

Luv Kashyap* och Parul Tripathi**

*Department of Biochemistry, Aligarh
Muslim University, Aligarh, India

** Immunology Group, International Centre for Genetic
Engineering and Biotechnology, New Delhi, India

Alternativ Splitsning

Hur en gen kan ge flera proteiner

Människor har bara dubbelt så många gener som den vanliga bananflugan. Hur kan det komma sig? Vi har ju trots allt hjärnor som gör att vi kan räkna gener, medan deras huvudsyssla är att hitta en fin banan. Alternativ splitsning förklarar både hur vår genetiska information kan ta en mycket liten plats och mekanismerna bakom flera mänskliga sjukdomar.

Sekvenseringen av den mänskliga arvsmassan (1) har väckt viktiga frågor om den genetiska komplexitetens natur. Tidigare trodde forskare att människans komplicerade DNA bestod av kanske upp till 150 000 olika gener. Uppskattningen baserades på det antal olika transkript, mRNA, som man hade hittat hos människa. Man antog att det borde finnas en gen för varje mRNA. När hela sekvenseringen var klar hade man bara hittat 32 000 mänskliga gener. (Senare resultat visar att antalet antagligen är ännu färre, mindre än 25 000.) Som en jämförelse har bananflugan *Drosophila* 14 000 gener och nematoden *Caenorhabditis elegans* 19 000. Rapporten om endast 32 000 mänskliga gener kom som en överraskning (2). Antalet mänskliga sekvenser som uttrycks (mRNA) var mycket större än antalet gener. Hur kan koden för människan, som är så mycket större och mer komplicerad, få plats i bara dubbelt så många gener som en fluga?

En förklaring till denna paradox är att det antal proteiner som ett genom kan ge upphov till vida överstiger antalet gener, om en stor andel av generna har förmågan att koda för flera proteiner. En sådan utvidgning av proteomet kan åstadkommas genom alternativ splitsning av prekursor-budbärrar-RNA (pre-mRNA).

Vad är alternativ splitsning?

Alternativ splitsning är en process som gör att olika varianter av proteiner kan produceras från en enda gen (3, 4) (**Fig. 1**). De flesta generna i eukaryota genom består av exoner och introner. Efter transkription måste intronerna tas bort från pre-mRNA via ett steg som kallas splitsning. Ibland kan ett exon antingen inkluderas eller exkluderas i det slutliga transkriptet, eller också kan det vid ett exons ena ände finnas två splitsningsställen som känns igen av splitsosomen (eng. spliceosome, det molekylkomplex som bär med sig splitsningsmaskineriet). På så sätt kan flera olika transkript framställas via "alternativ splitsning".

Adress för korrespondens:
Dr. Parul Tripathi,
Research Scientist
Immunology Group
International Centre for Genetic
Engineering and Biotechnology (ICGEB),
New Delhi, India
E-mail: parultripathi@rediffmail.com

Genom att foga ihop olika splitsningsställen kan enskilda gener uttrycka flera mRNA som kodar för proteiner med vitt skilda och till och med motverkande funktioner (5). Genom den här mekanismen kan den information som lagras i generna hos komplexa organismer redigeras på en mängd olika sätt, så att en enda gen kan producera två eller flera olika proteiner.

