International Rare Diseases Research Consortium (IRDiRC)

Genomic Medicine: Current Status and Future

Zhiyu Peng, BGI, China

7 November 2014

Starting From Human Genome Project

说起基因组学，不得不提人类基因组计划

20世纪人类三大科学计划

1945 USA
China1964

曼哈顿原子弹

1969 USA
China 2007

阿波罗登月

2000-2008
USA … China

人类基因组计划

人类基因组计划：
六国参与，耗时超10年，花费30亿美金

促进了诸多学科的发展
带动了8万亿的产业机会

1999年9月9日，
华大基因为了完成人类基因组计划（中国部分）而成立

全球第一的基因组学（研究）中心

排名	研究机构	高通量测序仪数量	国家
1	**BGI**	**166**	**中国**
2	Broad Institute	114	美国
3	The Genome Center at Washington University	62	美国
4	Wellcome Trust Sanger Institute	35	英国
5	Canada's Michael Smith Genome Sciences Centre	27	加拿大
6	DOE Joint Genome Institute	21	美国
7	Yale Center for Genome Analysis	20	美国
8	Human Genome Sequencing Centre, Baylor College of Medicine	18	美国
9	Ontario Institute for Cancer Research	17	加拿大
10	Beijing Institute of Genomics	15	中国

华大基因，拥有世界最大基因组测绘设备和能力的机构

新一代测序平台
服务产品涵盖动植物、微生物、海洋等多个领域
Illumina HiSeqTM 2000
Applied Biosystems SOLiDTM 4 System
Roche GS FLX Platform
Ion Torrent

3730 测序平台
采用最新的光路系统设计，结合更优化的试剂组合和全自动化操作流程，在提高产率、提高数据质量的同时最大限度的降低试剂的消耗。
ABI3730xl

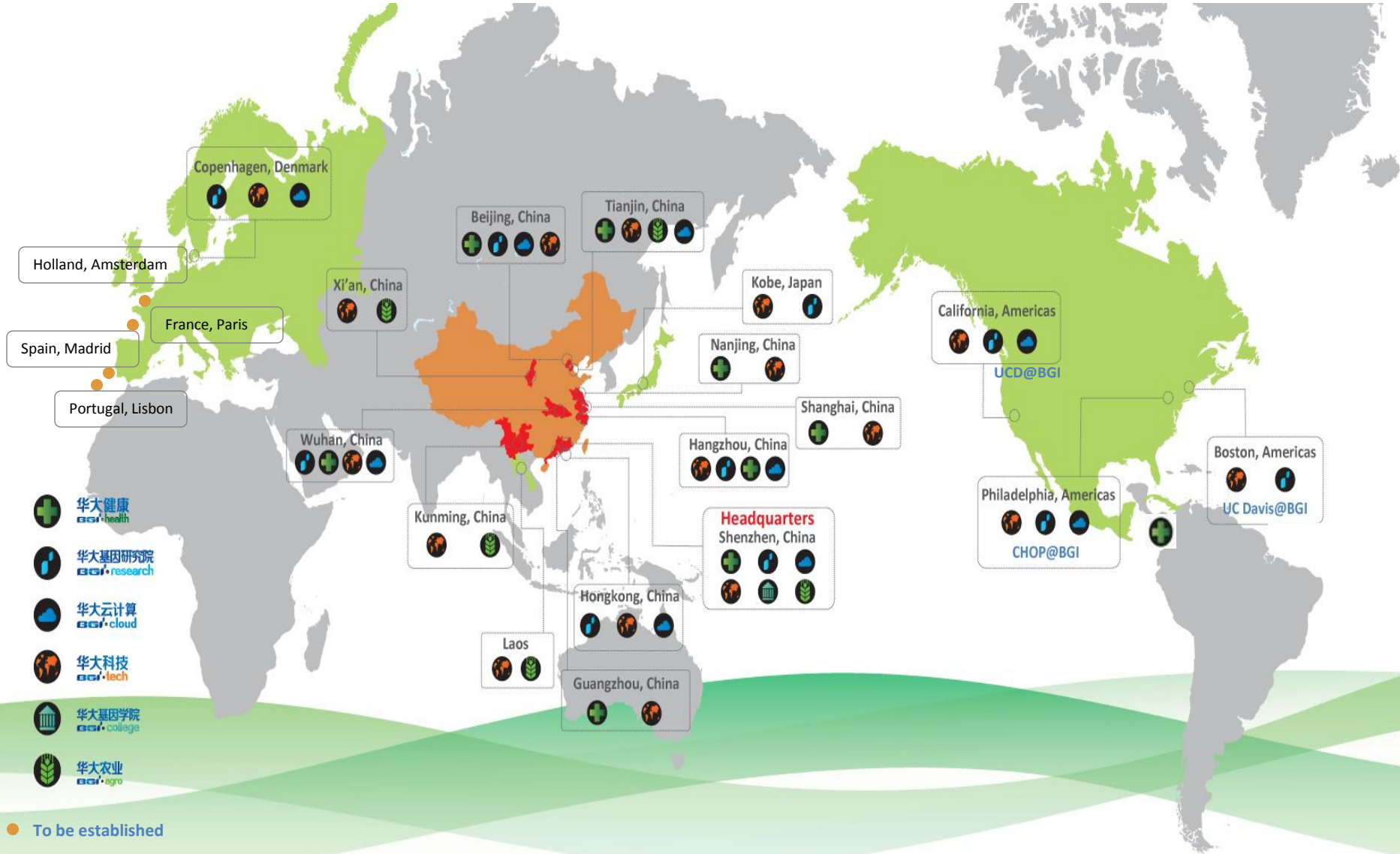
基因分型平台
在分析人类疾病相关位点与结构变异以及动植物农艺性状相关位点上有广泛应用。
Illumina iScan
Sequenom MassARRAY iPLEX Gold

全基因组酶切图谱平台
快速定位序列位置，已广泛应用于比较基因组学分析发现突变基因、辅助高通量测序的数据组装以及菌株分型。
OpGen Argus Optical Mapping System

蛋白质谱平台
实现对动植物、微生物等进行大规模的蛋白质组学研究。应用领域涉及营养学、药学、临床研究、法医毒理学、食品安全分析、环境监测等等。

信息技术平台
提供服务包括云存储、云计算、数据交付、生物信息培训和咨询等。

华大全球布局


Copenhagen, Denmark
Holland, Amsterdam
France, Paris
Spain, Madrid
Portugal, Lisbon
Beijing, China
Tianjin, China
Xi'an, China
Kobe, Japan
Nanjing, China
Shanghai, China
Hangzhou, China
Wuhan, China
Kunming, China
Headquarters
Shenzhen, China
Hongkong, China
Laos
Guangzhou, China
California, Americas
UCD@BGI
Boston, Americas
UC Davis@BGI
Philadelphia, Americas
CHOP@BGI
华大健康 BGI·health
华大基因研究院 BGI·research
华大云计算 BGI·cloud
华大科技 BGI·tech
华大基因学院 BGI·college
华大农业 BGI·agro
To be established

