

DOI: 10.3724/SP.J.1005.2008.00257

基于 EST 的新基因克隆策略

刘媛¹, 蔡嘉斌¹, 蒋国松², 童强松¹

1. 华中科技大学同济医学院附属协和医院小儿外科, 武汉 430022;
2. 华中科技大学同济医学院附属协和医院泌尿外科, 武汉 430022

摘要: 表达序列标签(expressed sequence tags, EST) 是从随机选择的 cDNA 克隆进行单向测序获得的短的 cDNA 序列, 代表一个完整基因的一部分。随着生物信息学和基因定位的迅猛发展, EST 已成为基因定位、基因克隆、基因表达分析的有力工具。近年来, 由于 EST 数据库的迅速扩张, 运用 EST 来克隆和定位基因, 使得新基因克隆的策略发生了革命性变革。尽管存在一些不足, 实践证明 EST 可大大加速新基因的发现与研究。本文将就 EST 技术尤其是它在新基因克隆中的应用策略作详细介绍。

关键词: 表达序列标签; 基因克隆; 基因

Strategies for cloning of novel genes based on EST

LIU Yuan¹, CAI Jia-Bin¹, JIANG Guo-Song², TONG Qiang-Song¹

1. Department of Pediatric Surgery, Union Hospital Affiliated to Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China
2. Department of Urology, Union Hospital Affiliated to Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430022, China

Abstract: Expressed sequence tags (EST) are short, randomly selected single-pass nucleotide sequence reads derived from cDNA libraries and represent a small part of a gene. Along with the development of bioinformatics and genetic localization, EST has already become a powerful tool for mapping, cloning and expression profiling of genes. Recently, because of the fast distension of EST databases, application of EST in gene mapping and cloning leads to revolutionary change in the strategies for cloning of novel genes. Despite of some insufficiencies, it has been proved that EST could promote the discovery and research of novel genes. In this article, an introduction about EST, especially EST-based strategies for cloning of novel genes will be given in details.

Keywords: expressed sequence tags; gene cloning; gene

表达序列标签(expressed sequence tags, EST)是从一个随机选择的cDNA 克隆进行 5'端和 3'端单次测序获得的短的cDNA 部分序列, 代表一个完整基因的一部分。这一概念由Adams等^[1]于 1991 年首次提出。随着人类基因组计划的开展, 生物学家 Venter JC 于 1991 年提出了发现表达基因的新战略,

即用EST克隆和定位新基因^[1], 使得新基因克隆的策略发生了革命性的变革^[2,3]。本文将就EST技术尤其是它在新基因克隆中的应用策略作一阐述。

1 EST 概述

随着人类基因组大规模cDNA测序计划的开展,

收稿日期: 2007-08-19; 修回日期: 2007-09-02

基金项目: 国家自然科学基金资助项目(编号: 30200284, 30600278) [Supported by National Natural Science Foundation of China (No. 30200284, 30600278)]

作者简介: 刘媛(1982-), 女, 广西人, 硕士研究生, 研究方向: 小儿泌尿系发育候选基因筛选。Tel: 027-85726197; E-mail: liuliurachel@yahoo.com.cn
通讯作者: 童强松(1974-), 男, 湖北人, 博士后, 副教授, 研究方向: 小儿泌尿系发育及调控。Tel: 027-63776478; E-mail: qs_tong@hotmail.com

Adams等^[2]最先提出了EST的概念。所谓EST即从一个随机选择的cDNA克隆进行5'端和3'端单一次测序获得的短的cDNA部分序列,其长度一般从20到7 000 bp不等,平均长度为 360 ± 120 bp。每个EST代表一个完整基因的一小部分。在人类庞大的基因组中表达基因仅5万到10万个,而EST正是这些表达基因的部分cDNA片段。1995年, Boguski和Schuler首先构思以EST为界标的人类基因组转录图谱计划,这使科学家们在人类基因组测序即将完成之际,提前进入对基因的功能研究。美国国家癌症研究中心(National Cancer Institute)的癌基因组剖析计划(Cancer Genome Anatomy Project)也开始存入从恶性肿瘤组织中发现的EST。截至2006年2月,dbEST数据库已拥有从559种生物体获得的32 889 225条EST,几乎涵盖了GenBank中3千万个已报导基因序列的70%^[4]。一个基因转录本的cDNA序列可能包含有多个序列重叠的EST,而一个基因mRNA剪接点不同可以获得多个cDNA克隆,因此EST既可能对应于一个cDNA的某一部分,又可能代表mRNA的不同剪接方式。此外,由于EST来源于一定环境下一个组织总mRNA所构建的cDNA文库,故EST也能说明该组织中各基因的表达水平及特异性。EST自身的局限性集中体现在序列分析中的精确程度,另外还有由高丰度与低丰度序列之间的巨大差异造成的重复测定与漏测^[5-7]。