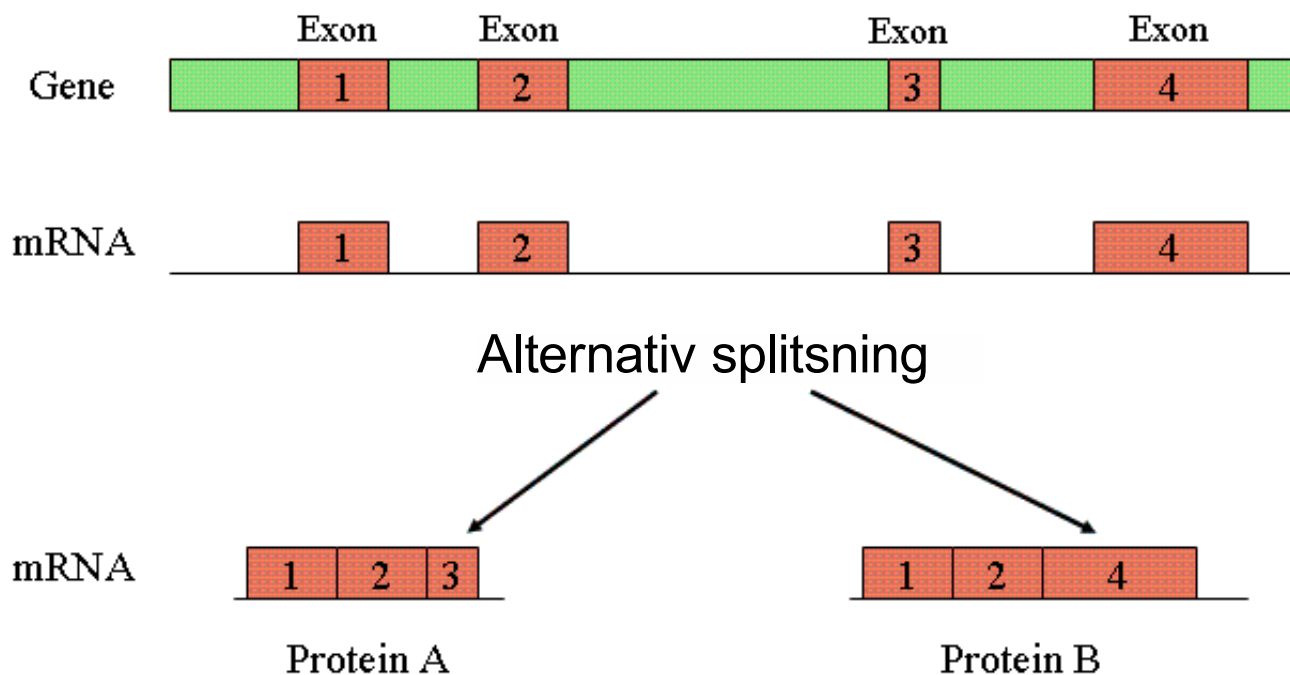


Fig. 1: Beskrivning av alternativ splitsning: I DNA finns den genetiska informationen som kodar för proteiner i fragment (exoner, röda boxar) och mellan dessa finns fragment som inte är kodande (introner, gröna boxar). Genom alternativ splitsning tas intronerna bort och exonerna splitsas ihop i olika kombinationer, vilket ger olika budbärar-RNA (mRNA) som avkodas (translateras) till olika proteiner.

Kanske var det här ett mycket viktigt steg för att eukaryoterna skulle kunna utvecklas mot högre effektivitet, eftersom informationen lagras på ett mycket mer ekonomiskt sätt. Att förstå vilken inverkan alternativ splitsning har på funktion, reglerande mekanismer som styr RNA-splitsning och vilken roll alternativ splitsning har spelat i utvecklingen av olika genom har blivit en av de mest utmanande och spännande uppgifterna inom genetik och bioinformatik i den post-genomiska eran. De olika varianter av alternativ splitsning som har observerats, är bland annat överhoppning av exoner, kvarhållning av introner, användning av alternativa donator- eller mottagarställen för splitsning.

En förvånande vanlig process

Alternativ splitsning ansågs länge vara något ganska ovanligt i eukaryota genom. Till exempel antog man att det skedde i mindre än 5 % av människans gener. På senare tid har dock analyser av stora mängder transkriptdata, från människa och andra organismer, visat att alternativ splitsning är vanligt hos alla däggdjur. Hos människa uppskattar man att alternativ splitsning sker i mer än 60 % av generna.

Den centrala dogmen **en gen, ett protein**, som har varit en ledande teori inom modern molekylärbiologi och haft en djupgående inverkan på hur biologer tänker, måste därför revideras. Till följd av detta måste även synen på många biologiska processer, så som proteininteraktioner och genuttryck, anpassas.

