

# 基因选择性剪接的生物信息学研究概况

林鲁萍,马飞,王义权

(厦门大学生命科学学院,厦门 361005)

**摘要:**基因选择性剪接现象是真核生物基本而又重要的调控机制。由于基因的选择性剪接在形成生物复杂性和多样性上具有极其重要的作用,同时选择性剪接与许多人类疾病也密切相关。因此,研究基因选择性剪接是一项十分重要的工作。生物信息学作为一门新兴的学科在研究基因选择性剪接上起关键的作用,尤其在研究基因表达调控机制、选择性剪接基因预测以及选择性剪接基因进化上。文章综述了这方面的最新研究进展,为更深入了解真核生物基因的表达调控机理提供依据。

**关键词:**选择性剪接;生物信息学;基因组

**中图分类号:**Q71

**文献标识码:**A

**文章编号:**0253-9772(2005)06-1001-06

## The Application of Bioinformatics in the Research of Alternative Splicing

LIN Lu - Ping, MA Fei, WANG Yi - Quan

(School of Life Sciences, Xiamen University, Xiamen 361005, China)

**Abstract:** Alternative splicing, a fundamental and important regulatory mechanism in eukaryotes, allows one pre-mRNA to be processed into many different mature forms within a cell, each of which can have distinct functions. As alternative splicing is associated with human diseases, the study of alternative splicing becomes quite important. Bioinformatics is a new subject for the study of alternative splicing, especially for its regulatory mechanism, prediction and origin. Of course, bioinformatics must be combined with experimental research so as to clarify these aspects of alternative splicing. This paper reviewed the recent research progress in this field in the hope to gain a deeper understanding of eukaryotic gene expression regulation.

**Key words:** alternative splicing; bioinformatics; genome

1977年,在Adenovirus hexon基因中发现内含子和外显子之后,Gilbert提出:一种基因外显子的不同组合方式可以产生不同的mRNA产物<sup>[1]</sup>。20世纪80年代早期,选择性剪接现象在免疫球蛋白mu基因和降钙素基因中被发现<sup>[2,3]</sup>。之后的几年,科学家们就估计在高等真核生物中大约有5%的基因是选择性剪接基因<sup>[4]</sup>,进一步研究发现,至少有30%的人类基因存在不同的剪接产物<sup>[5]</sup>,从决定性别到与细胞凋亡有关的基因都存在选择性剪接现

象<sup>[6,7]</sup>。mRNA前体的选择性剪接是真核生物的一种基本而又重要的调控机制,它精细协调基因的功能,高效调节基因的定量表达以及蛋白功能的多样化,影响主要发育方向的确定,对细胞的分化、发育、生理功能和病理状态都有重要意义。基因选择性剪接与许多人类疾病密切相关<sup>[8]</sup>,目前对基因选择性剪接现象的研究已经成为后基因组研究的重要组成部分。

科学家于2001年宣布完成了人类基因组框架图后,即已注意到一个有趣的事实,人类基因为什么

收稿日期:2004-12-16;修回日期:2005-05-22

基金项目:国家自然科学基金项目(编号:60305001,30470938)资助[Supported by the National Natural Science Foundation of China (No. 60305001,30470938)]

作者简介:林鲁萍(1982-),女,硕士研究生,专业方向:生物化学与分子生物学;E-mail:linluping@vip.163.com

通讯作者:王义权(1957-),教授,博士生导师,研究方向:动物分子遗传与进化;Fax:0592-2181015;E-mail:wangyq@xmu.edu.cn

