

新闻动态

SNP Array 6.0; Gene 1.0 ST,
i-Array项目; 昂飞专利胜诉

专题综述

GeneChip®生物芯片技术平台

技术介绍

GeneChip®表达谱芯片技术及应用

新技术与新启示

mRNA的生成; 选择性剪接和

全转录本表达谱芯片

实验指南

表达谱芯片实验和数据分析流程

应用案例

癌症分子分型; 发病机理

外显子跳跃; 植物基因组研究

本期相关技术要点

3' 表达谱芯片到全转录本表达谱芯片



GeneChip®通讯是由美国昂飞公司(Affymetrix)上海代表处主办的季刊, 我们希望用通讯作为一种交流途径和中国科技工作者共同探讨基因组学工具, 特别是生物芯片在科学研究和技术创新的应用。

正如大家熟知, 人类基因组计划(Human Genome Project)的成功和其他生物种基因组序列图绘的累积, 使得我们今天有效地在基因组水平或聚焦于特定的一组基因去探索以最终获得对生物系统更全面的理解, 并迅速应用到生物医学、农业科技、环境

科学等领域创新。生物芯片以其微型化、自动化、高通量、可靠性等特点, 日益成为基因组学研究的最主要手段之一。

美国昂飞公司(Affymetrix)目前是最先进的生物芯片(GeneChip®)系列为代表的生物芯片发源地。其产品已被国际上绝大部分领先的与生命科学有关的研究机构和政府部门所采用, 成为各个领域创新的核心工具之一, 并且几乎所有的世界著名医药公司将其广泛应用于药物开发的各个环节中。昂飞现已拥有国际性专利近400项。用昂飞芯片发表的科学文章已超过8500多篇, 而且许多发表在Science、Nature、Cell等世界最高级别的学术刊物上。

我们将努力使GeneChip®通讯成为向读者提供最新芯片技术和有关信息的有效工具, 并真诚地祝愿通讯作出贡献有助于加强昂飞和广大读者的科技合作, 为促进中国生命科学研究和生物科技产业的发展作出贡献。

新闻动态

· 至今世界上规模最大的全基因组疾病关联分析项目

2007年6月7日, Wellcome Trust Case Control Consortium刚完成了至今世界上规模最大的全基因组关联分析项目。此项目证实了全基因组关联分析是有效的手段去发现导致复杂疾病的基因。项目用昂飞500K芯片共分析了具有七种常见病(双极神经障碍, 冠心病, 克隆氏病, 高血压, 风湿性关节炎, I型糖尿病和II型糖尿病) 14, 000个个体和3, 000个共享的对照个体。详细报道, 请看Nature6月7日期第661页。

· 目前世界上信息量最大的基因芯片: 人类全基因组SNP 6.0芯片

2007年5月21日, 昂飞公司正式推出人类全基因组SNP6.0芯片。单张芯片可以含有同时检测超过1, 800, 000个遗传变异(SNP和拷贝数的变异)的生物标记, 其数量超过市场上其他同类产品的三倍, 并且在成本上具有优势, 从而能增加发现同药物反应及复杂疾病相关基因的可能, 为广大生命科学研究者提供了强有力的工具。

SNP6.0芯片能检测900, 600个SNP, 946, 000个用于检测拷贝数变异的非多态性探针。通过与哈佛大学合办的Broad研究所合作, SNP6.0芯片在准确性和一致性方面达到了新的高度。相应推出的Genotyping Console用来对处理SNP6.0芯片数据和全基因组遗传分析及质量控制。全基因组采样分析(WGSA)已应用于全世界多于200个实验室, 并有300多篇相应的论文。SNP6.0芯片是继Mapping10k、100k、500k和SNP5.0芯片后, 应用WGSA技术的最新的芯片。本刊下期提供详细介绍。详细信息请访问www.affymetrix.com网站或咨询昂飞公司上海代表处。

· Affymetrix (美国昂飞公司)专利胜诉Illumina, 请看新闻原文转载: (<http://www.marketwatch.com/news/story/>)

ILLUMINA TO PAY \$16.7M IN DAMAGES TO AFFYMETRIX

By Monica M. Clark

Last Update: 3:42 PM ET Mar 13, 2007

ILLUMINA INC. (ILMN32.04, -0.59, -1.8%) WAS ORDERED TO PAY ABOUT \$16.7 MILLION TO AFFYMETRIX INC. (AFFX) FOR INFRINGING UPON FIVE PATENTS. ILLUMINA, A SAN DIEGO BIOTECHNOLOGY COMPANY, INTENDS TO APPEAL THE INFRINGEMENT FINDING OF THE FEDERAL JURY IN WILMINGTON, DELAWARE. THE \$16.7 MILLION IN DAMAGES IS BASED ON A ROYALTY OF 15% FOR CERTAIN PRODUCTS LAUNCHED BEFORE AND SOLD BY ILLUMINA THROUGH 2005, THE COMPANY SAID. ILLUMINA SHARES WERE RECENTLY TRADING DOWN \$2.44, OR 7.8%, AT \$28.89