Alternativ splitsning är ett kraftfullt medel för att öka proteinernas mångfald. Man tror att över 70 % av alla gener är föremål för alternativ splitsning för att producera funktionellt olika proteiner från en enstaka gen. De flesta generna hos flercelliga djur kodar faktiskt för pre-mRNA som genomgår alternativ splitsning för att producera allt från två till tiotusentals isoformer av mRNA. Alternativ splitsning förekommer i hög grad hos alla högre eukaryoter, från organismer som *Caenorhabditis elegans* (6, 7), *Drosophila* (8) och mus (9) till människor. Det understryker hur viktigt alternativ splitsning har varit genom hela evolutionen. En av de mest dramatiska exemplen hittar man i genen *dscam* i *Drosophila* (8). Denna enda gen innehåller 116 exoner av vilka 17 bibehålls i det slutliga mRNA:t. Vissa exoner inkluderas alltid medan andra varierar. I teorin skulle det här systemet kunna producera 38 016 olika proteiner. Man har faktiskt hittat mer än 18 000 olika *dscam*-varianter i *Drosophila*. Dessa proteiner bidrar till att vägleda nervceller till sina rätta platser och är antagligen inblandade i igenkänning och fagocytos av invaderande bakterier. Om ett specifikt RNA-segment ska behållas som ett exon eller klippas bort som ett intron beror på omständigheterna.

Reglering av gener

Alternativ splitsning spelar en viktig roll i regleringen av hur gener uttrycks. Under det senaste decenniet har man visat att alternativ splitsning avgör bindningsegenskaper, intracellulär lokalisering, enzymaktivitet, stabilitet och posttranslationella modifieringar för ett stort antal proteiner. Många proteiner består av flera domäner, eller moduler, som tjänar en viss funktion. En domän kan till exempel göra så att ett protein binder till ett annat protein, medan en annan domän ger enzymaktivitet. I genomet motsvaras domäner av exoner. Genom alternativ splitsning av exoner, d.v.s. proteindomäner, kan dessa blandas och fogas ihop så att det bildas proteiner med skiftande egenskaper. Genom att reglera vilka splitsningsmönster som används i olika vävnader, kan organismen finjustera en enda gens funktion så att den kan utföra flera olika uppgifter.

Alternativ splitsning kan verka på två olika nivåer: **förändringar på proteinnivå** och **modifieringar på transkriptnivå**. På proteinnivå ger alternativ splitsning upphov till splitsningsvarianter som producerar olika proteinprodukter. Till exempel kan ett kortare protein bildas, på grund av en läsransmutation som orsakas av att ett alternativt exon. Ett protein med en annorlunda funktionell domän kan uppstå på grund av att ett specifikt exon från en helt annan grupp av exoner sätts in. På transkriptnivå kan olika splitsningsprodukter ha olika translations- eller stabilitetsegenskaper.

När till exempel ett transkripts livslängd ökar, får det effekten att dess protein är tillgängligt under en längre tid. Processen spelar alltså en utomordentligt viktig roll när det gäller att utvidga mångfalden av proteiner och kan därför, åtminstone delvis, förklara det skenbara motsatsförhållandet mellan antalet gener och hur komplicerad en organism är (1).

Alternativ splitsning och mänskliga sjukdomar

Alternativ splitsning har visat sig vara inblandat i ett stort antal sjukdomar hos människa, inklusive neurodegenerativa, kardiovaskulära, respiratoriska och metabola sjukdomar så väl som Alzheimers sjukdom, och det är en lovande måltavla för behandlingar. I komplicerade vävnader, som hjärnan, behövs ett stort antal proteiner för att utföra alla olika funktioner. Felaktig alternativ splitsning kan orsaka allvarliga problem för cellerna, eftersom sammansättningen av proteiner förändras. Många neurologiska sjukdomar orsakas också av defekt alternativ splitsning. Dessa defekter kan delas in i två huvudgrupper: primära och sekundära splitsningsdefekter (17).