科研成就

发表200多篇SCI学术论文，其中58篇发表于Nature系列与Science等国际顶级学术期刊

华大医学临床检验中心（从科技到生产力的转化）

华大临检中心

深圳华大临床检验中心
医疗机构执业许可证

PCR实验室技术验收
合格证书

基因检测实验室
ISO/IEC 17025:2005认证

基因组技术平台
ISO 9001认证

临床分子诊断实验室

出生缺陷筛查实验室

未知病原体检测重点实验室

HLA配型实验室

什么是基因？

遗传

- 种瓜得瓜
- 种豆得豆

变异

- 一窝生九仔，
- 各个不相同

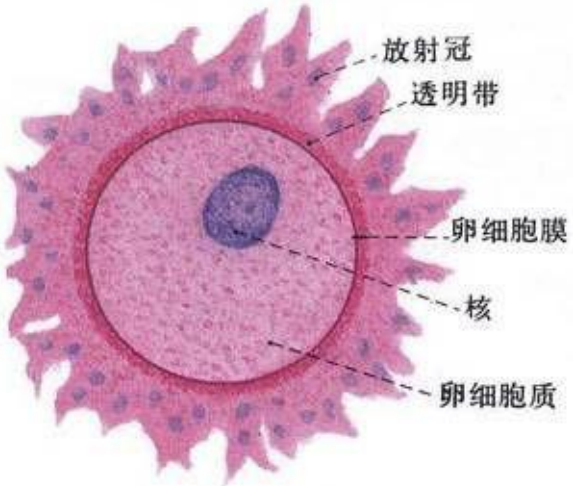
人不可貌相，基因来决定

拇指是否能弯曲

单双眼皮

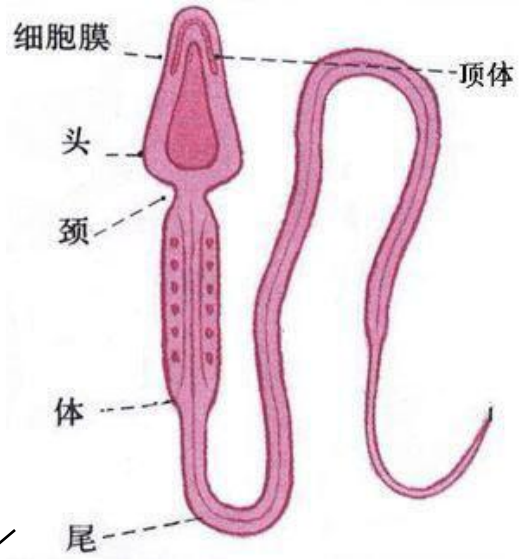
是否卷舌

基因与生命

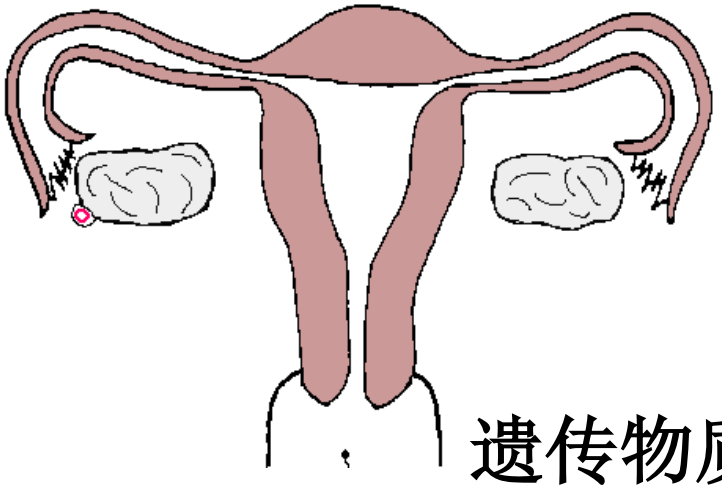


卵细胞

精卵结合
合二为一

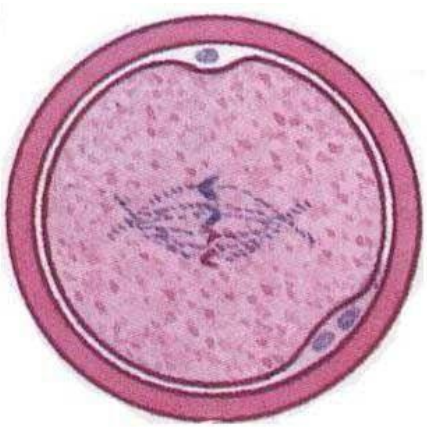


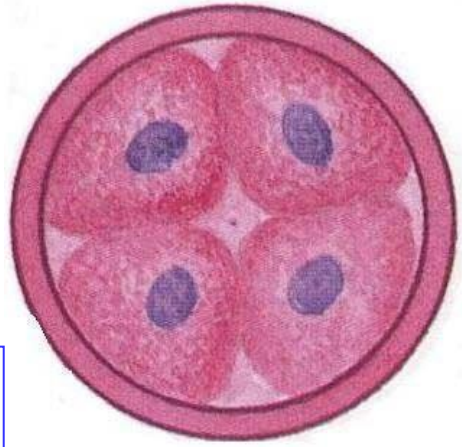
精细胞



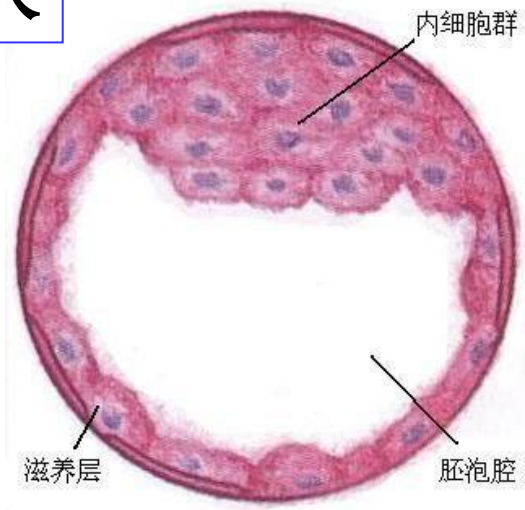
遗传物质DNA

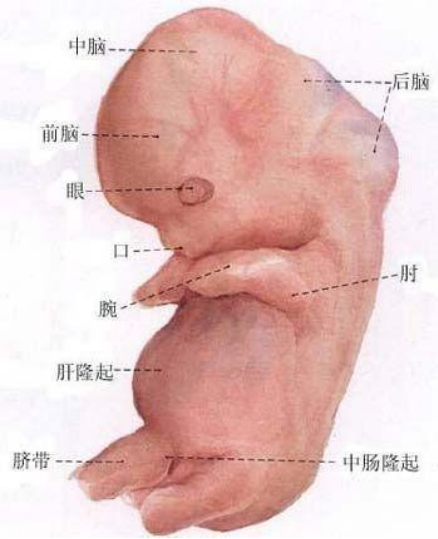
基因与生命





细胞分化：
一分为二，
直到变成人





什么是基因（gene）？

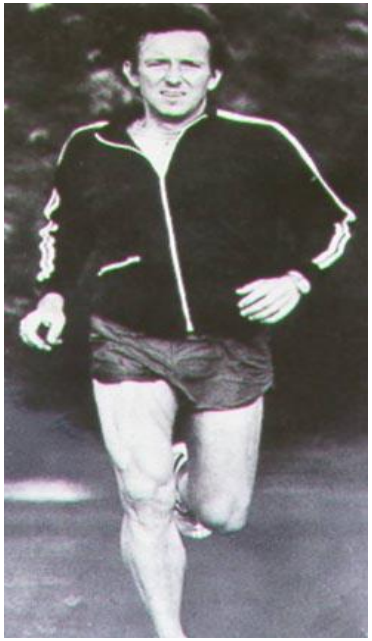
- 什么是生命？（合二而一，一分为二）
- **什么是基因，什么是基因组？**

（一变二的基本因素，组：所有基因）

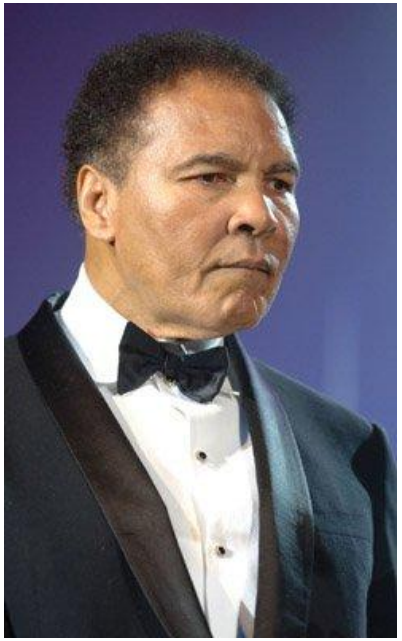
- 基因与生命（父母给后代的只有基因）

- 基因与健康（基因在50%的程度上决定健康）

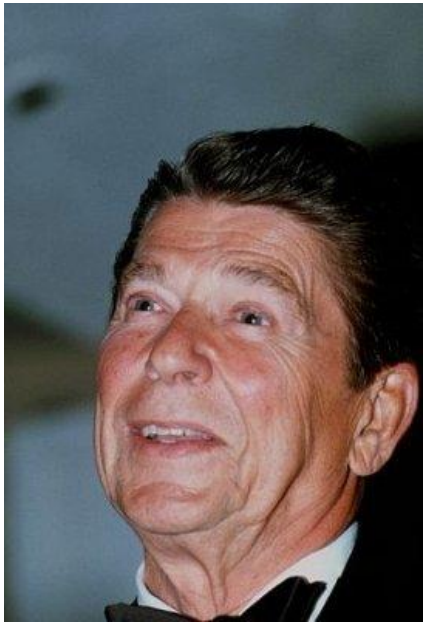
疾病面前，人人平等？



著名马拉松运动员
吉姆. 菲克斯：
心源性猝死



拳王阿里：
帕金森症



美国总统里根：
老年痴呆

英国首相丘吉尔
肥胖