不是原来预计的 8 万至 10 万个,而是 3 万个左右,而最近完成的人类基因组常染色质区测序进一步表明,人类的基因只有 2~2.5 万个左右<sup>[9]</sup>。人类基因组计划首席科学家 Collins 认为,人类基因数比预计要少得多,说明高等生物在使用基因上的简约性。与过去一个基因编码一种蛋白的假设相比,现在是每个基因平均负责制造 3 种蛋白。高等生物不单纯靠增加新基因来获取新功能,还可以通过重新编码扩充已有的可靠资源来达到创新的目的。选择性剪接现象的存在正是生物节约使用基因的最好例证,在生物复杂性形成上具有极其重要的作用,并且可以部分解释基因数与生物复杂性之间的矛盾。随着基因组学的发展,公共数据库中可利用的数据迅速增加,为用生物信息学方法研究基因选择性剪接、预测选择性剪接的结果、揭示选择性剪接的规律创造了机遇。

## 1 基因选择性剪接模式

选择性剪接通过以下几种方式来改变 mRNA 前体的产物(图 1):1) 外显子加入或转移,也称为外显子跳跃(exon skip)。这是选择性剪接最简单也是

最普遍存在的 1 种方式。2) 基因的 5 端或 3 端外显子可以被有选择地延长或缩短,也叫做 5 端或 3 端的选择性剪接(alternative 5 or 3' splicing)。3) 两个相邻的外显子间互相排斥,以至于只有其中 1 个外显子被包括在剪接产物中,这也被称为互相排斥的外显子(mutually exclusive exons)。4) 上述第 3 种方式可以衍生到从 1 个大的外显子集合中选出其中 1 个外显子,作为选择性剪接的产物,简称为多个外显子选一(one-of-N)。5) 在 1 个基因中使用两个或两个以上不同的启动子,这样在 5 端就会产生不同的转录产物,这也称为选择性起始(alternative initiation)。6) 同样的 1 个基因中出现两个或两个以上的 poly A 位点,这可以使转录产物有不同的 3 端,这也称为选择性终止(alternative termination)。7) 在剪接的过程中有的内含子没有被剪切掉,而是被保留下来,起到了外显子的作用,简称为内含子保留(intron retention)。在以上介绍的几种方式中,外显子跳跃是最常见的,其次是 5 端或 3 端的选择性剪接<sup>[10]</sup>;外显子间相互排斥的选择性剪接方式相对较少见;据统计选择性终止的方式存在于 24% 的基因中<sup>[11]</sup>;当然最少见的方式是内含子保留<sup>[12]</sup>。

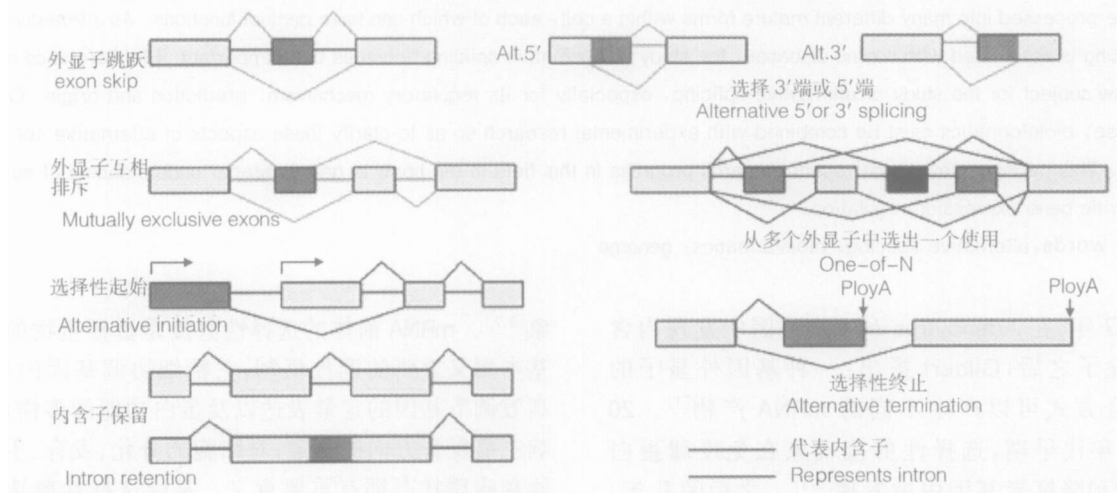


图 1 选择性剪接的几种不同的方式  
被选择性剪接添加或被移动的部分用深色表示<sup>[13]</sup>。

Fig. 1 Types of alternative splicing

Regions that are added or removed by alternative splicing are shaded<sup>[13]</sup>.