· “全转录本”策略设计，具价格优势的表达芯片：人类基因组Gene 1.0 ST芯片

2007年4月3日，昂飞公司正式推出人类Gene1.0ST芯片。不象3'端表达谱芯片一样专注于基因3'端附近序列表达情况的检测，Gene1.0ST芯片能对一条基因的所有转录产物进行整体表达分析。由于其成本优势以及简化的实验流程让更多的科学家可以在日常研究中能采用生物芯片这一工具。详细信息，请参考本期内容。

· i-Array 项目 -- 掌控您的实验

您想在您自己的实验室里轻松完成所有的芯片实验(芯片扫描除外)及数据处理工作吗？昂飞为您提供了一个新的选择——i-Array 项目，它可以 1) 使科学工作者有能力建立并掌握国际一流的基因芯片技术，并将其运用于日常的分子生物学研究中；2) 使科学家能掌控自己的实验，确保样品的安全性及实验思路的保密性，杜绝目前由于服务过程中产生的常见纠纷；3) 有利于实验室建设及人才培养；4) 大大降低使用基因芯片技术的前期投入，灵活地进行仪器配置。项目将于今年7月开始，详细信息请垂询代理商或直接致电昂飞上海办事处。

专题综述

本期介绍昂飞公司GeneChip®生物芯片技术平台。作为一整套遗传研究的工具，该平台由高密度生物芯片以及帮助处理和分析生物芯片的工具组成，包括标准化的生物芯片、实验试剂、仪器、数据管理和分析工具。全世界超过有1500个昂飞的技术平台系统，是世界上第一的经欧盟及美国食品和药物管理局(FDA)审批的可用作体外诊断的芯片系统。

芯片设计及种类

- 应用光刻技术和严格的流程控制原位合成高密度芯片，目前已达到约四百万探针/cm²增加密度的空间还很大。
- 使用多个探针来检测转录本或SNP，有效地减少探针杂交非专一性的影响，并通过合适的算法获得更为有力的数据。
- 在序列筛选、芯片设计和质量控制上确保芯片数据的准确性和重复性。在芯片制作过程中不采纳“随机”的原则，每块同类芯片都具有高均一性。
- 芯片物种齐全，几乎所有代表性生物都能找到合适的昂飞表达谱芯片。种类繁多，可用于整个转录组基因功能、另型剪接或转录调控(如ChIP-on-chip)的研究；全基因组关联研究，特定一组基因的定向基因分型或重测序。表达谱和重测序以及SNP芯片皆可自创定制。

仪器及实验检测

- 自动化的实验操作仪器保证了不断增加数据重复性，减少手工操作时间。同时，有最快速高通量的扫描平台，这个最经济、灵活的系统可以用来进行RNA和DNA多项应用研究。
- 低样量完成实验，最低只需要10~100ng RNA用于表达谱芯片。500ng用于全基因组SNP芯片。实验样本可以是福尔马林或液氮处理的。
- 结合于芯片设计中，实验各个阶段设置了合理的对照，方便了实验监控及使其趋于完美。
- 每个转录本或SNP具有多个探针芯片设计，为准确的qPCR验证提供了基础。

软件及数据分析

- 所有的探针测序的描述和注解都可通过NetAffx™分析中心获得，达成在芯片的设计和下游分析处理的紧密相连，从而快速获取其生物学意义。
- 公共数据库拥有大量Affymetrix芯片实验结果，如NCBI的基因表达库(GEO)，和其它的生物芯片平台相比，存有1000多倍的数据，能够提供更大量的比较数据以满足快速进行分析生物路径和目标靶的需要。
- 大量的商业软件和免费软件为GeneChip系统而构建了相应软件。使得复杂或者个性化数据分析成为可能，同时也大大促进了实验结果的交流。



美国食品和药物管理局(FDA)最近进行生物芯片质量监控研究(MAQC)来确定不同生物芯片平台的结果的重复性和为科学家提供开发临床用芯片标准。研究结果(Nature Biotechnology Sept. 8, 2006)表明为什么昂飞技术被视为生物芯片领域的黄金标准以及为什么科学家发表了超过8500多篇论文中用了昂飞技术。昂飞的技术使您能获得最灵敏、最高效和最具重复性的数据,并且以难以想象的创造性方式去研究基因组。

技术介绍

• 本期技术介绍为表达谱芯片技术及应用。图1描述了研究中应用生物芯片的实验的基本流程。

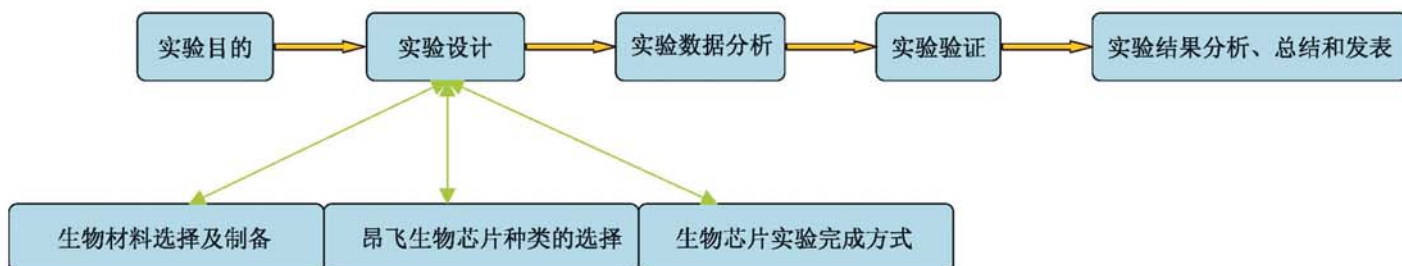


图1. 生物芯片实验设计流程图

昂飞表达谱芯片按芯片设计策略可分为两类: 3'端表达谱芯片(图2以U133 PLUS 2.0为例); 全转录本(WT)表达谱芯片(Exon 1.0 ST芯片和Gene 1.0 ST芯片为例)