却活到92岁

患乳腺癌的明星们

林妹妹：陈晓旭

南希·里根

奥黛丽 ·赫本

结果显示，ＢＲＣＡ１、ＢＲＣＡ２或ＴＰ５３基因变异是导致家族性乳腺癌和早发性乳腺癌的原因；有这３种基因变异者的终生乳腺癌发病危险为４３％～８５％。

患宫颈癌的明星们

宫颈癌是唯一可以早发现、早预防、早根治的癌症。

——国际抗癌联盟

然而

李媛媛

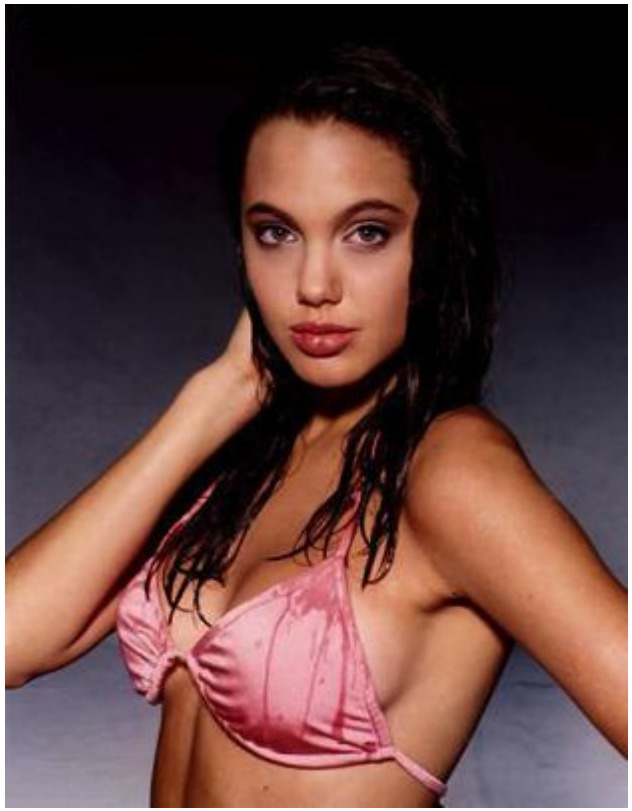
梅艳芳

都跟基因有关系

朱莉切除乳腺自保 “乳腺癌基因”是怎么回事

导语
PREFACE

好莱坞知名女星安吉丽娜·朱莉披露，她由于携带乳腺癌1号基因(BRCA1)，已经接受预防性的双侧乳腺切除手术，以降低罹癌风险。
今天科学姐就告诉你，所谓“乳腺癌基因”这东东真的有！



朱莉母亲曾经与癌症作斗争了近十年，于56岁时去世

她遗传了BRCA1基因，因此
患乳腺癌和卵巢癌的几率分别是87%和50%。

肿瘤与基因组学的发展（从上世纪说起）

如果我们想更多地了解肿瘤，我们从现在起必须关注细胞的基因组。…… 从哪个物种着手努力？如果我们想理解人类肿瘤，那就应从人类开始。……人类肿瘤研究将因对DNA的详细知识而得到巨大推动。

1986年，诺贝尔奖获得者Renato Dulbecco发表短文《肿瘤研究的转折点：人类基因组测序》（Science,231: 1055－1056）

20世纪人类三大科学计划

1945 USA
China1964

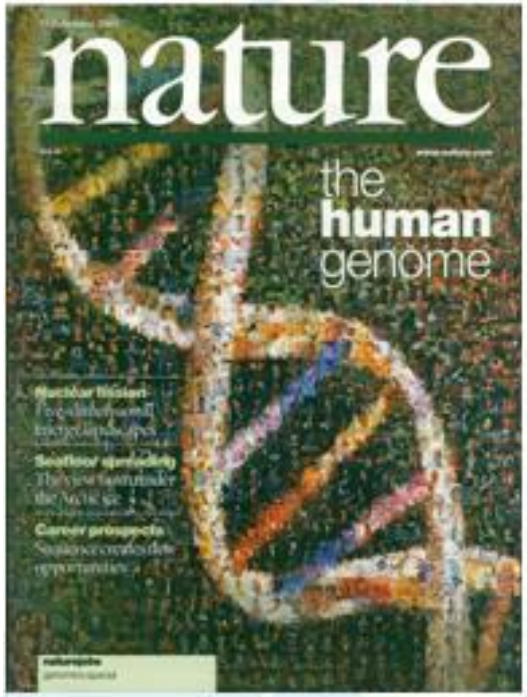
1969 USA
China 2007

2000-2008
USA … China

国际人类基因组计划: 1990—2003，美、英、日、德、法、中

2000年6月26日克林顿代表六国宣布人类基因组框架图完成

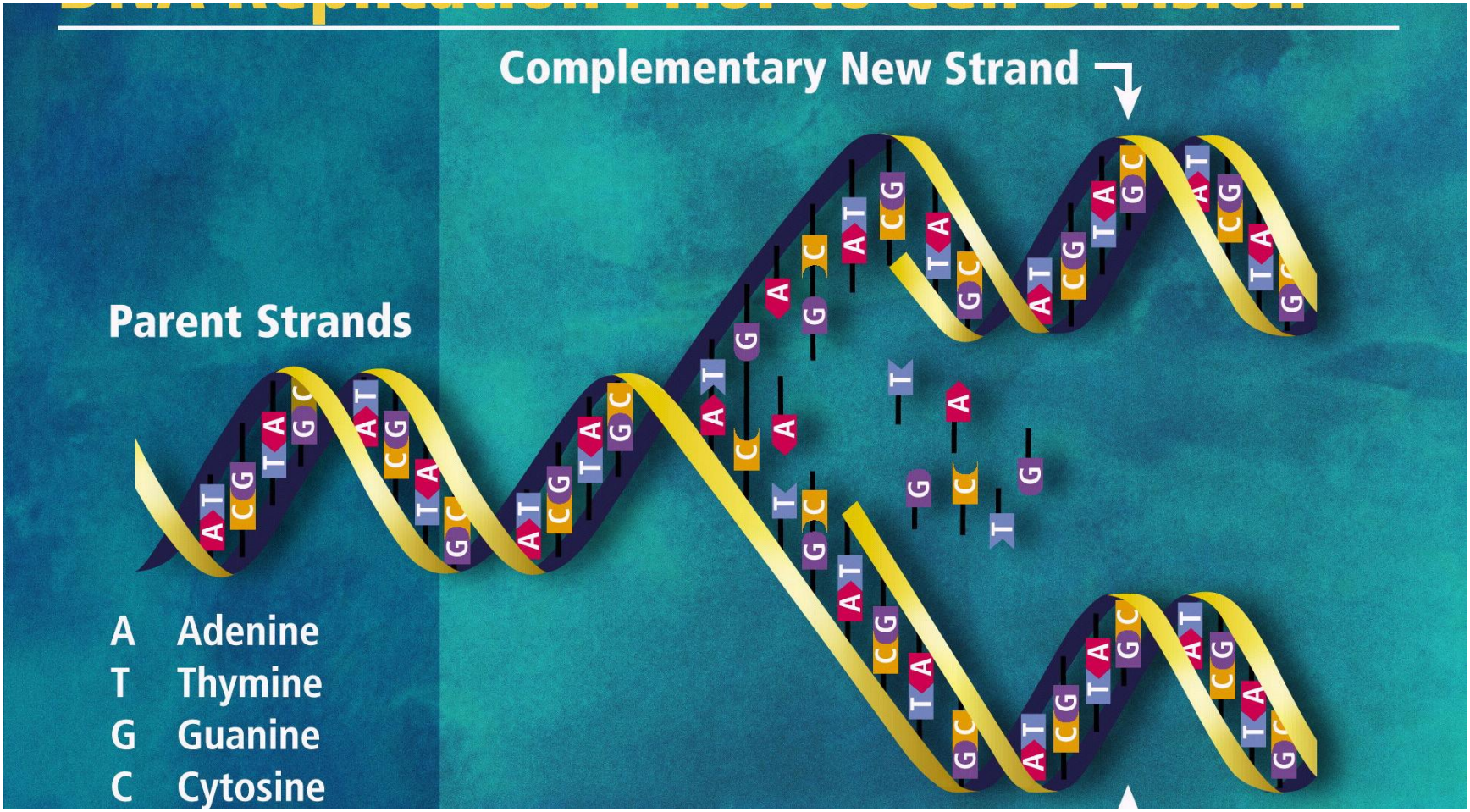
2000年6月28日江泽民表彰中国科学家的贡献



人类基因组计划科研成果发表

人类基因组长什么样？

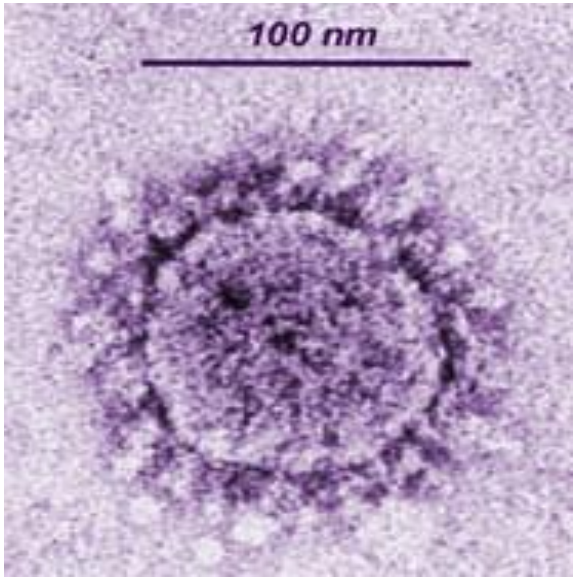
A=T，C=G （01，01）ATCG四个碱基的排列组合，构成基因和基因组, 通过时空变化，掌管着生命大千世界的千变万化及生老病死