Primära splitsningsdefekter

Vid primära splitsningsdefekter har sekvenser som är viktiga för splitsningen av pre-mRNA muterat. Det får till följd att riktig splitsning inte kan ske (**Fig. 2**). Många av de punktmutationer som orsakar sjukdomar hos människa anses vara splitsningsdefekter i pre-mRNA. Två exempel är ataxia telangiectasia (Louiss sjukdom) och neurofibromatos. Hälften av alla som lider av dessa sjukdomar bär på mutationer som påverkar splitsning av pre-mRNA (18,19). Störningar av primär splitsning förekommer också hos patienter med frontotemporal demens. Detta syndrom är en autosomal dominant rubbning som hänger ihop med kromosom 17. Felaktig splitsning gör att det bildas förändrade transkript av det mikrotubuliassocierade proteinet tau (MAPT) som anrikas i mogna och växande neuron. Förändrad splitsning av MAPT ingår även i andra sjukdomars patologi, så som Alzheimers sjukdom, myoton dystrofi, gliopati och degenerering av ryggmärgen. Duchennes muskeldystrofi (DMD) orsakas av mutationer i dystrofin-genen. Mutationerna stör splitsningen och felaktigt mRNA bildas. En annan mutation av dystrofin-genen är mindre allvarlig och leder till en mildare form av DMD: Beckers muskeldystrofi (BMD) (17).

A

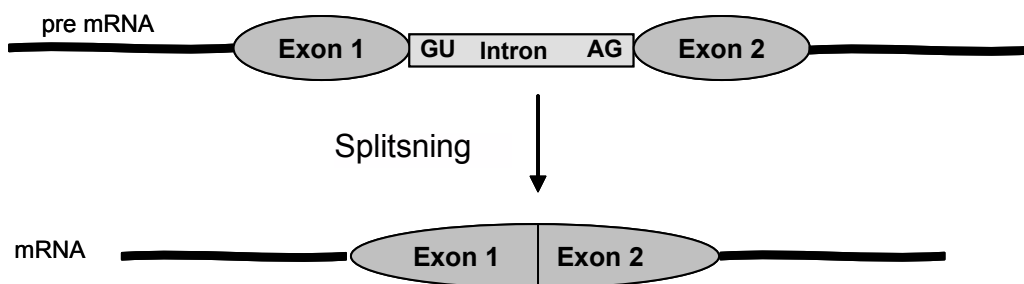
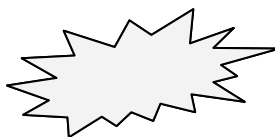


Fig. 2 (A): Exempel på korrekt splitsning. Exon 1 och exon 2 är de delar av pre-mRNA som bär den kodande informationen som behövs för att producera ett protein. Intronet är den icke-kodande sekvensen däremellan som tas bort. Exon 1 och exon 2 splitsas ihop. Efter splitsning används den färdiga mRNA-molekylen till proteinproduktion.

Translation



B

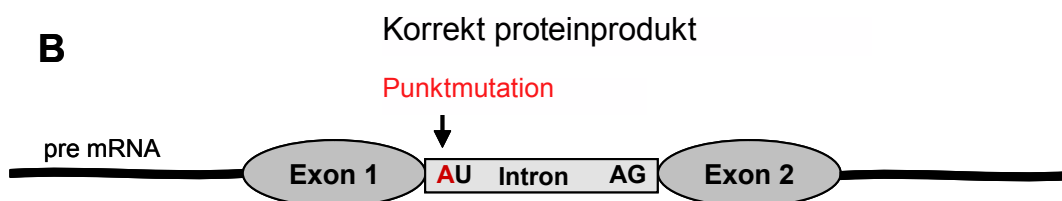
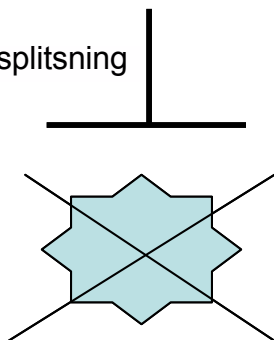


Fig. 2 (B): Exempel på en primär splitsningsdefekt: Sekvensen GU, som definierar intronets början, har muterat. Splitsningsstället är felaktigt (AU istället för GU till följd av en punktmutation i sekvensen) och kan inte längre användas. Intronet klyvs inte bort och inget korrekt protein kan bildas.