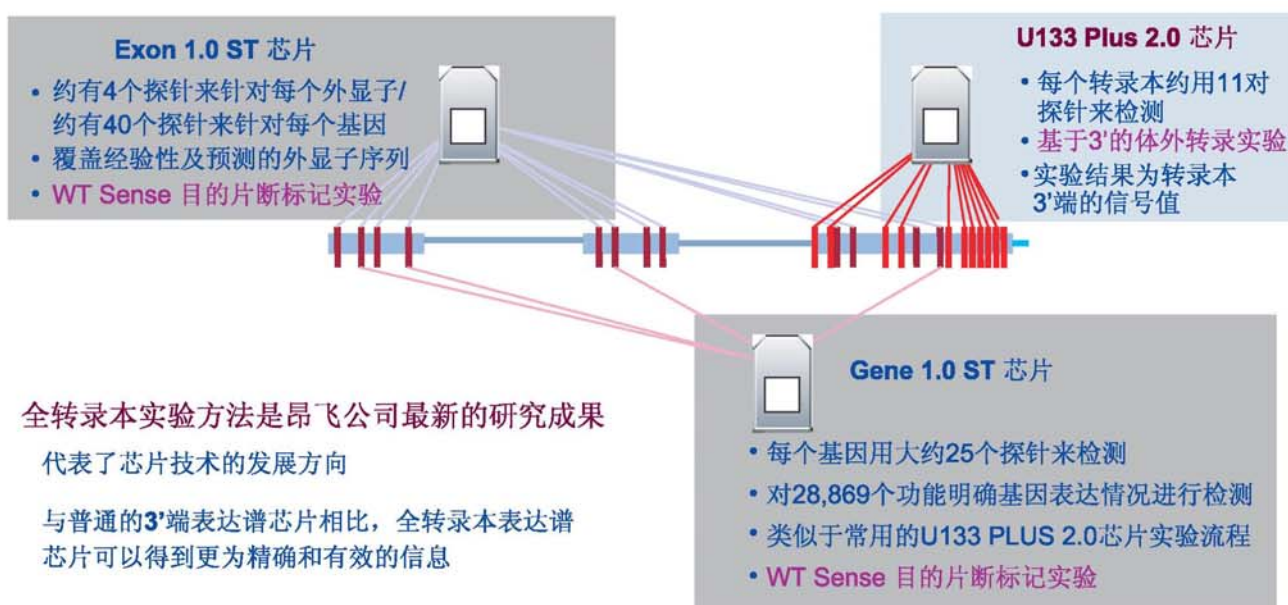


图2. 3'端表达谱芯片和全转录本(WT)表达谱芯片的比较

如图2所示, 3'端表达谱芯片的设计策略指的是通过RNA末端反转录机制, 将检测目标集中在RNA的3'端区域。传统表达谱芯片都采用3'端体外转录的设计策略, 包括人、大鼠、小鼠、狗、鸡、猪、牛、猴、拟南芥、水稻、小麦、大麦、大豆、杨树、柑桔、苜蓿、西红柿、葡萄、棉花、果蝇、线虫、斑马鱼、爪蟾、大肠杆菌、金黄色葡萄球菌、绿脓杆菌等, 以及按照实际需要定制的表达谱芯片。以人类基因组表达谱芯片U133 PLUS 2.0为例具体说明3'端芯片设计: 利用多种转录本序列, 其中包括表达序列标签(EST), 去产生共识转录本序列, 方向共识序列, 找出对于每个多聚腺苷酸合成位点, 然后在多聚腺苷酸合成位点5'端区域, 由11对探针组成的探针组完成正配和错配的选择。详细信息, 请参考技术文献Technical Note, ArrayDesign for the Human Genome U133 Set. www.affymetrix.com/support/. 本刊物中提到的所有技术文献和实验手册均可在此网址查阅到。

依赖专注于3'端目的序列的操作方法有其局限性。因为这种方法是利用已知的选择性的多聚腺苷酸加尾(Polyadenylation)和选择性的末端外显子来设定多个偏向3'端序列的探针组。一种最新的杂交样本制备方法, 全转录本(Whole Transcript, WT)实验, 能针对整条转录本产生目的片断, 而不是仅仅局限于3'末端。全转录本(WT)表达谱芯片设计策略是通过随机引物对全长RNA进行整体反转录, 将检测目标放宽到整个转录本的所有区域(如图2所示)。详细信息, 请参考技术文献: Technical Note, Human Exon 1.0 ST Array and WT Sense Tar-

get Labeling Assay for Genome-Wide, Exon-Level Expression Analysis.