## 2 基于生物信息学的基因选择性剪接研究

由于选择性剪接与人类的一些疾病基因相关,研究选择性剪接有以下 3 个目的:1) 发展一些有效

的计算方法,用这些方法和大量的 ESTs 数据预测人类的选择性剪接基因;2) 建立一个实验验证过的选择性剪接数据库,或创建一个预测性的选择性剪接基因数据库,它可以当作最初的资源去检测特定基因的选择性剪接情况,也可以用做疾病基因的诊

断标记,对研究选择性剪接事件的起源也十分重要。  
3) 用实验验证生物信息学方法预测的选择性剪接基因,确定剪接位点预测的正确性,进而可以验证预测方法是否正确,以获得可靠的预测方法。

### 2.1 对选择性剪接调控机制的初步研究

到目前为止,分子生物学家已经发现了与选择性剪接有关的顺式调控元件(如外显子剪接增强子)和反式调控元件(如SR蛋白和PTB等)<sup>[6,13]</sup>。分子生物学研究发现仅局限于一个或者少数几个调控事件。生物信息学则将大量更新中的选择性剪接数据和多种生物的全基因组序列结合起来,对揭示选择性剪接调控机制起到非常重要的作用。

Brudro 等人<sup>[14]</sup>从基因数据库中找出 25 个高表达的脑特异性选择的外显子,将包含这些外显子和它们邻近内含子的 DNA 序列编撰在一起,并用文字对比的算法分析它们核苷酸序列的主要特征,在对比了构成外显子的调控区序列后,发现脑特异性外显子常包括以下特征:5 端剪接位点的不同;下游内含子富含嘧啶;下游内含子含有极少量的 GGG 结构区;最接近下游内含子的六聚核苷酸 UGCAUG 被过量表达,同时 UGCAUG 在肌肉特异性外显子组中也被高表达,有趣的是在早先被发现的少数内含子的增强子中 UGCAUG 也广泛存在。他们的分析结果表明,这些元件比很多以往发现的组织内特异性选择的外显子调控元件起到更广泛的作用。

Fairbrother 等人用统计分析方法分析了剪接位点上外显子-内含子的碱基组成,来预测外显子中的外显子剪接增强子(ESE, Exon Splicing Enhancer),ESE是指那些特异的短寡聚核苷酸序列,当这些序列出现在外显子中时可以增强 mRNA 前体的剪接作用,它在组成性剪接和选择性剪接中都起着重要的作用。他们分析了大量的人类基因序列,发现了 10 个 ESEs 基序,并发现所有这 10 个基序在生物体内都有增强子的活性,而且这些序列的点突变会大大降低它的活性,基序的确定可以用于预测人类基因外显子突变的剪接表型<sup>[15]</sup>。

### 2.2 发现和预测选择性剪接基因

高通量的人类基因组序列,特别是表达序列标签(EST)的测序,产生了 1 种完全不同的以生物信息学为基础的选择性剪接研究方法。因为 EST 是从整个 mRNA 群体来的序列标签,它们提供大量不同 mRNA 的信息,这种多样性可以用计算机技术来分析。在过去的两年里,用生物信息学的手段发现了大量的选择性剪接基因,比过去 20 年中发现的要