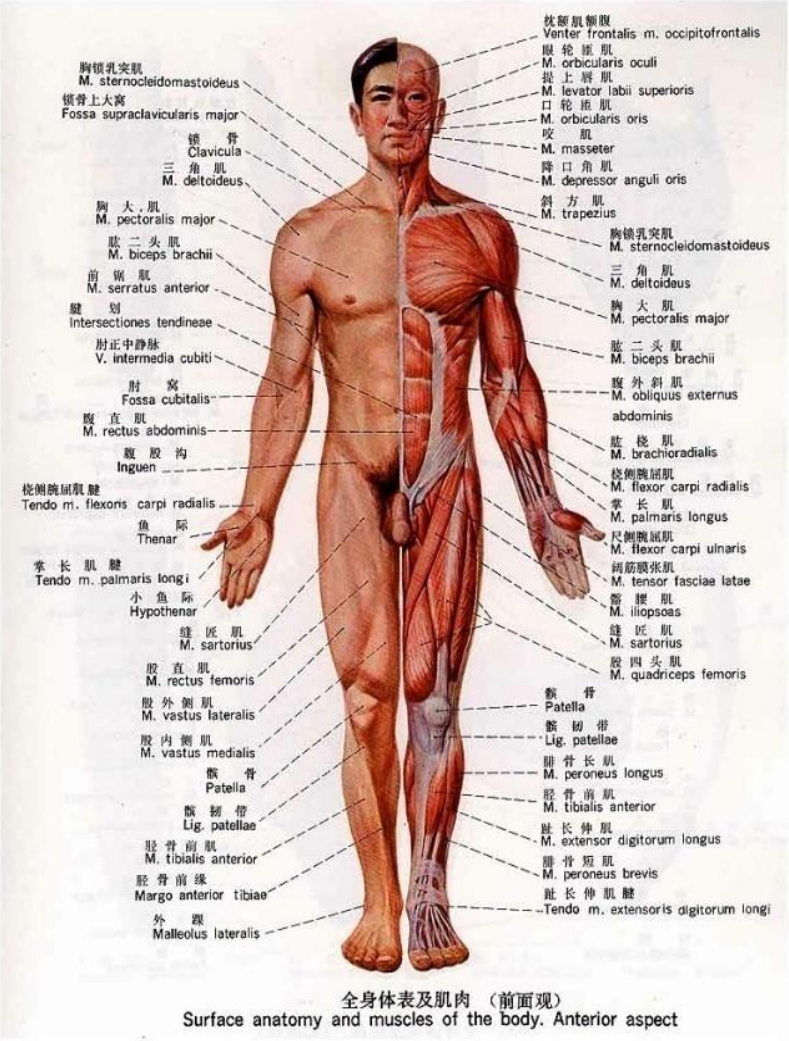


个人基因组的信息量

生命大千世界只是ATCG排序和多少不同

- SARS病毒基因组：32,000碱基
- 人类基因组：3,000,000,000碱基





全身体表及肌肉（前面观）
Surface anatomy and muscles of the body. Anterior aspect

- 有了人类基因组我们就解决了肿瘤的问题吗？
  - 没有，到今天为止仍然没有

为什么仍然没有解决问题

ATCGTAGCATGCGTAAATCGTAGCATGCGTAAATCGTAGCA
TGCGTAAATGTAGCATGCGTAAATCGTAGCATGCGTAAATC
GTAGCATGCGTAAATCGTAGCATGCGTAAATCGTAGCATGC
GTAAATCGTAGCATGCGTAAATCGTAGCATGCGTAAATGTA
GCATGCGTAAATCGTAGCATGCGTAAATCGTAGCATGCGTA
AATCGTAGCATGCGTAAATCGTAGCATGCGTAAATCGTAGC
ATGCGTAAATCGTAGCATGCGTAAATGTAGCATGCGTAAAT
CGTAGCATGCGTAAATCGTAGCATGCGTAAATCGTAGCATG
CGTAAATCGTAGCATGCGTAAATCGTAGCATGCGTAAATCG
TAGCATGCGTAAATGTAGCATGCGTAAATCGTAGCATGCGT
AAATCGTAGCATGCGTAAATCGTAGCATGCGTAA

基因组就像一部天书。

我们不是上帝，看不懂天书。解读天书需要对海量数据作海量计算

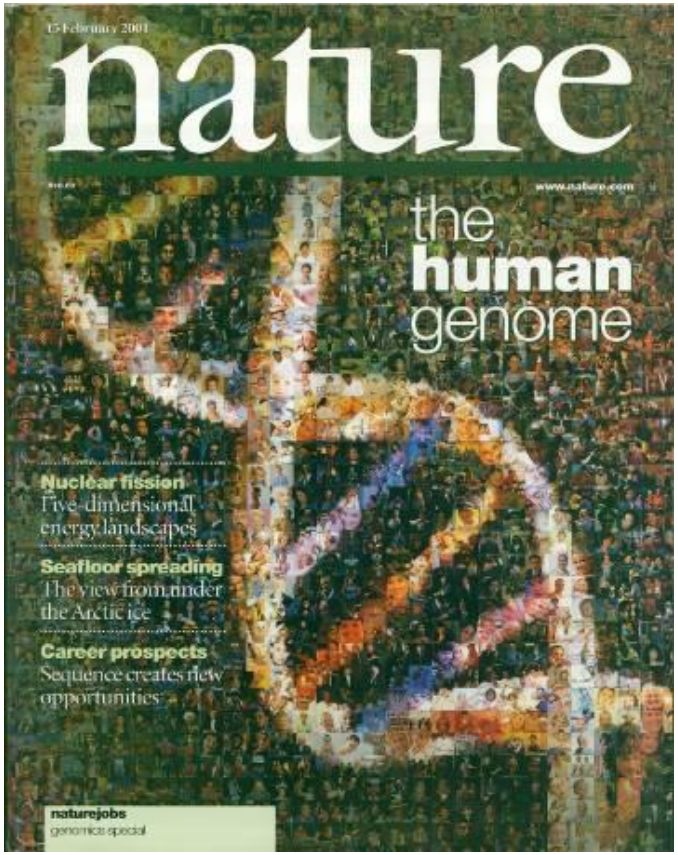
从表型差异找基因差异（如：东西方人差异）

流感、SARS

肿瘤的治疗

易瑞沙（吉非替尼）

东西方人基因组有什么差异？



2000
第一个人基因组(1%)

2008
第一个亚洲人基因组(100%)