Ingen splitsning

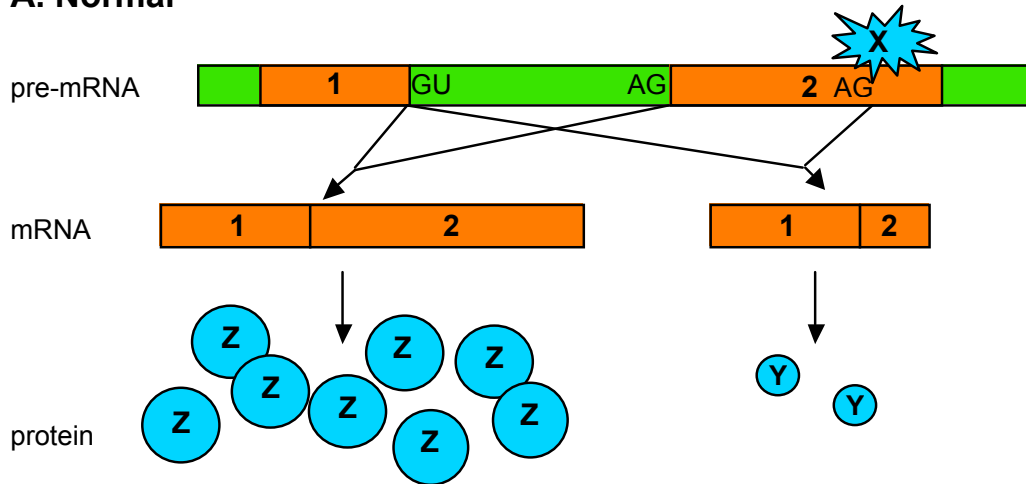


Ingen korrekt proteinprodukt

Sekundära splitsningsdefekter

Vid sekundära splitsningsrubbningar muteras en reglerande faktor som är avgörande för processen så att splitsningsfunktionen störs. Förändringar i faktorer som aktiverar eller hämmar splitsning kan påverka signalvägar för alternativ splitsning. Beroende på vilken reglerande faktor som påverkas förändras mönstret av alternativ splitsning av särskilda pre-mRNA. I vissa fall kan det leda till allvarliga rubbningar medan andra ger mildare effekter.

A. Normal



B. Sjukdom

muterad, kan inte binda

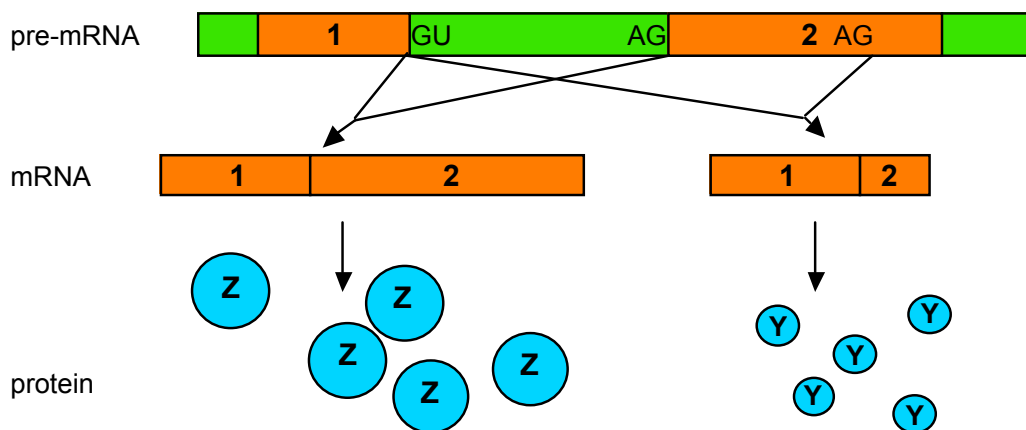


Fig. 3: Sekundära splitsningsdefekter. I det här pre-mRNA:t finns två olika splitsningsställen (AG) i slutet av pre-mRNA. De två olika produkterna som bildas efter splitsning och translation (protein Z och protein Y) skiljer sig åt beroende på vilket splitsningsställe som använts. I normalfallet (övre delen, A) är det ena splitsningsstället delvis blockerat av en proteinfaktor X. Endast det andra stället vid intronets slut är lätt åtkomligt och därför bildas mer av protein Z än av protein Y. Efter en mutation (nedre delen, B) kan den reglerande faktorn X inte längre binda till pre-mRNA:t. Båda AG-sekvenserna är tillgängliga och kan användas för splitsning. När andelen protein Z minskar och proteinnivåerna av Y ökar orsakar det sjukdom.

