-3983.
39. Leparc GG and Mitra RD (2007) Non-EST-based prediction of novel alternatively spliced cassette exons with cell signaling function in

Caenorhabditis elegans and human. *Nucleic Acids Res*. 35, 3192-3202.
40. Liu W, Qian C and Francke U (1997) Silent mutation induces exon skipping of fibrillin-1 gene in Marfan syndrome. *Nature Genet.* 16, 328–329.
41. Lopez AJ (1998) Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. *Annu. Rev. Genet.* 32, 279-305.
42. Matlin AJ, Clark F and Smith CWJ (2005) Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol.* 6, 386-398.
43. Mironov AA, Fickett JW and Gelfand MS (1999).Frequent alternative splicing of human genes *Genome Res.* 9, 1288–1293.
44. Modrek B and Lee CJ (2003) Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss. *Nat. Genet.* 34, 177-180.
45. Modrek B, Resch A, Grasso C and Lee C (2001) Genome-wide analysis of alternative splicing using human expressed sequence data. *Nucleic Acids Res.* 29, 2850−2859.
46. Ohler U, Shomron N, Burge CB (2005) Recognition of unknown conserved alternatively spliced exons. *PLoS Comput. Biol*. 1(2), e15.
47. Pajares MJ, Ezponda T, Catena R, Calvo A, Pio R and Montuenga LM (2007) Alternative splicing: an emerging topic in molecular and clinical oncology. *The lancet Oncology*. 8, 349-357.
48. Philipps DL, Park JW, Graveley BR (2004) A computational and experimental approach toward a priori identification of alternatively spliced exons. RNA 10(12), 1838–1844.
49. Ratsch G, Sonnenburg S, Scholkopf B (2005) RASE: recognition of alternatively spliced exons in C.elegans. *Bioinformatics.* 21 (1), i369–i377.
50. Sanford JR and Caceres F (2004) Pre-mRNA splicing: life at the centre of the central dogma. *J. Cell Sci*. 117(26), 6261- 6263.
51. Schmucker D, Clemens JC, Shu H, Worby CA, Xiao J, Muda M, Dixon JE and Zipursky SL (2000) Drosophila Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* .101, 671-684.
52. Sharp PA (1994) Split genes and RNA splicing. *Cell.* 77, 805–815.
53. Singh RN (2007) Unfolding the mystery of alternative splicing through a unique method of in vivo selection. *Frontiers in Bioscience.* 12, 3263-3272.
54. Smith CW and Valcarcel J (2000) Alternative pre-mRNA splicing: the logic of. combinatorial control. *Trend Biochem. Sci*. 25, 381-388.
55. Sorek R, Shamir R and Ast G (2004) How prevalent is functional alternative splicing in the human genome? *Trends Genet*. 20, 68–71.
56. Sosnowski BA, Belote JM and McKeown M (1989) Sex-specific alternative splicing of RNA from the *transformer* gene results from sequence-dependent splice site blockage. *Cell.* 58, 449-459.
57. Srinivasan K, Shiue L, Hayes JD, Centers R, Fitzwater S, Loewen R, Edmondson LR, Bryant J, Smith M, Rommelfanger C, Welch V, Clark TA, Sugnet CW, Howe KJ, Mandel-Gutfreund Y and Ares M Jr (2005) Detection and measurement of alternative splicing using splicing-sensitive microarrays. *Methods.* 37, 345-59.
58. Stallings-Mann ML, Ludwiczak RL, Klinger KW, Rottman F (1996) Alternative splicing of exon 3 of the human growth hormone receptor is the result of an unusual genetic polymorphism. *Proc. Natl. Acad. Sci. USA.* 93, 12394–12399.
59. Sternberg PW (2001) Working in the post-genomic *C. elegans* world. *Cell.* 105, 173–176.
60. Stetefeld J and Ruegg MA (2005) Structural and functional diversity generated by alternative mRNA splicing. *Trends Biochem. Sci*. 30, 515–521.
61. Sylvain F and Thomas S (2005) Integrating alternative splicing detection into gene prediction. *BMC Bioinformatics*. 6, 25.
62. Tabish M and Ticku MK (2004) Alternate splice transcripts of mouse NR2B gene. *Neurochem. Int.* 44, 339-343.
63. Tabish M, Clegg RA, Rees HH and Fisher MJ (1999) Organization and alternative splicing of the *Caenorhabditis elegans* cAMP-dependent protein kinase catalytic-subunit gene (*kin-1*). *Biochem. J.* 339, 209-216.
64. Talavera D, Hospital A, Orozco M and de la Cruz X (2007) A procedure for the identification of homologous alternative splicing events. *BMC Bioinformatics*.8(1),260.
65. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, *et al.* (2001) The sequence of the human genome. *Science.* 291, 1304-1351.
66. Wilson CA, Payton MN, Elliott GS, Buaas FW, Cajulis EE, Grosshans D, Ramos L, Reese DM, Slamon DJ and Calzone FJ (1997) Differential sub-cellular localization, expression and biological toxicity of BRCA1 and the splice variant BRCA1-D11b. *Oncogene.* 14, 1-16.

67. Xu Q, Modrek B and Lee C (2002) Genome-wide detection of tissue-specific alternative splicin g in the human transcriptome. *Nucleic Acids Res.* 30, 3754-3766.
68. Yamakawa K, Huot YK, Haendelt MA, Hubert R, Chen XN, Lyons GE and Korenberg JR (1998) DSCAM: a novel member of the immunoglobulin superfamily maps in a Down syndrome region and is involved in the development of the nervous system. *Hum. Mol. Genet*. 7(2),227-237.
69. Yeo GW, Van Nostrand E, Holste D, Poggio T and Burge CB (2005) Identification and analysis of alternative splicing events conserved in human and mouse. *Proc. Natl. Acad. Sci. USA.* 102 (8), 2850-2855.
70. Zolezzi F, Valli M, Clementi M, Mammi I, Cetta G, Pignatti PF and Mottes M (1997) Mutation producing alternative splicing of exon 26 in the COL1A2 gene causes type IV osteogenesis imperfecta with intrafamilial clinical variability. *Am. J. Med. Genet.* 71, 366-370.

Dr.Luv Kashyap, a Post-Doctoral Associate at University of Pittsburgh, has spent over 5 years in working on "*Caenorhabditis elegans*" and finds it as a wonderful model organism. His current theme of research is centered on *C. elegans* as model to identify the genes involved in Sarcopenia using RNAi. He obtained his Ph.D. from Aligarh Muslim University in the field of Biochemistry with specialization in Bioinformatics. During that period, he succesfully developed a new method to predict novel and previously undetected alternatively spliced transcripts from *C.elegans* genome using a combination of a large array of bioinformatics tools and molecular biology techniques.