全转录本 (WT) 表达谱芯片目前已有了基因组外显子芯片 (Human Exon 1.0 ST芯片、Mouse Exon 1.0 ST芯片、Rat Exon 1.0 ST芯片) 以及最新的人类基因芯片 (Human Gene 1.0 ST)。前三种基因组外显子, 以人类基因组外显子 Human Exon 1.0 ST 芯片为例, 其设计是基于完整的人类基因组信息, 包含了已知外显子, 以及更多的基因推测性部分。这类外显子芯片可以得到外显子水平基因表达分析结果。Gene 1.0 ST 芯片则是一种专注于基因-水平表达分析的芯片, 其含有从 RefSeq、Ensembl 以及 GenBank 假定完整转录本编码序列库中挑选出来 28,869 个功能清晰的全长基因。在人类 Gene 1.0 ST 芯片设计时尽可能地使用 Human Exon 1.0 ST 芯片相同探针的子集, 来更好地在基因水平上将目光集中在功能明确、研究充分的基因上。针对全长转录本设计的探针, 可以为基因位点提供一个更为精确代表了整体转录活性信号值。详细信息, 请参考技术文献: Technical Note, Array Design for the Human Exon 1.0 ST Array, 和 GeneChip® Human Gene 1.0 ST Array Design.

新技术与新启示

mRNA 的生成, 选择性剪接和全转录本表达谱芯片众所周知, 信使 RNA (mRNA) 是一种生命周期很短的分子, 它可以将细胞核以 DNA 形式储存的遗传信息转移到细胞质中, 然后将这个信息翻译成蛋白质。mRNA 生成涉及两个步骤: 转录和 RNA 处理。

转录的过程就是以 DNA 分子的一条单链作为模板, 通过 RNA 聚合酶来指导互补 RNA 链合成的过程。加强对可能转录方式的理解, 例如 3' 末端不确定的转录本、截短转录本、转位、没有多聚-腺苷酸 (PolyA tail) 的转录本 (图3), 对于获得涉及到整个基因组位点转录活动的完整情况至关重要。

在真核生物中, 从 DNA 转录这一过程中而得到的起始 RNA 产物 (最初转录本) 含有内含子和外显子。最初转录本要经历一个 RNA 处理过程来形成具有生理功能的活性的 RNA 种类--功能性 RNA。RNA 处理包括主要的 3 个方面是 5' 端加帽、切割和 3' 端加多聚腺苷酸尾, 以及 RNA 选择性剪接 (图3)。选择性剪接是高等真核生物蛋白质多样性的主要来源, 经常是以发育阶段特异性或者组织特异性方式进行调节。目前估计大约有 40%-60% 的人类基因具有不同的异构体形式。同一基因来源的剪接异构体可以产生在特征上有明显区别、功能上差异显著 (甚至是相互拮抗的) 的蛋白质。此外, 涉及到人类疾病的大量遗传突变被定位在可调节剪接的剪接信号或者序列上。因此, 对剪接方式改变的理解在深入探讨生物调节和疾病机制方面尤为重要。