2 EST 的应用

2.1 基因识别

EST最早也最常用于基因识别^[8]。基因识别即使用生物学实验或计算机等手段识别DNA序列上的具有生物学特征的片段。由于EST代表一个完整基因的一部分,以其为查询序列在基因组内进行同源性分析,可快速地识别基因。Medzhitov等^[9]通过果蝇黑胃TOLL蛋白进行dbEST数据库检索,该蛋白已证实成熟果蝇抗真菌反应中发挥重要作用^[10],通过同源分析的方法,找到相应的人类同源EST,这为接下来研究人类TOLL同源蛋白的功能提供了很好的基础。

2.2 物理图谱的构建

物理图谱是指有关构成基因组的全部基因的排列和间距的信息,它是通过对构成基因组的DNA分子进行测定而绘制的。EST本身来源于cDNA的表达

序列,因编码DNA序列高度保守而具有自身的特殊性质,利用其作为遗传标记可以真实并直接地反映基因在染色体上的位置,与来自非表达序列的标记,如扩增片段长度多态性(amplified restriction fragment polymorphism, AFLP)、随机扩增多态性DNA(random amplified polymorphic DNA, RAPD)以及微卫星标记(simple sequence repeats, SSR)等相比,更可能突破家系与物种的限制,因此EST标记在亲缘关系较远的物种间比较基因组连锁图和比较质量性状信息方面具有优势;同样,对于一个DNA序列缺乏的目标物种,来源于其他物种的EST也能用于该物种有益基因的遗传作图,加速物种间相关信息的迅速转化,因此有着其它标记方法无法比拟的优越性^[11,12]。邵筱等^[13]应用EST及电子克隆策略研究了血吸虫的表达基因谱。Tuggle等^[14]将EST应用于鉴别基因在不同组织或细胞类型间的表达差异以及人类与猪比较物理图谱的构建。

2.3 “电子”基因克隆

“电子”基因克隆(*silico cloning*),即采用生物信息学的方法延伸EST序列,以获得基因部分乃至全长的cDNA序列。EST数据库的迅速扩张,使利用其来识别与克隆新基因迅猛发展起来。陈大福等^[15]以果蝇腺苷酸转移酶基因cDNA序列作为模板,对家蚕EST数据库进行同源检索筛选,克隆出家蚕腺苷酸转移酶基因的cDNA序列。通过EST分析,还可用于预测基因结构^[16,17],挖掘mRNA加工和成熟的分子机制(如mRNA加工信号元件、可变剪接和3'末端加工的多样性等)^[18-20]及存在于同一物种个体之间的单核苷酸多态性(single nucleotide polymorphism, SNP)信息^[21,22]。近来,将EST结合其他技术还可用于蛋白的鉴定,Edwards^[23]运用EST及序列数据库压缩的方法找到了一种迅速而有效鉴别蛋白的新方法。正是由于EST表现出了这些巨大潜能,使其得到了充分的利用与发展。

3 基于 EST 的新基因克隆策略

利用EST进行电子基因克隆的具体步骤可总结为:首先,用已知的EST序列为起始序列,采用BLAST(Basic Local Alignment Search Tool)程序检索数据库中与其同源或有部分重叠的EST序列,以确定哪些代表已知基因,哪些代表已知EST,哪些代表新发现的EST序列;还可以多个EST序列电子

延伸组装为基序(contigs),即通过与已知的 EST 序列重叠区(overlapping sequence)的配对延伸,组装成毗邻序列;以此基序为被检序列再进行 BLAST 检索;重复以上过程,直至没有更多的重叠 EST 检出,基序不能继续延伸为止,由此可以获得尽可能长或全长的 cDNA 序列。而应用 EST 来克隆新基因,关键在于感兴趣 EST 序列的选取和获得,以下将着重介绍这两个方面。