Ett sådant exempel visas i **Fig. 3**. Två möjliga splitsningsställen finns, AG. Det kan ge två olika färdiga mRNA och därför två olika proteinprodukter (protein Z och protein Y). Den reglerande faktorn (X) binder till det ena splitsningsstället, som finns i den kodande delen av pre-mRNA. Detta splitsningsställe är nu inte tillgängligt för splitsningen och kan inte användas för splitsning (Fig. 3). Faktorn X kan förändras (till exempel genom en punktmutation) så att den inte längre binder till exon 2. Ett exempel på en sekundär splitsningsdefekt är Prader Willis syndrom (PWS). Det är en genetisk rubbning som karakteriseras av intellektuella och beteendemässiga störningar. Sjukdomen kan uppstå vid brist på en liten RNA-molekyl som i normala fall reglerar pre-mRNA-splitsning av serotoninreceptorns RNA.

Enzymet acetylcholinesteras (AChE) är en kombinatorisk serie proteiner som produceras via alternativ splitsning. Man har hittat tre olika splitsade isoformer som är inblandade i fortskridandet av neurodegenerativa sjukdomar som Alzheimers sjukdom och Parkinson. Hos Alzheimerpatienter kan obalans i de faktorer som reglerar splitsning orsaka modifierad alternativ splitsning av AChE pre-mRNA som leder till överskott av isoformen AChE-R. Vid Parkinsons sjukdom förknippas skada och psykologisk stress (viktiga orsaker till sjukdomen) med obalans i alternativ splitsning av AChE och överuttryck av den alternativa isoformen AChE-R (20).

Box 1

EST (= expressed sequence tags) är korta subsekvenser av en transkriberad nukleotidsekvens. De kan användas för att identifiera gentranskript och är användbara vid upptäckt av gener och bestämning av gensekvens.

EST-sekvenser gör molekylärbioologiska studier mer effektiva, särskilt när det gäller undersökningar av genuttryck och analys av gener som är inblandade i virulens och infektioner.

En DNA-mikromatris (kallas också gen- eller genomchip, DNA-chip, eller genmatris) är en samling mikroskopiska DNA-fläckar, som vanligtvis motsvarar enstaka gener, utspridda på en fast yta till vilken de är kovalent bundna till en kemisk matris. (Taget från Wikipedia)

Framtidsutsikter

Alternativ splitsning har inte bara visat sig vara en nyckelfunktion för kontroll av genuttryck hos däggdjur, utan också den huvudsakliga källan till mångfald i den mänskliga genomet.

De bästa metoderna för storskalig undersökning av alternativ splitsning är idag EST (Expressed Sequence Tags) och mikromatrisanalys (microarray analysis) (**Box 1**). Metoderna har sina brister och det är svårt att upptäcka all förekomst av alternativ splitsning (12-15). Nya, mer effektiva metoder som ska kunna hitta alla möjliga splitsningsvarianter av en gen utvecklas hela tiden. Därför ändras siffrorna för hur vanligt alternativ splitsning är varje år. Eftersom alternativ splitsning är så allmänt spritt är det en viktig uppgift att skapa en fullständig förteckning över alla splitsningsvarianter och utveckla en större förståelse för dess ursprung, funktion och reglering.

Om kontrollen över den alternativa splitsningen störs, kan resultatet bli att cellernas och vävnadernas behov inte uppfylls, vilket leder till funktionsrubbnings och sjukdom. Studier av hur felreglering av alternativ splitsning orsakar sjukdomar är ett komplement till studier av de normala regleringsprocesserna och berikar vår förståelse för reglerande mekanismer i allmänhet. I det långa loppet kommer förståelse av hur alternativ splitsning förändras vid sjukdom att underlätta strategier som inriktar sig på att motverka eller kringgå sådana defekter.