多得多。目前公共数据库里有代表不同组织的,约 1 500 000 条以上选择性剪接的 ESTs,其中还包括人发育的不同阶段的 ESTs(从胚胎到老年)。Brett 等人建立了一个广泛的人类选择性剪接基因数据库,他们分析了 7 867 条 mRNA 序列,发现其中 3 011 条是选择性剪接基因(占总分量的 38%),从对总的 12 572 条 ESTs 的分析来看,发现可能存在 4 560 条不同可能的选择性剪接。很有趣的是,它们中的 70%是由外显子删除事件引起,而仅有 30%为外显子插入事件引起的。而后他们又用实验加以验证,发现 16 条基因的 19 种选择性剪接产物与疾病有关<sup>[16]</sup>。

随着人类基因组测序的完成,将 ESTs 与基因组序列进行比对成为基因选择性剪接预测中实用而有效的方法。Kan 等人<sup>[17]</sup>设计了一种软件——表达集合程序(TAP),它可以用全基因组与 ESTs 比对的方法描绘基因结构,并可以分析选择性剪接事件。他们还分析了 1 124 条基因,表明一半以上的人类基因是选择性剪接基因,令人惊讶的是,大部分选择性剪接事件影响着编码区,并且用 TAP 软件辅助 ESTs 比对的方法分析人和小鼠选择性剪接的进化保守性。生物信息学的一个重要研究领域就是设计软件辅助解决生物学问题和阐明生物学现象,设计软件发现选择性剪接的位点可以为寻找选择性剪接基因提供实验设计的依据。

Hanke 等人为了寻找新的选择性剪接变体和候选诊断目标,用 TBLASTN 软件分析 475 条与人类疾病相关蛋白质的氨基酸序列,预测其中 162 条可能有 222 个选择性剪接位点,而且还预测出这 162 条序列中有 137 条是选择性剪接的产物<sup>[18]</sup>。Mironov 等人将人的 ESTs 与 392 条已知基因的全基因组进行比对,发现这些基因中有 33 条存在选择性剪接现象<sup>[19]</sup>。Croft 等人从 GenBank 中收集已经注释过的内含子序列建立了内含子序列信息系统 ISIS (intron sequence information system),然后用这些内含子序列寻找与之匹配的 ESTs,结果发现很多 ESTs 与这些内含子匹配,认为这些内含子可能是未被定义的可变外显子,并分析了 2 698 条非冗余的人类基因序列,发现 1 124 条基因(42%)中的 3 119 条 ESTs 包含之前被注释为内含子的序列,认为这些基因都是选择性剪接基因<sup>[20]</sup>。Brett 等人没有用 ESTs 与全基因组比对的方法,而是寻找一些与已知 mRNAs 有关的 ESTs 的插入和缺失作为选择性剪接基因的指示器,发现了 3 011 条选择性剪接基因<sup>[16]</sup>。

国际人类基因组测序联盟用 EST 与全基因组比对的方法分析了 22 号染色体,发现有 145 条选择性剪接基因<sup>[11]</sup>。Modrek 等人将有效的人类 ESTs 和 mRNA 序列(共 2 100 000 条)与全基因组进行比对,运用严格匹配寻找剪接位点和选择性剪接检测规则从 2 272 条基因中鉴别出 6 201 种不同剪接产物<sup>[21]</sup>。

### 2.3 选择性剪接的起源研究

Zhuang 等人<sup>[22]</sup>在对比了 6 种生物(人、老鼠、田鼠、果蝇、线虫和牛)的选择性剪接产物和非选择性剪接产物的氨基酸序列和蛋白长度后,发现选择性剪接产物和非选择性剪接产物都偏爱使用氨基酸 A、E、G、K、L、P、S、V、R、T 和 D,而且选择性剪接产物蛋白平均长度比非选择性剪接产物蛋白平均长度长,他们提出基因的选择性剪接是来源于非选择性剪接基因的 DNA 突变或基因融合。