3.1 感兴趣 EST 序列的选取方法

3.1.1 选取用定位候选克隆策略得到的疾病相关基因的 EST

定位候选克隆策略是在定位克隆的基础上部分融合了功能克隆的策略。先以基因的染色体定位信息和相应的染色体区段物理图和遗传图为基础,在分子生物学信息网上查询这一区域内所有的已知基因EST和部分cDNA序列作为候选基因,然后在病变组织或细胞中运用常规分子生物学方法检测这些候选基因的改变,从而迅速筛选出致病基因的方法。由于EST可能是某一细胞或组织中大量未知基因的部分表达序列,故通过定位候选克隆策略选取EST可以方便快捷地发现新的疾病相关基因。阳剑波等^[24]即应用该策略筛选出在鼻咽癌细胞株HNE1和鼻咽癌活检组织中表达下调的、定位于共同缺失区内的3'末端EST,克隆出9p21-22区域内鼻咽癌候选抑瘤基因。

3.1.2 选取模型生物基因组研究成果发现功能相近基因的 EST

生物在进化过程中保持有某些功能基因的保守性,也即同源性,因此我们可以从各种模型生物基因组研究中获取精确信息从而对人类可能存在的相应功能基因的克隆提供极有价值的线索,并成为人类基因组研究的重要辅助手段之一。郭丽丽等^[25]运用这一策略克隆出人类新的锌指蛋白基因ZNF18。Li等^[26]利用随机纯合子剔除技术在小鼠纤维母细胞中发现了抑瘤基因*tsg101*,将该基因的cDNA全序列输入dbEST数据库中进行BLAST分析,得到10个与小鼠*tsg101*基因高度同源的人类EST,对这些EST进行组装、克隆和测序,最终获得一个人类抑瘤基因*TSG101*。因此,选取模型生物基因组研究成果发现的功能基因的EST作为感兴趣EST,有可能发现人类基因组中功能相近的新基因。

3.1.3 选取基因家族同源性分析得到的 EST

基因家族是指一组具有相似或相关外显子的基因,它们具有高度相似的cDNA序列和基因表达产

物。以基因家族某个已知基因序列为查询链,运用BLAST程序进行同源性分析获得感兴趣EST,然后进行电子克隆,可以快速、经济地获得基因家族的新成员。通过这种方法,Hardiman等^[27]成功克隆了人类WNT基因家族的新成员*WNT10B*,而Still等^[28]则获得了*TACC*家族另外两个新成员*TACC2*和*TACC3*基因的全序列。

3.1.4 选取差异表达或协同表达基因的 EST

EST来源于一定环境下一个组织总mRNA所构建的cDNA文库,故EST能说明该组织中各基因的表达水平及特异性。基于此特性,结合mRNA差异显示、cDNA代表差异分析及抑制性消减杂交等技术,可从快速获取的差异表达cDNA片段筛选出存在差异表达的未知基因片段,为克隆新的差异表达基因提供依据。Aouacheria等^[29]利用网络资源,假定某EST上调或下调均可能致肿瘤,由此指定正常组织和肿瘤组织中的基因差异表达EST文库,并将所得肿瘤组织EST文库按组织特异性分组。对组织特异性肿瘤EST进行电子基因克隆,可能得到相应组织新的致癌基因。Wu等^[30]则从协同表达的基因着手,构建了一个能够实现GBA(Guilt-by-Association)协同表达运算的网络服务系统:输入一条与某疾病相关的已知基因序列或一条新的EST序列,能够找到与其协同表达的已知或未知基因的部分序列,从而可能得到新基因。

3.2 感兴趣 EST 的具体获取途径

3.2.1 通过 EST 数据库中寻找

目前,在互联网上存在有很多公共的EST数据库,常用的EST数据库包括:NCBI(National Center for Biotechnology Information) GenBank 的dbEST^[31]、意大利Tigem的EST machine(包括EST提取者和EST组装机器)、THC(Tentative Human Consensus Sequences)数据库。而找出EST的最有效途径是寻找同源序列,其选取标准是:长度 100 bp,同源性 50%以上、85%以下。这一过程可通过使用BLAST检索程序实现。