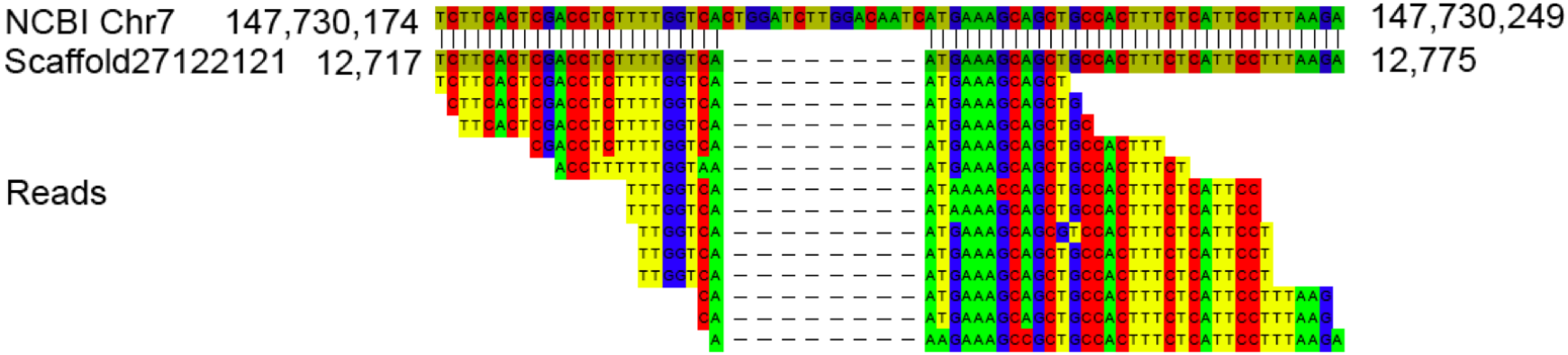
Q:第一作者是谁？

东西方人基因组有什么差异？

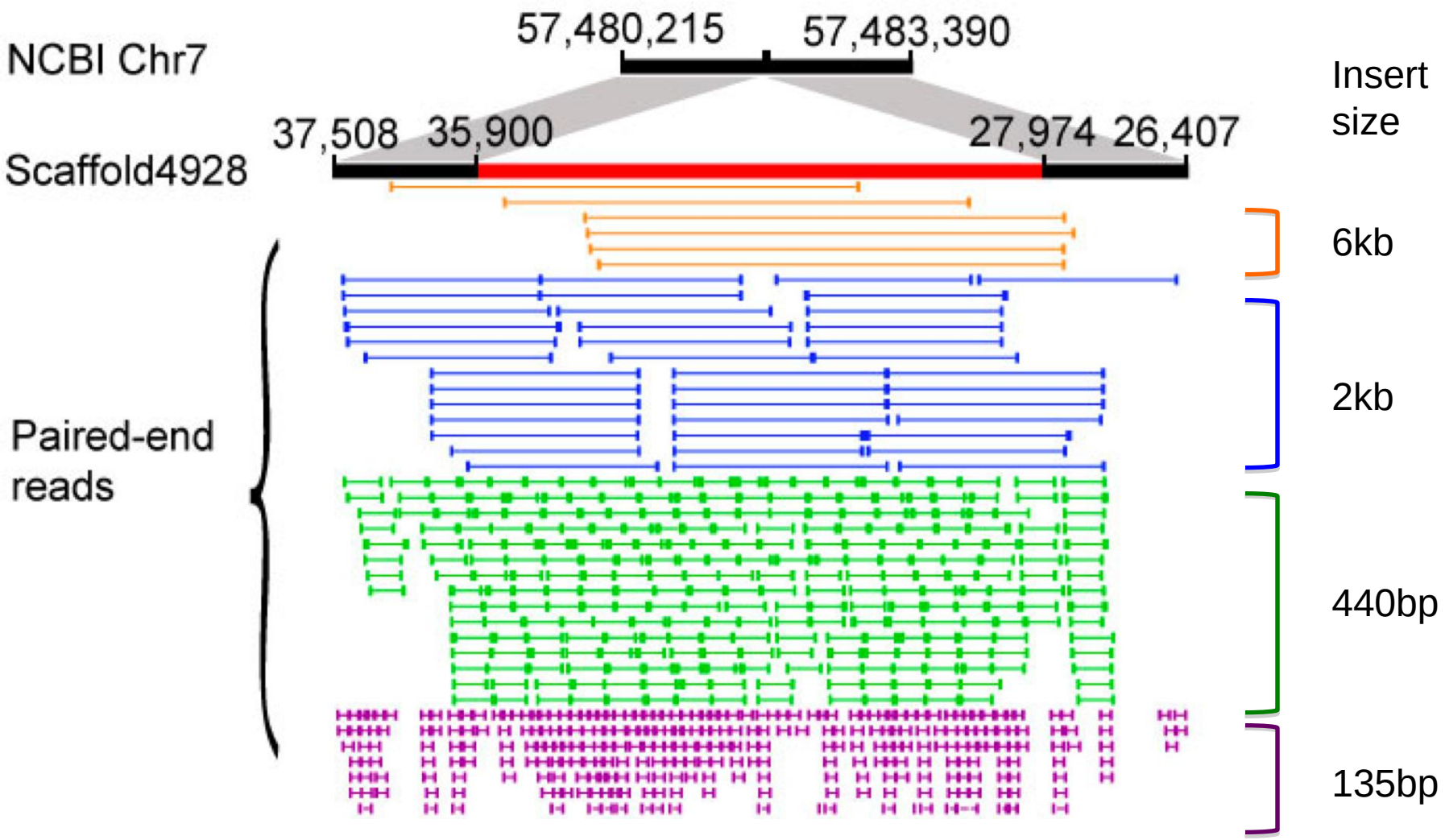
- 300万单碱基差异，其中40万为首次发现
- 2,682 个结构变异;
- 7,000 个基因序列有微小差异

2008.11

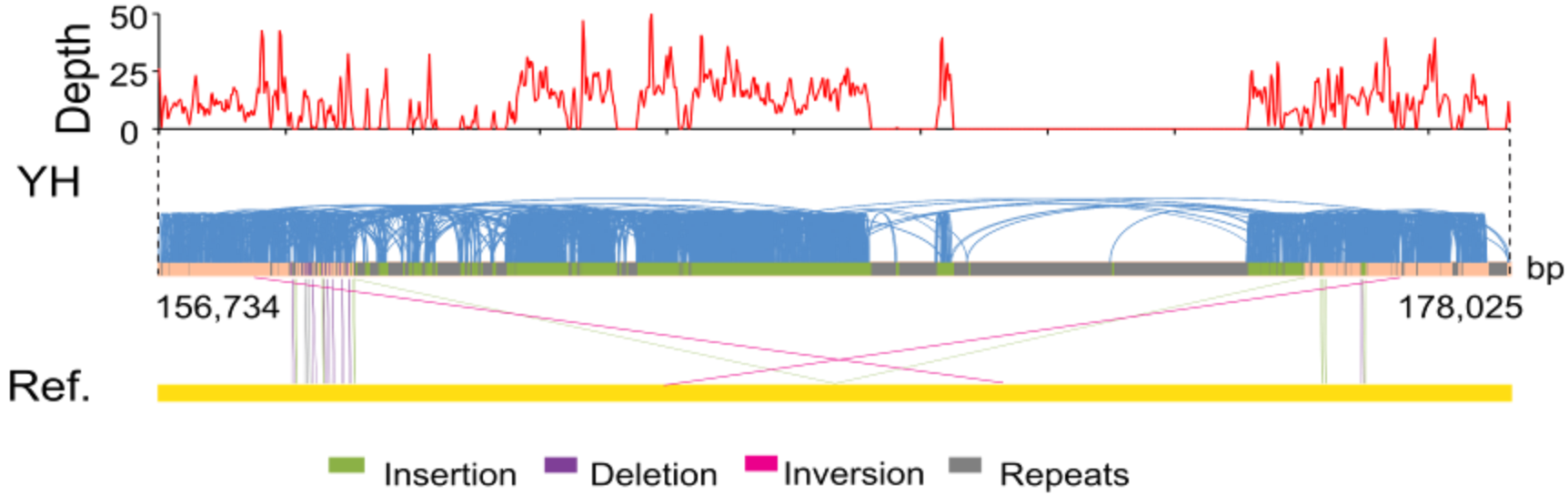
序列删除


NCBI Chr7 147,730,174
147,730,249
Scaffold27122121 12,717
12,775
Reads

8-KB序列插入


NCBI Chr7
57,480,215
57,483,390
Scaffold4928
37,508
35,900
27,974
26,407
Insert size
6kb
2kb
Paired-end reads
440bp
135bp

更复杂案例：插入、删除、反转、重复



A case of complex structural variation in YH genome. The figure illustrate an ~22kbp inversion (pink line, as a cross between assembly and reference) with repetitive sequences (gray block) and several other events, including insertions range from 1bp to ~18kbp (green line and block), and deletions (violet line and block) among a hyper-mutation region (with over 10 insertion and deletion events spacing less than 200bp). Read depth (uppermost line chart) and paired-end reads alignment (medium curve chart) shows the difficulty for this SV event to be detected by previous approaches including RD, PEM and SR.

千人基因组计划：中英启动、美国加入

基于高通量测序构建更精细的人类基因组遗传图谱

- *Find >95% accessible SNPs at allele frequencies above 1%, down towards 0.1% in coding regions*

为基因组研究提供**一个可供参比的基线**

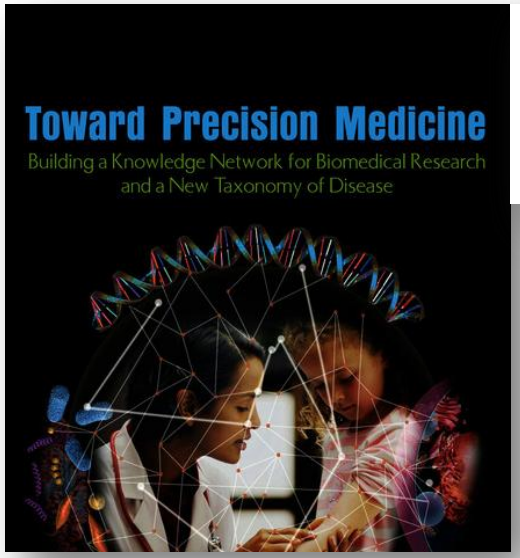
百万人基因组计划

中国：百万中国人基因组计划（2011年）

美国：百万美国人基因组计划（2011年）

英国：10万英国人基因组计划（2012年）

美国：百万美国人基因组计划（2011年）



Toward Precision Medicine:

Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease (2011)

2011年**美国**基因组学与生物医学界的智库发表了《迈向精准医学：建立生物医学与疾病新分类学的知识网络》报告，宣示运用生物医学和临床医学研究的交汇，来编织新的知识网络，并建议启动**“百万美国人基因组计划”**。