有大于 10% 的选择性剪接属于置换型的,即只有一段氨基酸序列或另一段氨基酸序列被包含在蛋白序列中,两段氨基酸序列不被同时包在一条蛋白序列中。Kondrashov 等人将全部已确定的多于 9 个氨基酸单位的置换型选择性剪接产物的蛋白序列进行比对分析,提出选择性剪接是在外显子一前一后复制过程中出现的,通过将不同物种同源的关于选择性剪接有效实验数据结合起来分析外显子的复制,计算性的分析表明大部分复制事件至少比哺乳动物产生早甚至比脊椎动物的出现还早,他们认为:外显子一前一后复制只是置换型选择性剪接起源的机制,而其他类型的剪接形式可能有其他不同的起源。连同基因复制一起,外显子一前一后复制可能是多细胞生物进化过程中产生功能分化的主要原因<sup>[23]</sup>。

Kondrashov 等人还将 73 条长度不同的选择性剪接(LDAS, Length Difference Alternative Splicing)基因序列与原核生物或酵母的全序列进行比对,分析了由选择性剪接基因编码蛋白的进化保守性,提出长度不同的选择性剪接起源于序列较长的核苷酸的删除和插入。由于插入性选择性剪接产生的氨基酸序列与其周围序列的保守性一致,由此他们认为选择性剪接的进化是由内含子插入导致产生新功能的蛋白序列引发的<sup>[24]</sup>。

选择性外显子的产生可能通过 DNA 重组(如外显子复制、外显子转座和冗余插入)或内含子序列突变为剪接位点而来的。Sorek 等人最新的研究表明:至少有 5% 的选择性外显子是来自占人类基因的重复序列 10% 的 Alu;而且,单核苷酸突变可以将可变的或沉默的 Alu 转变为组成性的外显子<sup>[25]</sup>,

这种机制可能导致人类疾病<sup>[26]</sup>。人类基因组包含一百万个 Alu,大部分 Alu 位于编码蛋白基因的内含子中,这些位于内含子的 Alu 不仅是新形成的选择性外显子的起源,而且为加速进化铺路<sup>[27]</sup>。

### 3 生物信息学方法与实验方法在研究选择性剪接中的关系

大量确定选择性剪接的实验数据需要用生物信息学的方法来分析,其中最有效的方法就是用较长探针(大于 60 nt)的微阵列来分析。Schoemaker 等人<sup>[28]</sup>用这种技术检测了在人类 22q 染色体上注释的 8 183 个外显子。这种技术很适合去检测选择性剪接,设计一条跨越外显子-外显子连接处的探针,由于给定基因选择性剪接产物会造成不同的外显子-外显子连接处,与不同组织 mRNA 样品杂交就可以检测选择性剪接(图 2)。尽管大部分给定基因的外显子-外显子连接处的杂交率是不变的,但选择性剪接会导致一些连接处的上移或下移。这些芯片的快速出现使从不同组织的选择性剪接基因编写出选择性剪接形式目录成为可能。Affymetrix 公司用 20 种探针(25 nt)代表同一基因的不同外显子,尽管一个基因的大部分探针的强度在不同组织中会有所变化,但某个组织中的特定外显子的探针被不规则地杂交可以指示选择性剪接<sup>[29]</sup>。但要指出的是只用基因芯片的方法,不结合生物信息学分析,是无法解决选择性剪接识别的问题<sup>[30]</sup>。

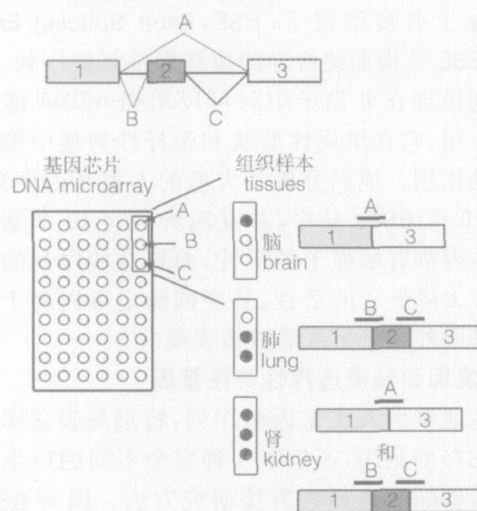


图 2 用基因芯片识别高通量的选择性剪接<sup>[30]</sup>

Fig. 2 High-throughput identification of alternative splicing can be carried out by using microarrays<sup>[30]</sup>