3.2.2 通过网上的二级 EST 数据库获取

随着EST技术广泛深入的发展,很多研究组织也根据自己的研究领域,构建出该领域内专门的EST数据库,或是联合运用多种工具或综合多个数据库构建出二级数据库,以便我们能从特异性数据库中获取信息,这两种途径均使得我们能更快更好的寻找和获得感兴趣EST。Lazzari等^[32]综合了GenBank nr 数据库、UniProt KB数据库和室内培养

真菌基因组数据库构建了ESTuber数据库, 让我们能够方便快捷的找到勃氏块菌相关EST。Chen等^[33]结合分析EST Unigene数据库信息和Ensembl数据库同源关系后构建了一个ZooDDD数据库。这一数据库使我们能够找到某单一种属差异表达基因转录子, 并为我们比较研究动物基因组的功能、调节

和进化以及寻找疾病相关新基因或差异表达的未知基因提供了一个很好的平台。表 1^[34]显示了更多可获取的EST资源。当得到了感兴趣EST之后就可以按照电子基因克隆的一般方法, 得到完整的基因序列。

表 1 EST 数据资源

Table 1 Resources for EST data

资源 Resource	网 址 Web site
ApiEST-DB	http://www.cbil.upenn.edu/apidots/
dbESTat NCBI	http://www.ncbi.nlm.nih.gov/dbEST/
Diatom EST database	http://avesthagen.sznbowler.com .
ESTree	http://www.itb.cnr.it/estree/
Fungal genomics project	https://fungalignomics.concordia.ca/home/index.php
Honey bee brain EST project	http://titan.biotec.uiuc.edu/bee/honeybee_project.htm
Nematode ESTs at the Sanger Institute	ftp://ftp.sanger.ac.uk/pub/pathogens/nem_ests/
NEMBASE-parasitic nematode ESTs	http://www.nematodes.org
Parasitic and free-living nematode EST resource	http://www.nematode.net/
Phytopathogenic Fungi and Oomycete EST database	http://cbr-rbc.nrc-cnrc.gc.ca/services/cogeme/
Plant Gene Research, Kazusa DNA Research Institute	http://www.kazusa.or.jp/en/plant/database.html
Plant Genome database	http://www.plantgdb.org/
Rat EST data at University of Iowa	http://ratest.eng.uiowa.edu
Sanger Institute Xenopus tropicalis EST project	http://www.sanger.ac.uk/Projects/X_tropicalis/
The TIGR Gene Indices	http://www.tigr.org/tdb/tgi/
UniGene database at NCBI	http://www.ncbi.nlm.nih.gov/UniGene

4 EST 应用中应注意的问题

4.1 提高预测的精确度

由于mRNA在不同时空的丰度差异、反转录酶的误配、同功密码子的存在、电子杂交对碱基错配等原因, 常造成预测的精度较低。为了克服这些问题, 在序列分析时应用多种综合软件如GeneScan、Phred、Polyphred等, 使候选序列同见于 2 个或 2 个以上的文库, 以有效排除假阳性结果^[35]。

4.2 减少冗余

4.2.1 避免重测和漏测

由于在对cDNA克隆随机测序时, 高丰度的基因容易被反复测序, 而低丰度的基因则有可能漏测^[6,7], 故需建立标准化的cDNA文库, 使cDNA文库中基因出现的频率基本一致, 可采用以下 4 项措施: (1) 建立差减文库, 扣除管家基因, 使所研究的差异基因富集; (2) 用出现频率高的基因作探针, 文库杂交筛选后再测序; (3) 用mRNA差异显示技术获得特异性片段, 测序分析; (4) 在EST数据库中剔

除冗余重复序列^[36]。

4.2.2 剔除非正常转录序列

由于构建cDNA文库过程中可能引入种种污染, 如载体序列等, 应用EST数据时必须将其剔除。通过应用本地BLAST等程序将EST与非冗余载体数据库(如UniVec 和EMVEC)比较即可识别出载体序列并将其剔除^[34]。