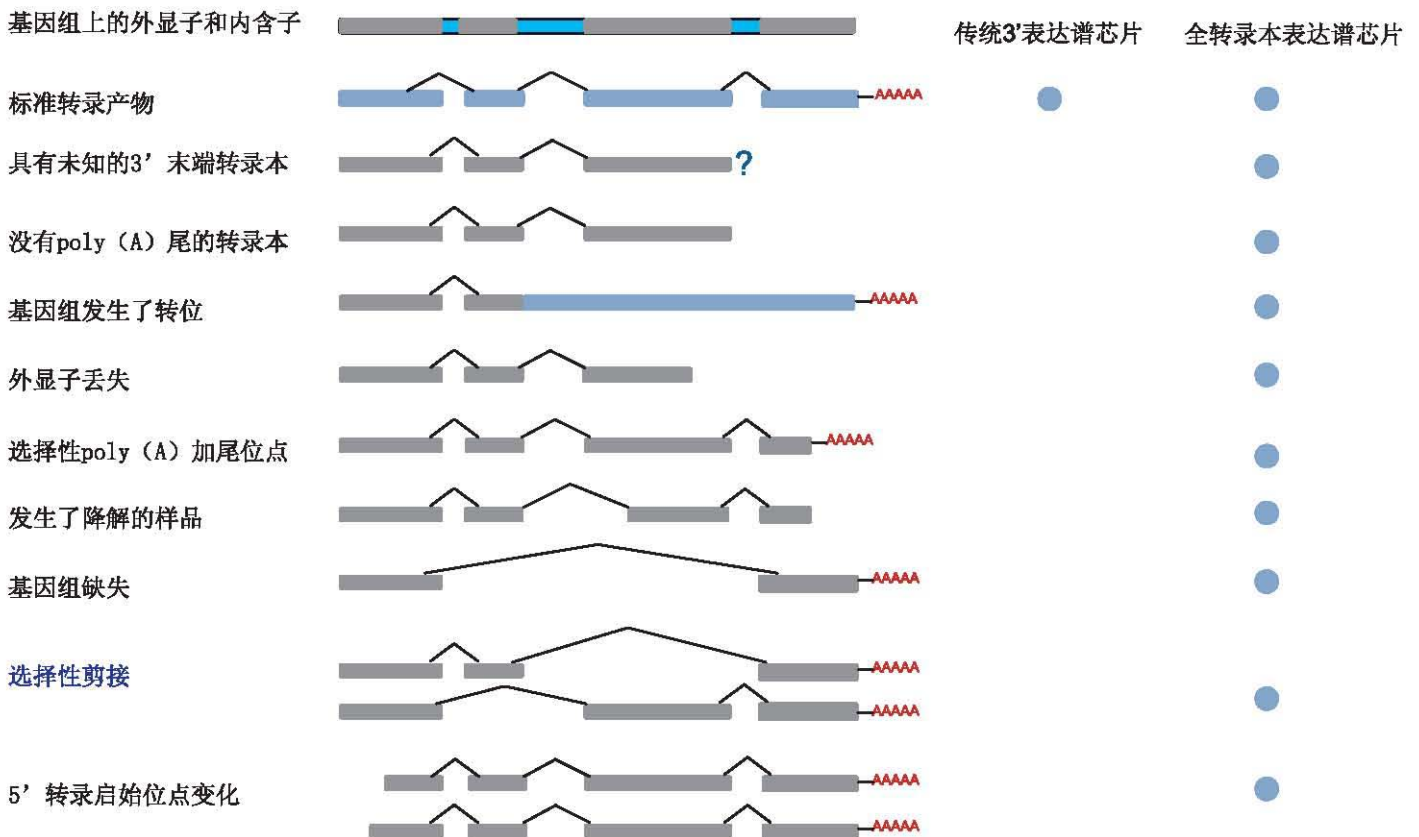


图3. 全转录本表达谱芯片所具有的优势

从图3中我们可以看出，传统的3'端表达谱芯片仅仅能对标准转录本进行有效的分析。全转录本表达谱芯片（Exon 1.0 ST 和 Gene 1.0 ST）则是强有力工具去检测转录方式的可能性。而其一的Exon 1.0 ST 芯片更可以获得与选择性剪接相关信息，进而可以进行（1）选择性剪接调控的分子机理分析。（2）药靶基因的新型选择性剪接异构体的发现，以及获得其组织特异性表达图谱。（3）增进对遗传性变异所导致的下游效应的理解，从而可以将选择性剪接方式同生物表型变化联系起来。请参阅应用案例之四。

实验指南

本期实验指南为表达谱芯片实验和数据分析流程方面，与传统性3'端表达谱芯片相比，如图4所示，全转录本表达谱芯片实验的特点为：1) RNA的起始用量大幅度的下降（推荐采用100ng或1μg Total RNA）。2) 如用EXON 1.0芯片，流程要求有rRNA消滅步骤，用以降低背景信号以及提高特异性（用Gene 1.0芯片时省略rRNA消滅步骤）。3) 与3'端表达谱芯片步骤类似—cDNA 1链、2链合成及体外转录（IVT），区别在于全转录本表达谱芯片体外转录合成cRNA时并不对其进行标记。4) 其后部分步骤是全转录本芯片所特有的：通过随机引物来合成掺入了dUTP的正义链DNA，水解cRNA模板后对正义链DNA进行片断化和标记。5) 随后完成同样杂交、洗涤和扫描等一系列流程以获得到结果。实验操作作用的详细信息，务请参考实验手册：GeneChip® Expression Analysis Technical Manual 和 GeneChip® Whole Transcript (WT) Sense Target Labeling Assay Manual。