生物信息学一个很大的贡献就是建立数据库,给设计实验提供依据并让生物学家从众多的数据中总结出生物学规律。同时数据库的建立又要以实验结果为依据,如何调节好数据库资源与实验结果的关系十分重要。生物信息学可以推断出一些可靠的功能,在检测已知结构域的添加和删除上,可以推测出用实验如何验证这些结构域是否存在,以及它们

可能导致的疾病和组织特异性。生物学家如果对预测的基因感兴趣,则可以用一些技术(如 PCR, Southern 或 Northern blot)来检验这些预测是否正确。选择性剪接注释程序被当作最好的中心知识库,为生物信息学的预测和以后的实验验证以及功能明确提供服务(图 3),这种程序可以快速发展成为预测和实验之间的桥梁。

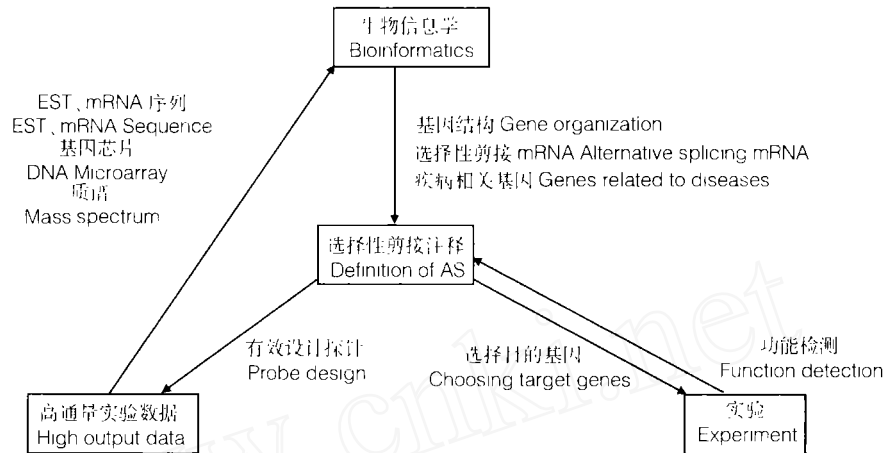


图 3 生物信息学方法与实验方法在选择性剪接注释计划中的关系

对特定基因有兴趣的个人研究者可以从数据库中找到目标基因,并将他们的实验结果提交给数据库。选择性剪接数据库可以帮助设计大量的实验,在用计算机分析实验结果后,数据库也就得到了更新<sup>[30]</sup>。

Fig. 3 Cooperative roles for bioinformatics and experimentation in an alternative splicing annotation project (ASAP)

Individual researchers interested in special genes can find computationally derived alternative splices in the database, allowing them to characterize and report results back into the database. High-throughput experiments can be designed using the database, and after computational analysis of the experimental data, the database can be updated with new results<sup>[30]</sup>.

## 4 展 望

研究选择性剪接对发现致病基因,进而从剪接机制上找到治疗该疾病的方法是未来几年这一研究领域的重点。但由于目前对选择性剪接的研究还在起步阶段,用 ESTs 研究选择性剪接还存在一些争议<sup>[31]</sup>,对选择性剪接的研究还需科学家们共同的努力找到新的突破口。从生物信息学的角度来看,首先是对选择性剪接调控机制的研究,如收集大量选择性剪接基因上游和下游序列,设计软件发现它们的共性,又或者用选择性剪接基因的内含子序列与组成性剪接基因的内含子序列作比较,发现选择性剪接基因内含子的特点,这不仅对选择性剪接调控机制的研究,而且对发现选择性剪接基因也很有帮助;其次就是发现选择性剪接基因并建立选择性剪