4.3 注意实验过程和描述的标准化

由于核苷酸测序存在的 1%~3% 碱基错误^[37], 计算机软件和操作中限定匹配条件的不同, 运用EST电子克隆获得的新基因必须通过进一步的鉴定才能证实其存在及功能。这就要求注意实验过程和描述的标准化, 以建立标准注释的公共数据库, 使实验数据具有更准确的可比性。而标准化的基因表达数据库要有以下 5 类必要的信息: 联系信息、杂交靶探针信息、杂交样本、细胞类型和组织来源要用标准语言描述、mRNA转录定量及统计学意义^[36]。研究者应注意实验描述的标准化, 使研究结果快速登陆数据库。

5 展望

由于 EST 数据库中包含着丰富的分子生物学信息, 使其在新基因克隆中展现出明显的优势。随着基因组学的研究从结构基因组学过渡到功能基因组学, EST 作为基因组学与功能基因组学及后基因组学之间的桥梁, 将为大规模基因识别、克隆和表达分析提供空前的动力并起到更加显著的作用。

参考文献(References):

- [1] Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merrill CR, Wu A, Olde B, Moreno RF. Complementary DNA sequencing: expressed tags and human genome project. *Science*, 1991, 252(5013): 1651–1656. [\[DOI\]](#)
- [2] Collins FS, Patrinos A, Jordan E, Chakravarti A, Gesteland R, Walters L. New goals for the U.S. human genome project: 1998–2003. *Science*, 1998, 282(5389): 682–689. [\[DOI\]](#)
- [3] Hattori M, Tsukahara F, Furuhashi Y, Tanahashi H, Hirose M, Saito M, Tsukuni S, Sakaki Y. A novel method for making nested deletion and its application for sequencing of a 300 kb region of human APP locus. *Nucleic Acids Res*, 1997, 25(9): 1802–1808. [\[DOI\]](#)
- [4] Nagaraj SH, Gasser RB, Ranganathan S. A hitchhiker's guide to expressed sequence tag (EST) analysis. *Brief Bioinform*, 2007, 8(1): 6–21. [\[DOI\]](#)
- [5] Bonaldo MF, Lennon G, Soares MB. Normalization and subtraction: two approaches to facilitate gene discovery. *Genome Res*, 1996, 6(9): 791–806. [\[DOI\]](#)
- [6] Yamamoto K, Sasaki T. Large-scale EST sequencing in rice. *Plant Molecular Biology*, 1997, 35 (1): 135–144. [\[DOI\]](#)
- [7] Hoog C. Isolation of a large number of novel screening strategy. *Nucleic Acids Res*, 1991, 19 (22): 6123–6127. [\[DOI\]](#)
- [8] Wilcox AS, Khan AS, Hopkins JA, Sikela JM. Use of 30 untranslated sequences of human cDNAs for rapid chromosome assignment and conversion to STSs: implications for an expression map of the genome. *Nucleic Acids Res*, 1991, 19(8): 1837–1843. [\[DOI\]](#)
- [9] Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA. The dorsoventral regulatory gene cassette *spätzle/Toll/cactus* controls the potent antifungal response in drosophila adults. *Cell*, 1996, 86(6): 973–983. [\[DOI\]](#)
- [10] Her C, Doggett NA. Cloning, structural characterization, and chromosomal localization of the human orthologue of *saccharomyces cerevisiae MSH5* gene. *Genomics*, 1998, 52(1): 50–61. [\[DOI\]](#)