Frågor till elever:

- 1 Hur förklarar man att högre organismer, som människor kan vara så komplicerade trots att deras genpool inte är särskilt mycket större än den hos nematoden (19 000) och mindre än risets (38 000-40 000)?
- 2 Vad är Molekylärbiologins centrala dogm? Varför gäller denna inte längre?
- 3 Vilka är skillnaderna mellan primära och sekundära splitsningsdefekter?
- 4 Sök på nätet efter några av de sjukdomar som orsakas av felaktig alternativ splitsning och vad som är utmärkande för dem.
- 5 Ta reda på mer om mikromatristeknik och hur den kan användas vid studier av alternativ splitsning!

Några intressanta webbsidor att utforska:

<http://www.alternative-splicing.org/org/index.html>

http://en.wikipedia.org/wiki/Alternative_splicing

http://www.rnabioinformatics.com/2004/10/intro_what_is_alternative_spli.html

<http://scienceandreason.blogspot.com/2006/11/alternative-splicing.html>

Referenser

- International Human Genome Sequencing Consortium. (2001) Initial sequencing and analysis of the human genome. *Nature* 409 860-921
- Venter, J.C. Adams, MD. Myers, EW. Li, PW. Mural, RJ. et al (2001) The sequence of the human genome. *Science* 291 1304-1351.
- Maniatis, T. and Tasic, B. (2002) Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature* 418 236-243.
- Matlin, A.J. Clark, F. and Smith, CWJ. (2005) Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol* 6 386-398.
- Faustino, NA. and Cooper, TA. (2003) Pre-mRNA splicing and human disease. *Genes Dev* 17 419-437.
- 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.
- Park, YS. S. Kim, Y. Shin, BC. and Cho, NJ. (2003) Alternative splicing of the muscarinic acetylcholine receptor GAR-3 in *Caenorhabditis elegans*. *Biochem Biophys Res Commun* 308 961-965.
- Celotto, AM and Graveley, BR. (2001) Alternative splicing of the *Drosophila* Dscam pre-mRNA is both temporally and spatially regulated. *Genetics* **159** 599-608.

- Waterston, RH. Lindblad-Toh, K. Birney, E. et al., (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* 420 520–562
- Resch, A. Xing, Y. Modrek, B. Gorlick, M. Riley, R. and Lee, C. (2004) Assessing the impact of alternative splicing on domain interactions in the human proteome. *J Proteome Res* 3 76–83.
- Johnson, J. Castle, J. Garrett-Engele, P. et al (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* 302 2141-2144.
- Kashyap, L. and Tabish, M (2007c) Alternatively spliced isoforms encoded by cadherin genes from *C. elegans* genome. *Bioinformation.*, 2, 50-56.
- 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.
- 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.
- Kashyap, L. and Tabish, M. (2006). Comparative analysis of various gene finders specific to *Caenorhabditis elegans* genome. *Bioinformation.*, 1, 203-207.
- Xia, H., Bi, J., and Li, Y. (2006) Identification of alternative 5'/3' splice sites based on the mechanism of splice site competition. *Nucleic Acids Res.* 34, 6305-6313.
- Cáceres JF, Kornblihtt, AR. (2002). Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet.*,18, 186-93.
- Ars E, Serra E, Garcia J, Kruyer H, Gaona A, Lazaro C, Estivill X. (2000) Mutations affecting mRNA splicing are the most common molecular defects in patients with neurofibromatosis type 1. *Hum Mol Genet.*, 9,237-247.
- Teraoka SN, Telatar M, Becker-Catania S, Liang T, Onengut S, Tolun A, Chessa L, Sanal O, Bernatowska E, Gatti RA, Concannon P. (1999) Splicing defects in the ataxia-telangiectasia gene, ATM: underlying mutations and consequences. *Am J Hum Genet.*, 64,1617-1631.
- Meshorer E, Soreq H. (2006) Virtues and woes of AChE alternative splicing in stress-related neuropathologies. *Trends Neurosci.*, 29,216-224.



Tack

Volvox-projektet finanseras av EU:s Sjätte Ramprogram.

Nyckelord:

människans genom, mRNA, bioinformatik, introner, exoner, mutationer