英国：10万英国人基因组计划（2012年）

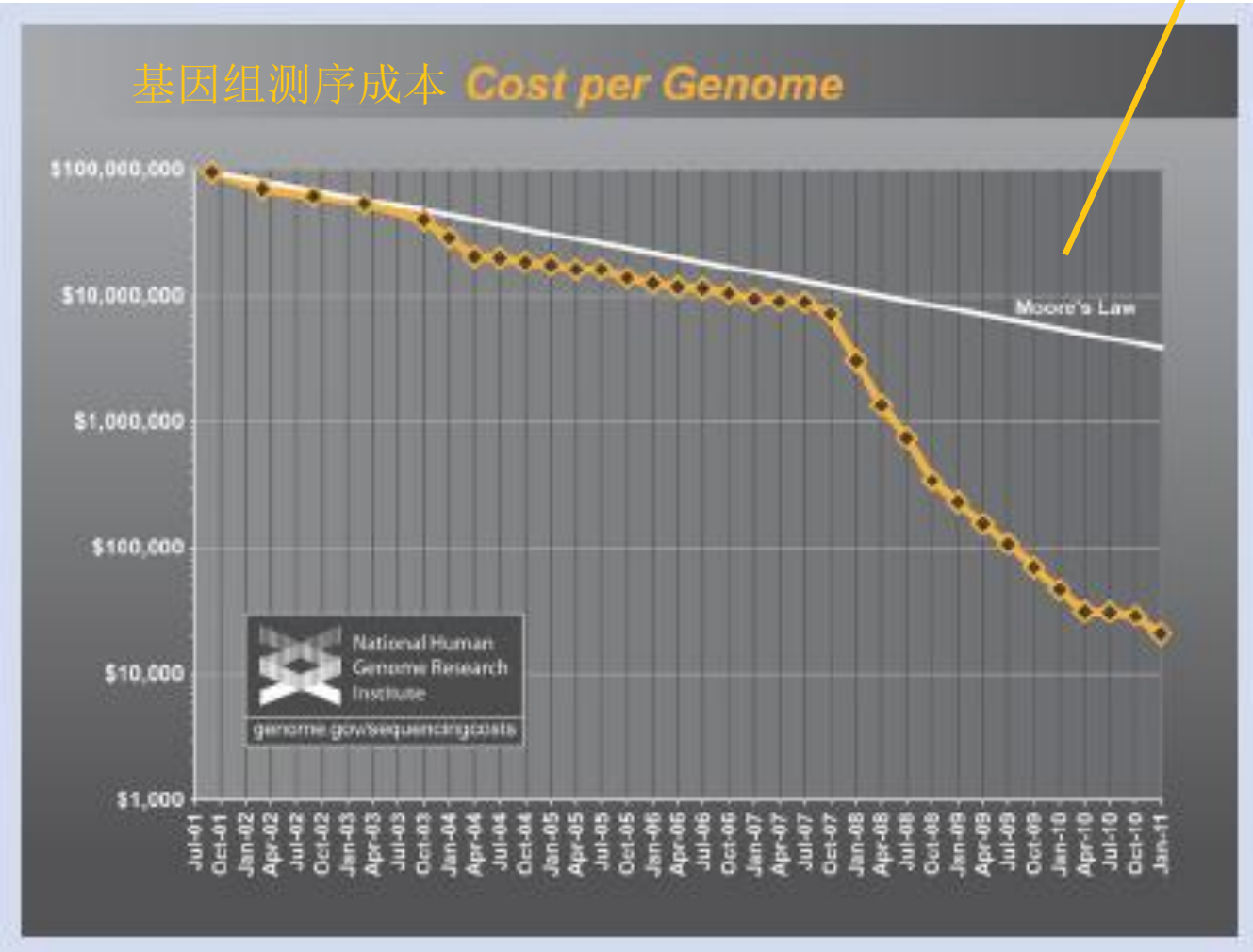
Genomics england

U.K. Unveils Plan to Sequence Whole Genomes of 100,000 Patients

10 December 2012 4:55 pm | 0 Comments

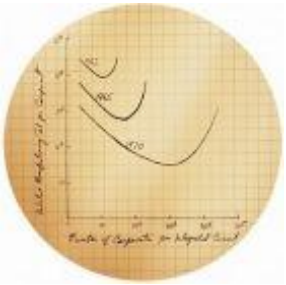
鉴于**英国**医学临床资源不仅规范而且丰富，首相卡梅伦2012年斥资一亿英镑启动了**“十万人基因组测序计划”**。

基因组成本：“超”摩尔定律=？？定律



摩尔定律
(Moore's Law):
1965年由Intel创始人之一戈登·摩尔（Gordon Moore）提出。

当价格不变时，集成电路上可容纳的晶体管数目，约每隔18个月增加一倍，性能也提升一倍。即每一美元所能买到的电脑性能，将每隔18个月翻两倍以上。这一定律揭示了信息技术进步的速度。



DNA测序数据库的大小每18个月会增加10倍. (NBB,2012)

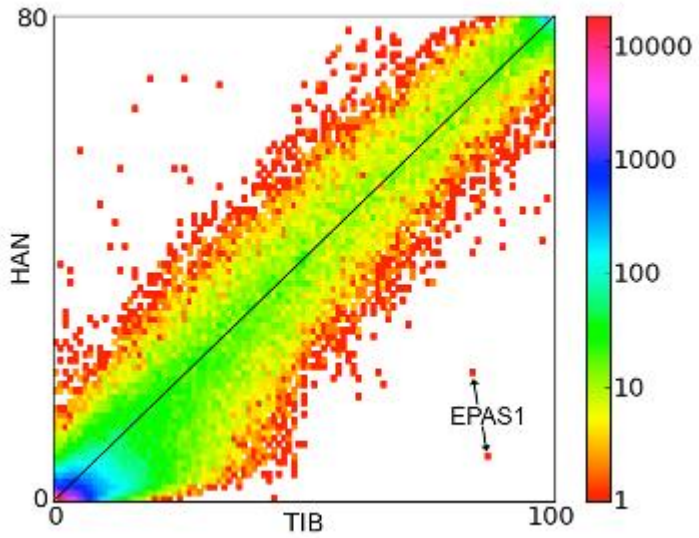
通过对大量样本做基因组的比较分析，挖掘基因与生命，基因与健康的关系，从而推动基因组医学的发展

基因与高原适应

What is the genetic basis of difference in gifted ability to adapt for high altitude?