- [11] Schuler GD, Boguski MS, Stewart EA, Stein LD, Gyapay G, Rice K, White RE, Rodriguez-Tomé P, Aggarwal A, Bajorek E, Bentolila S, Birren BB, Butler A, Castle AB, Chiannilkulchai N, Chu A, Clee C, Cowles S, Day PJ, Dibling T, Drouot N, Dunham I, Duprat S, East C, Edwards C, Fan JB, Fang N, Fizames C, Garrett C, Green L, Hadley D, Harris M, Harrison P, Brady S, Hicks A, Holloway E, Hui L, Hussain S, Louis-Dit-Sully C, Ma J, MacGilvery A, Mader C, Maratukulam A, Matise TC, McKusick KB, Morissette J, Mungall A, Muselet D, Nusbaum HC, Page DC, Peck A, Perkins S, Piercy M, Qin F, Quackenbush J, Ranby S, Reif T, Rozen S, Sanders C, She X, Silva J, Slonim DK, Soderlund C, Sun WL, Tabar P, Thangarajah T, Vega-Czarny N, Vollrath D, Voyticky S, Wilmer T, Wu X, Adams MD, Auffray C, Walter NA, Brandon R, Dehejia A, Goodfellow PN, Houlgatte R, Hudson JR Jr, Ide SE, Iorio KR, Lee WY, Seki N, Nagase T, Ishikawa K, Nomura N, Phillips C, Polymeropoulos MH, Sandusky M, Schmitt K, Berry R, Swanson K, Torres R, Venter JC, Sikela JM, Beckmann JS, Weissenbach J, Myers RM, Cox DR, James MR, Bentley D, Deloukas P, Lander ES, Hudson TJ. A gene map of the human genome. *Science*, 1996, 274(5287): 540–546.
- [12] Agyare FD, Lashkari DA, Lagos A, Namath AF, Lagos G, Davis RW, Lemieux B. Mapping expressed sequence tag sites on yeast artificial chromosome clones of *Arabidopsis thaliana* DNA. *Genome Res*, 1997, 7(1): 1–9. [\[DOI\]](#)
- [13] SHAO Xiao, WU Zhong-Dao, LIU Han-Teng, ZOU Sai-De, YU Xin-Bing. Application of EST and *silico* clone for the analysis of the gene expression profile of *Schistosoma japonicum*. *Bas Clin Med*, 2005, 25(7): 602–606.
邵筱, 吴忠道, 刘翰腾, 邹赛德, 余新炳. 应用 EST 和电子克隆策略研究血吸虫表达基因谱. *基础医学与临床*, 2005, 25(7): 602–606.
- [14] Tuggle CK, Wang Y, Couture O. Advances in swine transcriptomics. *Int J Biol Sci*, 2007, 3(3): 132–152.
- [15] CHEN Da-Fu, NIU Bao-Long, WENG Hong-Biao, MENG Zhi-Qi, LV Shun-Lin. Cloning of adenine nucleotide translocase gene from *Bombyx mori* based on EST database. *Sci Sericulture*, 2004, 30(2): 151–155.
陈大福, 牛宝龙, 翁宏彪, 孟智启, 吕顺霖. 利用 EST 库资源克隆家蚕腺苷酸转移酶基因. *蚕业科学*, 2004, 30(2): 151–155.
- [16] Kan Z, Rouchka EC, Gish WR, States DJ. Gene structure prediction and alternative splicing analysis using genomically aligned ESTs. *Genome Res*, 2001, 11(5): 889–900. [\[DOI\]](#)
- [17] Jiang J, Jacob HJ. EbEST: an automated tool using expressed sequence tags to delineate gene structure. *Genome Res*, 1998, 8(3): 268–275.
- [18] 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. [\[DOI\]](#)