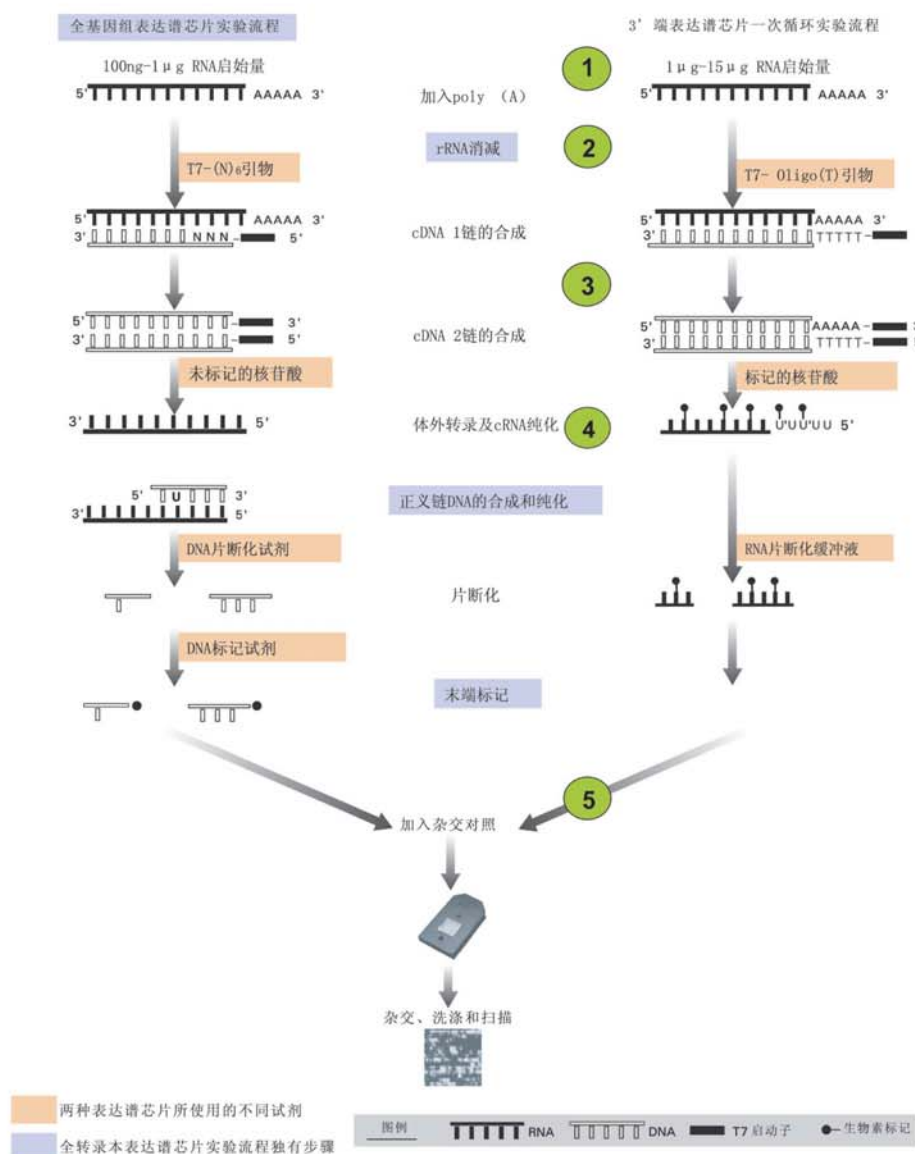


图4. 全转录本表达谱芯片和传统3'端表达谱芯片的实验流程比较

基因水平的表达数据分析方面，全转录本表达谱芯片（Gene 1.0 和 Exon 1.0）的分析流程同传统3'端表达谱芯片类似（图5），而同一张Exon 1.0芯片还能检测外显子水平上表达数据变化。外显子水平表达谱分析请参考：Technical Note, Identifying and Validating Alternative Splicing Events.