X. Yi, et.al. 2010. *Science* 329:75-8.

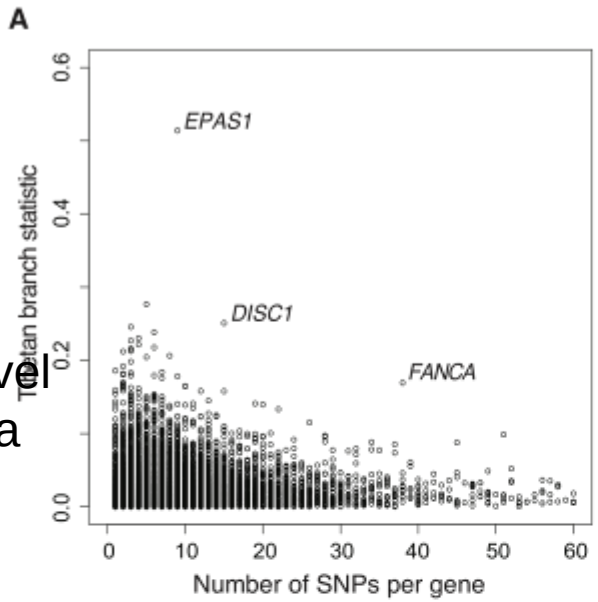
50个藏族人的全部基因测序

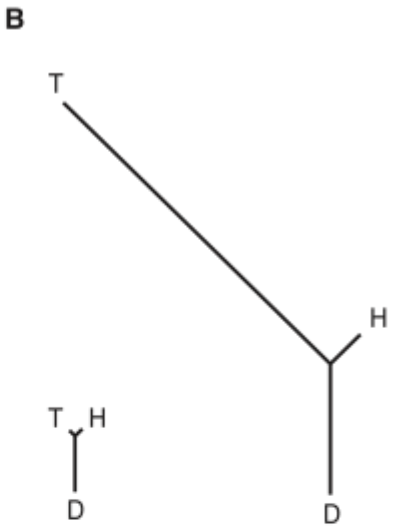


EPAS1: endothelial Per-Arnt-Sim (PAS) domain protein 1

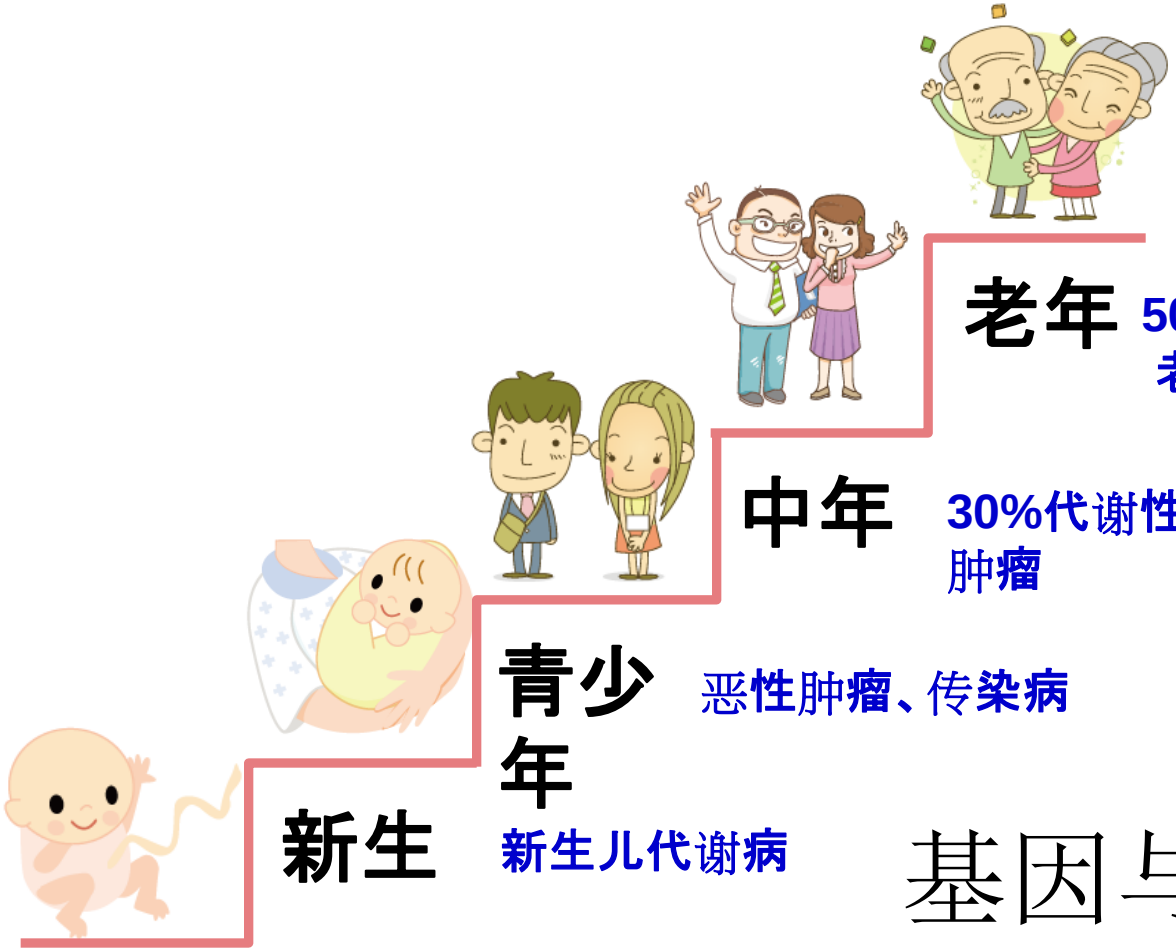
通过分析，找到高原适应相关基因：
EPAS1基因突变：89% 藏族人 VS 9% 汉族人

The signal of selection





Function further validated in
-Association with blood hemoglobin level
-Expression level difference in placenta



老年 50%心脑血管病，肿瘤、老年退型性病

中年 30%代谢性疾病、心脑血管病、肿瘤

青少年 恶性肿瘤、传染病

新生 新生儿代谢病

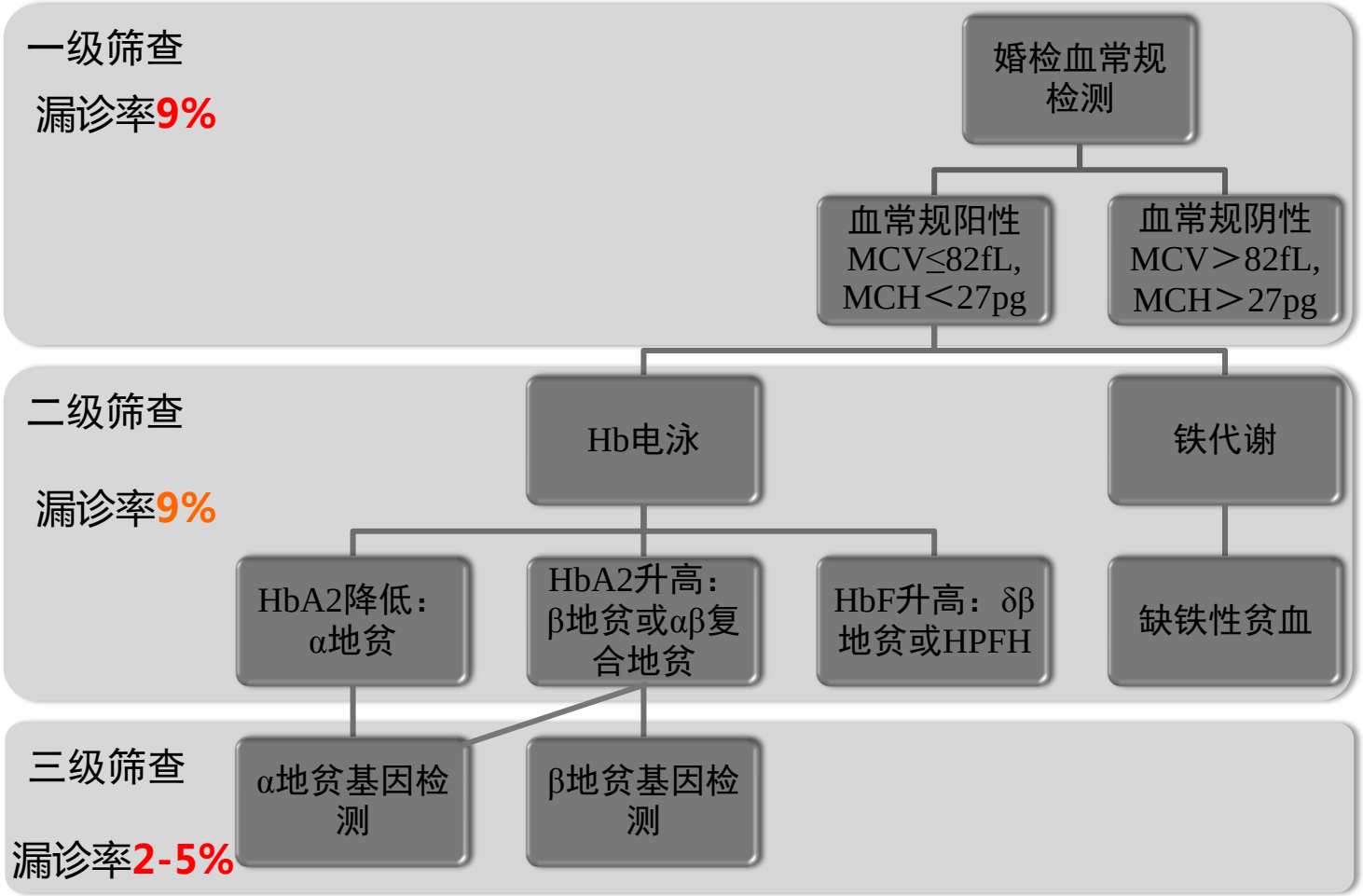
孕中 15%不孕不育
5%出生缺陷

基因与生老病死

40～50%决定性

地中海贫血：
遗传病，贫血，严重的致死或致残

现有地贫筛查流程



为什么不从根本（基因）出发去解决问题？

一、二级筛查漏诊率文献： MCV、MCH、RBC 计数和血红蛋白电泳在地中海贫血筛查中的诊断价值（刘丹，广东佛山市妇幼保健院，中国妇幼保健，2010年25卷）
三级筛查漏诊率出处：广州金域地贫基因检测报告

唐氏综合征

血清学筛查，B超 → 核型分析

唐氏综合征本质上是染色体异常（基因病）

为什么不从根本出发去解决问题？

现代医学的倒行逆施？舍本逐末？

地中海贫血

- 检查指标（MCV，MCH等各项指标）→ 蛋白电泳 → 基因检测

唐氏综合征

- 甲胎蛋白、。。。NT等指标 → 基因检测

为什么？

科学发现；技术限制；成本

基因组医学：从倒行逆施到内外兼修

果

肉眼
CT
光镜
电镜
PHYNOTYPE
生理
生化
机能
检测

因

DNA
RNA
PROTEIN

外

内

基因组医学（GENOME MEDICINE）

基因组医学（Genome Medicine）：人类基因组图谱完成后，在2003年纪念DNA双螺旋结构发现50年时，由世界上600多著名科学家提出的一个医学研究的新概念。