接数据库,为研究疾病基因提供数据,同时也丰富了生物数据库的内容;此外,选择性剪接起源和进化的研究,在揭示遗传物质起源和进化方面也有非常重要的作用,由于选择性剪接现象仅在较高等的真核生物中普遍存在,选择性剪接的起源和进化对揭示生物进化的分子机制也不可缺少。

## 参 考 文 献 (References):

- [1] Gilbert W. Why genes in pieces? Nature, 1978, 271 (5645): 501.
- [2] Early P, Rogers J, Davis M, Calame K, Bond M, Wall R, Hood L. Two mRNAs can be produced from a single immunoglobulin mu gene by alternative RNA processing pathways. Cell, 1980, 20(2): 313 ~ 319.
- [3] Rosenfeld M G, Lin C R, Amara S G, Stolarsky L, Roos B A,

- Ong E S, Evans R M. Calcitonin mRNA polymorphism: peptide switching associated with alternative RNA splicing events. *Proc Natl Acad Sci USA*, 1982, 79(6): 1717 ~ 1721.
- [4] Sharp P A. Split genes and RNA splicing. *Cell*, 1994, 77(6): 805 ~ 815.
- [5] Gelfand M S, Dubchak I, Dralyuk I, Zorn M. ASDB: database of alternatively spliced gene. *Nucleic Acids Res*, 1999, 27(1): 301 ~ 302.
- [6] Lopez A J. Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. *Annu Rev Genet*, 1998, 32: 279 ~ 305.
- [7] Boise L H, Gonzalez-Garcia M, Postema C E, Ding L, Lindsten T, Turka L A, Mao X, Nunez G, Thompson C B. *bcl-x*, a *bcl-2*-related gene that functions as a dominant regulator of apoptotic cell death. *Cell*, 1993, 74(4): 597 ~ 608.
- [8] YANG Jir-Chun, CHEN Yir-Hua, MO Jir-Dan, CHEN Cheng. The research progress on alternative splicing of pre-mRNA. *Life Science*, 2002, 4(3): 150 ~ 153.  
延锦春, 陈誉华, 未今丹, 陈澄. mRNA 前体选择性剪接的研究进展. *生命科学*, 2002, 14(3): 150 ~ 153.
- [9] International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*, 2004, 431(7011), 931 ~ 945.
- [10] Modrek B, Resch A, Grasso C, Lee C. Genome-wide analysis of alternative splicing using human expressed sequence data. *Nucleic Acids Research*, 2001, 29(13): 2850 ~ 2859.
- [11] International Human Genome Sequencing Consortium: Initial sequencing and analysis of the human genome. *Nature*, 2001, 409(6822): 860 ~ 921.
- [12] Christopher J L, Irizarry K. Alternative Splicing in the Nervous System: An Emerging Source of Diversity and Regulation. *Biol Psychiatry*, 2003, 54(8): 771 ~ 776.
- [13] Smith C W, Valcarcel J. Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem Sci*, 2000, 25(8): 381 ~ 388.
- [14] Brudno M, Gelfand M S, Spengler S, Zorn M, Dubchak I, Conboy J G. Computational analysis of candidate intron regulatory elements for tissue-specific alternative pre-mRNA splicing. *Nucleic Acids Res*, 2001, 29(11): 2338 ~ 2348.
- [15] Fairbrother W G, Yeh R F, Sharp P A, Burge C B. Predictive Identification of Exonic Splicing Enhancers in Human Genes. *Science*, 2002, 297(9): 1007 ~ 1013.
- [16] Brett D, Hanke J, Lehmann G, Haase S, Delbruck S, Krueger S, Reich J, Bork P. EST comparison indicates 38 % of human mRNAs contain possible alternative splice forms. *FEBS Lett*, 2000, 474(1): 83 ~ 86.