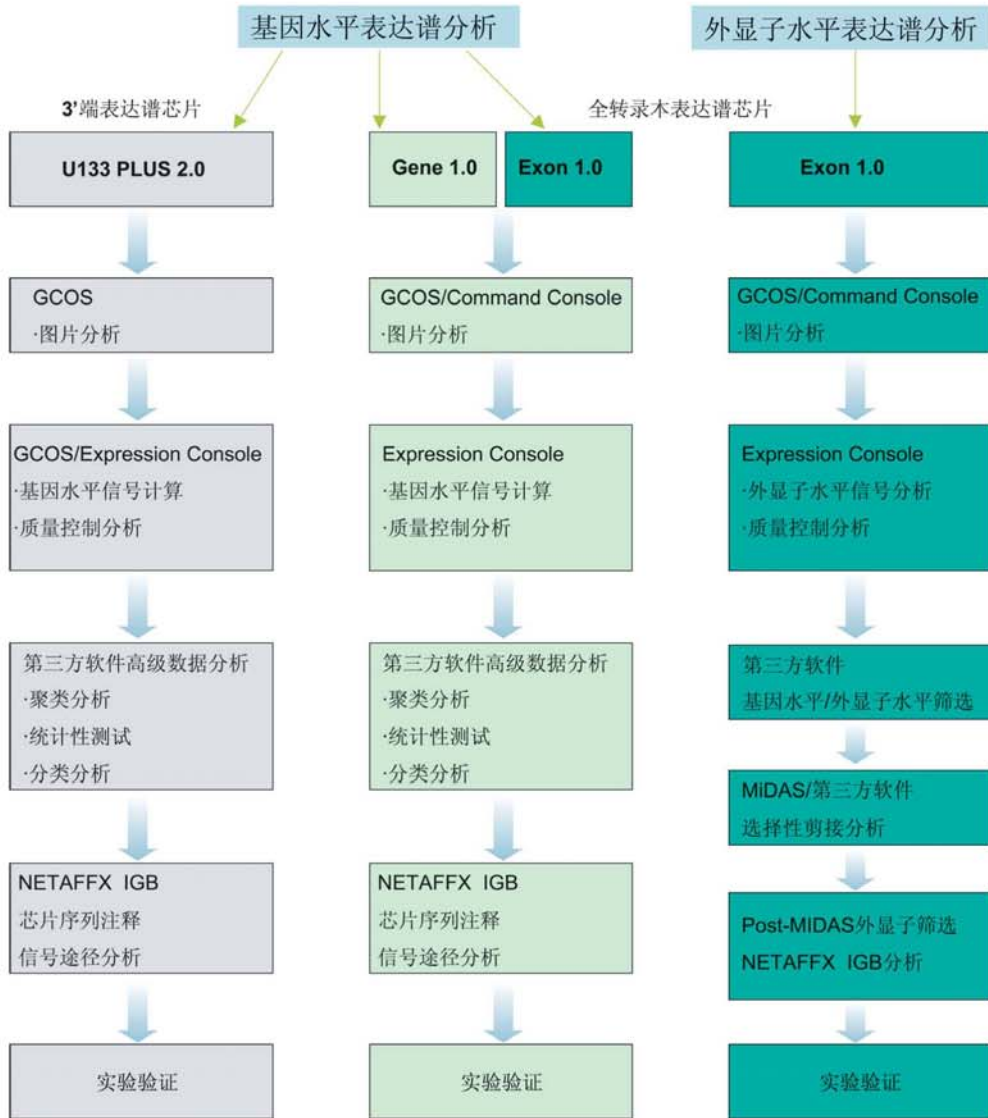


图5. 全转录本表达谱芯片和传统3'端表达谱芯片的分析流程比较

以上简述图5中提到的免费的昂飞数据存储和分析工具。GCOS—GeneChip[®]操作系统采用数据库结构存储和管理样本、样本属性、相应芯片和数据结果，微软WindowsXP环境下收集和初步处理数据。Command Console[™]软件使用最先进的软件开发技术，包括基于网络的用户界面、文件索引以及通过网络服务(Web Services)的整合。Expression Console[™]软件为使用者提供了常用的探针组计算算法。Expression Console[™]和Command Console[™]软件的整合使Command Console[™]界面收集的样本的属性信息在Expression Console[™]显示出来，并可被用来引导使用者进行分析和质量控制。MiDAS是Affymetrix Power Tools (APT)中用于外显子水平和选择性剪接分析的工具。IGB（集成基因组浏览器）是基因组环境、生物芯片设计和生物芯片数据的可视化软件。NetAffx[™]分析中心是昂飞的网络工具，用于从生物芯片测得的数据中快速获取其生物学意义。有关各软件的详细信息请查阅 www.affymetrix.com/support/。

第三方软件高级数据分析是GeneChip-Compatible[™]软件。compatible[™]Application program罗列了超过25家为昂飞平台构建软件的商业软件供货商。（www.affymetrix.com/products/software/compatible/snp.affx）提供软件包名录。

应用案例

在癌症分子分型研究方面，早在1999年，Golub在Science杂志的文章中表明，他的研究小组使用昂飞生物芯片，在鉴定癌症种类以及肿瘤分型方面取得了重大进展。Golub等在没有获得白血病病例背景信息的条件下，通过昂飞表达谱生物芯片，针对急性巨核细胞白血病（acute myeloid leukemia）和急性淋巴细胞性白血病（acute lymphoblastic leukemia）的可能性进行了预测性分析。研究表明，单纯依赖基因表达芯片生物芯片检测来对癌症进行分型是具有可行性的。这类研究提出了一条不依靠早期生物学表型的癌症类型预测和鉴定的策略。

T. R. Golub., et al., Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. Science 15 October 1999: Vol. 286. no. 5439, pp. 531 – 537.

Kimchi等学者使用人类基因组U133A芯片发现了与食道癌发病机理相关的一个新通路。他们对正常食道上皮、巴瑞特组织(Barret)变形和食道腺癌所采样本进行评测，发现有214个基因标记可用于区分三种样品型。他们的数据显示从巴瑞特氏食道转化为腺癌与表皮分化复合物相关的基因抑制有联系。对于在巴瑞特氏食道和腺癌中都有表达的基因进行相关分析可进一步发现21个可代表潜在遗传标记和治疗干扰的药物靶标的基因。

Kimchi, E.T., et al. Progression of Barrett's metaplasia to adenocarcinoma is associated with the suppression of the transcriptional programs of epidermal differentiation. Cancer Res 65: 3146-54 (2005).

在植物学基因组研究方面，昂飞芯片也有广泛的使用。Plant cell 杂志曾经报道过Caldo等研究人员利用昂飞公司含有22,000个大麦基因的表达谱芯片对植株基因型、小麦白粉菌(Bgh)及时间之间36个组合的相互作用进行了基因表达谱分析，深入研究了植物抗病原体感染的分子机制。通过比较植物通用感染抗性和特定感染抗性之间的关系，作者认为在病原体相关的通用识别和特定识别之间存在着联系，并且提出证据支持下面这个假说：病原体对植物基本防御的抑制这一状况能被宿主识别和防止，并从这个基础进化出了宿主特异性抗性。

Rico A. Caldo, et al., Interaction-Dependent Gene Expression in Mla-Specified Response to Barley Powdery Mildew. The Plant Cell, Vol. 16, 2514-2528, Sep 2004.