- [19] Modrek B, Resch A, Grasso C, Lee C. Genome-wide detection of alternative splicing in expressed sequences of human genes. *Nucleic Acids Res*, 2001, 29(13): 2850–2859. [\[DOI\]](#)
- [20] Kan Z, Castle J, Johnson JM, Tsinores NF. Detection of novel splice forms in human and mouse using cross-species approach. *Pac Symp Biocomput*, 2004: 42–53.
- [21] Buetow KH, Edmonson MN, Cassidy AB. Reliable identification of large numbers of candidate SNPs from public EST data. *Nat Genet*, 1999, 21(3): 323–325. [\[DOI\]](#)
- [22] Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA, Boyce-Jacino M. Mining SNPs from EST databases. *Genome Res*, 1999, 9(2): 167–174.
- [23] Edwards NJ. Novel peptide identification from tandem mass spectra using ESTs and sequence database compression. *Mol Syst Biol*, 2007, 3: 102.
- [24] YANG Jian-Bo, ZHANG Bi-Cheng, QIAN Jun, ZHANG Xiao-Hui, LI Zhong-Hua, BIN Liang-Hua. Refined localization and cloning of a novel putative tumor suppressor gene associated with nasopharyngeal carcinoma on chromosome 9p21-22. *Chin J Cancer*, 2000, 19(1): 6–9.
阳剑波, 张必成, 钱骏, 张小慧, 李忠花, 宾亮华. 精细定位和克隆 9p21-22 区域内鼻咽癌候选抑瘤基因. *癌症*, 2000, 19(1): 6–9.
- [25] GUO Li-Li, CI Hong-Liang, SHAN Hong-Shuang, ZOU Xing, ZHAI Yong-Gong, LI Yi-Ping. Molecular cloning and expression analysis of a novel human gene *ZNF18*. *Hereditas (Beijing)*, 2005, 27(4): 523–530.
郭丽丽, 慈宏亮, 单红爽, 邹星, 翟永功, 李亦平. 人类锌指结构新基因 *ZNF18* 的克隆和表达谱分析. *遗传*, 2005, 27(4): 523–530.
- [26] Li L, Cohen SN. Tsg101: A novel tumor susceptibility gene isolated by controlled homozygous functional knockout of allelic loci in mammalian cells. *Cell*, 1996, 85(3): 319–329. [\[DOI\]](#)
- [27] Hardiman G, Kastelein RA, Bazan JF. Isolation, characterization and chromosomal localization of human WNT10B. *Cytogenet Cell Genet*, 1997, 77(3-4): 278–282.
- [28] Still IH, Hamilton M, Vince P, Wolfman A, Cowell JK. Cloning of TACC1, an embryonically expressed, potentially transforming coiled coil containing gene, from the 8p11 breast cancer amplicon. *Oncogene*, 1999, 18(27): 4032–4038. [\[DOI\]](#)
- [29] Aouacheria A, Navratil V, Barthelaix A, Mouchiroud D, Gautier C. Bioinformatic screening of human ESTs for differentially expressed genes in normal and tumor tissues. *BMC Genomics*, 2006, 7: 94. [\[DOI\]](#)
- [30] Wu X, Walker MG, Luo J, Wei L. GBA server: EST-based digital gene expression profiling. *Nucleic Acids Res*, 2005, 33(web server issue): W673–W676. [\[DOI\]](#)
- [31] Boguski MS, Schuler GD. ESTablishing a human transcript map. *Nat Genet*, 1995, 10(4): 369–371. [\[DOI\]](#)
- [32] Lazzari B, Caprera A, Cosentino C, Stella A, Milanese L, Viotti A. ESTuber db: an online database for Tuber borchii EST sequences. *BMC Bioinformatics*, 2007, 8(Suppl.1): S13. [\[DOI\]](#)
- [33] Chen YC, Hsiao CD, Lin WD, Hu CM, Hwang PP, Ho JM. ZooDDD: a cross-species database for digital differential display analysis. *Bioinformatics*, 2006, 22(17): 2180–2182. [\[DOI\]](#)
- [34] Nagaraj SH, Gasser RB, Ranganathan S. A hitchhiker's guide to expressed sequence tag (EST) analysis. *Brief Bioinform*, 2007, 8(1): 6–21. [\[DOI\]](#)
- [35] CHEN Wei, ZHANG Ge, ZHANG Si-Zhong. Discovery of candidate SNP by bioinformatic methods. *Hereditas (Beijing)*, 2001, 23(2): 153–156.
陈伟, 张戈, 张思仲. 基于生物信息学的 SNP 候选位点搜寻方法. *遗传*, 2001, 23(2): 153–156.
- [36] YANG Ke-Qiang, WANG Yue-Jin, ZHANG Jin-Jin, WANG Xi-Ping, ZHANG Jian-Xia, WAN Yi-Zhen. Research advance of gene cloning and gene expression analysis based on EST. *J Northwest A & F Univ (Nat Sci Ed)*, 2002, 30(4): 141–145.
杨克强, 王跃进, 张今今, 王西平, 张剑侠, 万怡震. 基于表达序列标签(EST)的基因克隆基因表达分析研究进展. *西北农林科技大学学报*, 2002, 30(4): 141–145.
- [37] Hillier LD, Lennon G, Becker M, Bonaldo MF, Chiapelli B, Chissoe S, Dietrich N, DuBuque T, Favello A, Gish W, Hawkins M, Hultman M, Kucaba T, Lacy M, Le M, Le N, Mardis E, Moore B, Morris M, Parsons J, Prange C, Rifkin L, Rohlfing T, Schellenberg K, Marra M. Generation and analysis of 280,000 human expressed sequence tags. *Genome Res*, 1996, 6(9): 807–828. [\[DOI\]](#)