- [17] Kan Z, Rouchka E C, Gish W R, States D J. Gene structure prediction and alternative splicing analysis using genomically aligned ESTs. *Genome Res*, 2001, 11(15): 889 ~ 900.
- [18] Hanke J, Brett D, Zastrow I, Aydin A, Delbruck S, Lehmann G, Luft F, Reich J, Bork P. Alternative splicing of human genes: more the rule than the exception. *Trend Genet*, 2002, 15(10): 389 ~ 390.
- [19] Mironov A A, Fickett J W, Gelfand M S. Frequent alternative splicing of human genes. *Genome Res*, 1999, 9(12): 1288 ~ 1293.
- [20] Croft L, Schandorff S, Clark F, Burrage K, Arctander P, Mattick J S. ISIS, the intron information system, reveals the high frequency of alternative splicing in the human genome. *Nature Genet*, 2000, 24(4): 340 ~ 341.
- [21] Modrek B, Resch A, Grasso C, Lee C. Genome-wide analysis of alternative splicing using human expressed sequence data. *Nucleic Acids Res*, 2001, 29(13): 2850 ~ 2859.
- [22] Zhuang YL, Ma F, Li-Ling J, Xu XF, Li YD. Comparative Analysis of Amino Acid Usage and Protein Length Distribution Between Alternately and Non-alternately Spliced Genes Across Six Eukaryotic Genomes. *Mol Biol Evol*, 2003, 20(12): 1978 ~ 1985.
- [23] Kondrashov F A, Koonin E V. Origin of alternative splicing by tandem exon duplication. *Hum Mol Genet*, 2001, 10(23): 2661 ~ 2669.
- [24] Kondrashov F A, Koonin E V. Evolution of alternative splicing: deletions, insertions and origin of functional parts of proteins from intron sequences. *Trend Genet*, 2003, 19(3): 115 ~ 119.
- [25] Sorek R, Ast G, Graur D. Alu-containing exons are alternatively spliced. *Genome Res*, 2002, 12(7): 1060 ~ 1067.
- [26] Lev-Maor G, Sorek R, Shomron N, Ast G. The birth of an alternatively spliced exon: 30 splice-site selection in Alu exons. *Science*, 2003, 300(5623): 1288 ~ 1291.
- [27] Kneehling J, Graveley B R. The origins and implications of Alternative splicing. *Trend Genet*, 2004, 20(1): 1 ~ 4.
- [28] Shoemaker D D, Schadt E E, Armour C D, He Y D, Garrett-Engel P, McDonagh P D, Loercher P M, Leonardson A, Lum P Y, Cavet G, Wu L F, Altschuler S J, Edwards S, King J, Tsang J S, Schimmack G, Schelter J M, Koch J, Ziman M, Marton M J, Li B, Cundiff P, Ward T, Castle J, Krolewski M, Meyer M R, Mao M, Burchard J, Kidd M J, Dai H, Phillips J W, Linsley P S, Stoughton R, Scherer S, Boguski M S. Experimental annotation of the human genome using microarray technology. *Nature*, 2001, 409(6822): 922 ~ 927.
- [29] Hu G K, Madore S J, Moldover B, Jatkoie T, Balaban D, Thomas J, Wang Y X. Predicting splice variant from DNA chip expression data. *Genome Res*, 2001, 11(7): 1237 ~ 1245.
- [30] Modrek B, Lee C. A genomic view of alternative splicing. *Nature Genet*, 2002, 30(1): 13 ~ 19.
- [31] MENG Xian-Fang, ZHOU Guang-Yun. The Alternative Splicing of Pre-mRNA. *Chemistry of Life*, 2001, 21(4): 271 ~ 273.  
孟宪芳, 周光云. mRNA 前体的选择性剪接. *生命的化学*, 2001, 21(4): 271 ~ 273.