在应用全转录本表达谱芯片的研究中，虽然此类新型芯片推出的时间还不长，但是有很多科学家利用其得到了许多激动人心的成果。荷兰Erasmus医疗中心French等研究者最近使用人类Exon 1.0芯片对两种不同类型的神经胶质瘤进行了剪接异构体表达特征分析，已得到了突破性的结果。他们发现肿瘤中大约有20%的突变引起了外显子跳跃；在不同的个体中会观测到外显子跳跃这一现象的发生；那些相关基因的测序结果表明存在着同外显子跳跃现象相关的突变。这些发现指明了一个新的研究方向：导致疾病相关剪接异构体发生的基因可能会成为癌症治疗的药靶目标。Affymetrix Microarray Bulletin, Vol2, Issue 3, Summer 2006.

本期相关技术要点

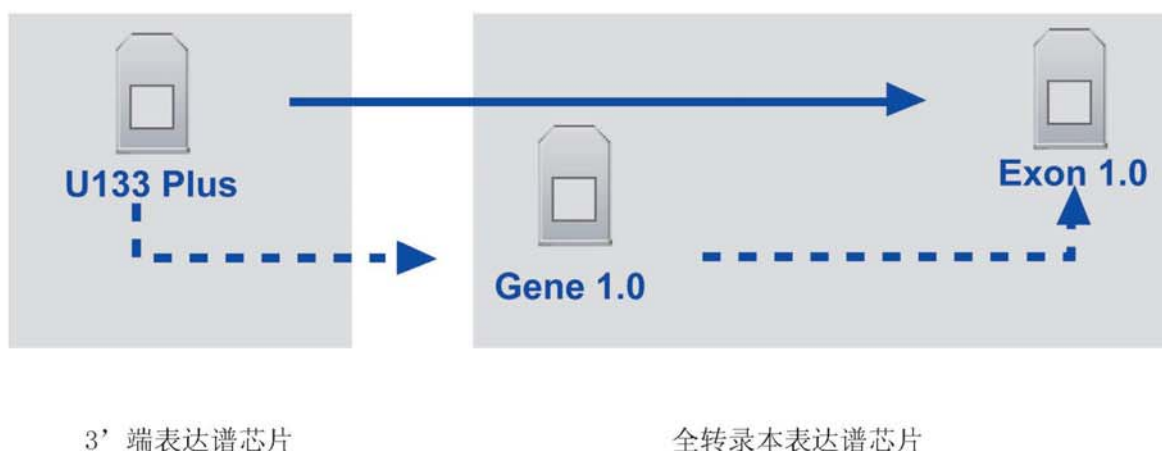


图6. 人类表达谱生物芯片发展过程。

图6中可以看出，以人类表达谱芯片为例，从传统3'表达谱芯片发展到全转录本表达谱芯片是一个不可逆转的技术发展趋势。以下二表对两类芯片的参数和特点进行了一个简要总结比较。

表1. 各种人类表达谱芯片的参数

表达谱芯片种类	规格	Feature size	所检测目标的数量	探针数目/目标序列	检测链的方向
U133 PLUS 2.0	49	11 μ M	~47,000个转录本	11对/转录本	反义链
Gene 1.0 ST	169	5 μ M	28,869条全长基因	平均25个/基因	正义链
Exon 1.0 ST	49	5 μ M	~1,000,000个外显子	平均4个/外显子	正义链

表2. 各种人类表达谱芯片的特点和优势

表达谱芯片种类	特点和优势
U133 PLUS 2.0	<ul style="list-style-type: none"> • 已被广泛使用，并有大量发表的文献，易于比较分析
Gene 1.0	<ul style="list-style-type: none"> • 更敏感和准确的分析结果，可以做到一个基因一个信号值 • 基于最新数据库进行芯片设计 • 分析方法和前一代表达谱芯片相比，没有复杂化 • 已有完整的解决方案 • 低廉的价格
Exon 1.0	<ul style="list-style-type: none"> • 一张芯片可以同时检测基因水平和外显子水平的检测 • 对基因表达可以进行更为精确的检测 • 基于最新数据库进行芯片设计 • 检测方法和前一代表达谱芯片相比，没有复杂化 • 可以得到外显子水平基因表达分析结果

通过技术的进步，全转录本芯片探针合成的密度比上一代3'端表达谱芯片有了极大的提高（表1）。所挑选的探针覆盖整个转录本的范围，从而能提供强有力的基因水平和外显子水平上的分析结果（表2）。昂飞公司上海代表处将极力促进3'表达谱芯片顺利过渡到全转录本表达谱芯片平台上，努力帮助科学家尽快将全转录本表达谱芯片技术成为研究的有力工具。

美国昂飞公司上海代表处

地址：中国上海陆家嘴环路1233号汇亚大厦1508室（邮编：200120）
 电话：86-21-68881980 传真：86-21-68881981
 saleschina@affymetrix.com supportchina@affymetrix.com
 www.affymetrix.com

Let Professionals Serve Professionals

香港：852-28966283
 广州：020-85524840
 武汉：027-87166462
 济南：0531-86560825
 西安：029-82501170

北京：010-51665161
 天津：022-88293136
 昆明：0871-5349992
 南京：025-83248692
 杭州：0571-87229824

上海：021-64951899
 成都：028-85431195
 沈阳：024-23341315
 长沙：0731-4476562
 重庆：023-68614842



Gene Company Limited
基因有限公司
 A Gene Group Company