以人类基因组为基础，将生命科学与临床医学整合在一起，使基因组的研究成果迅速高效地转化并应用于临床医学实际。

最重要的特点：

数据驱动

1. 单个**样本的数据量很大；**
2. **从大样本分析中找出规律并建立应用标准**

“蜜糖天使”——唐氏综合征

智力低下、生长发育迟缓、常伴有五官、四肢、内脏等方面畸形。

没有有效的治疗方案

最好的办法就是通过**产前筛查**和**产前诊断**来进行预防。

著名指挥家:舟舟

产前筛查 -- 生化指标+B超软指标

唐氏血清学筛查：

母亲外周血中AFP、HCG和uE3含量

唐氏B超检测：

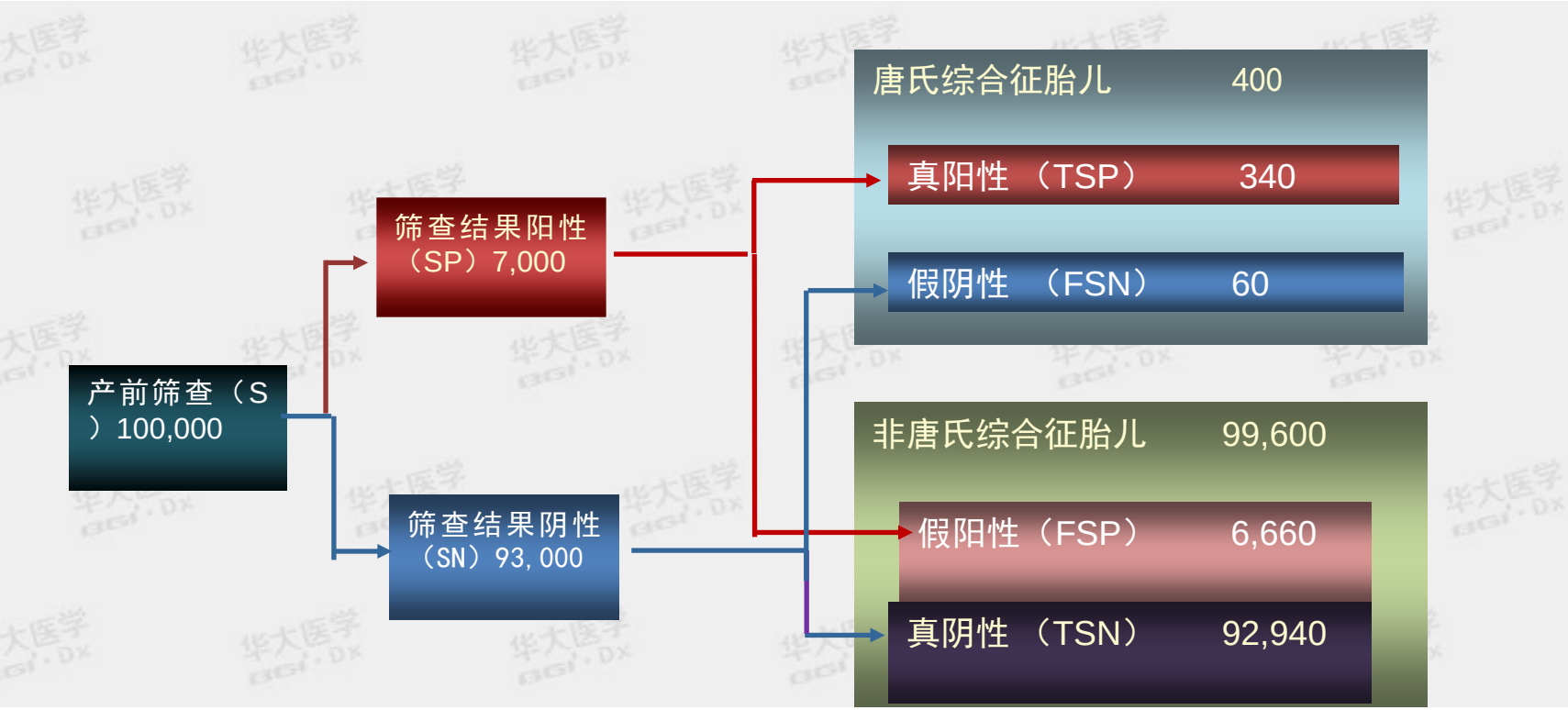
NT/NF

结合孕妇的年龄，体重，孕周，利用公式计算出胎儿为唐氏儿的风险值。

孕妇检查唐氏筛查的高风险率约为5%，但是高风险人群中仅1~2%左右胎儿确实存在染色体异常。

中华人民共和国卫生行业标准
WS 322.1-2010

从过往大数据看现有模式的不足



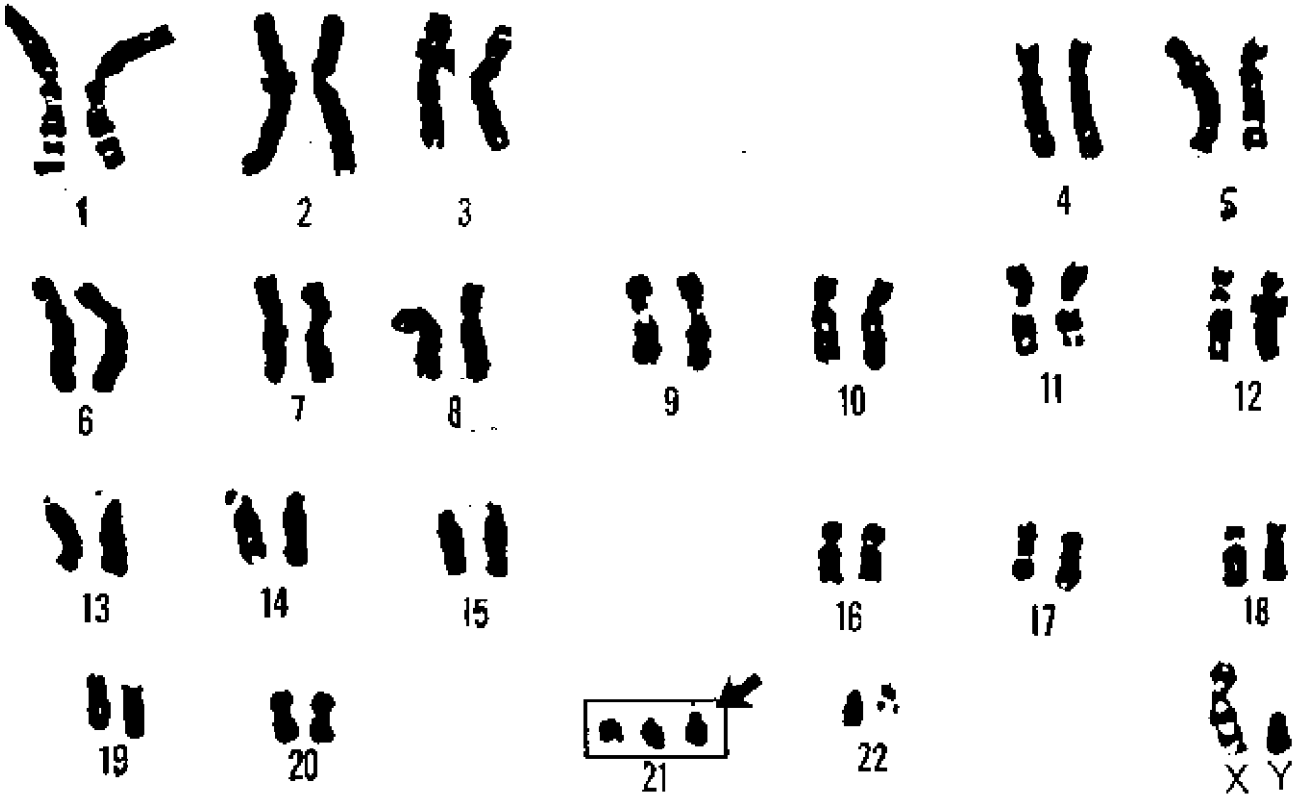
1991-2006
每10万孕妇

假阴性率高，400个唐氏儿有60个漏诊

假阳性率高，6660个胎儿进行了不必要的穿刺，59个正常胎儿流失

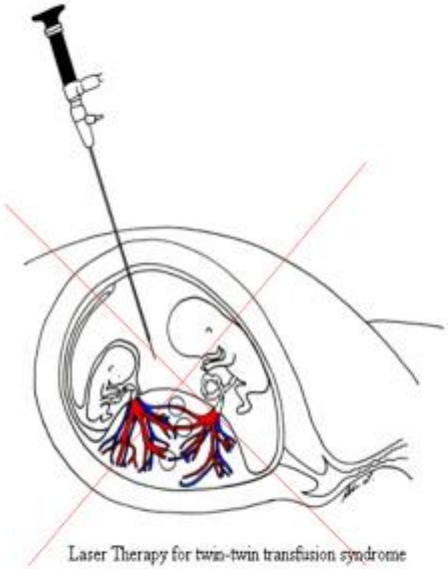
来源：Estimated prenatal screening performance 1991-2006 (2006 data are provisional), England and Wales.

唐氏综合征： 21号染色体为三倍体



正常人23对染色体，每对两条。“唐氏综合征”患者第21号染色体有三条染色体

非损伤性产前诊断



“羊水穿刺”
No !!!

你相信么？？

抽我手臂上的血就能检测我腹中孩儿的健康状况！！！

羊水穿刺致畸率为1－3％！

1997年：发现孕妇外周血中存在胎儿DNA

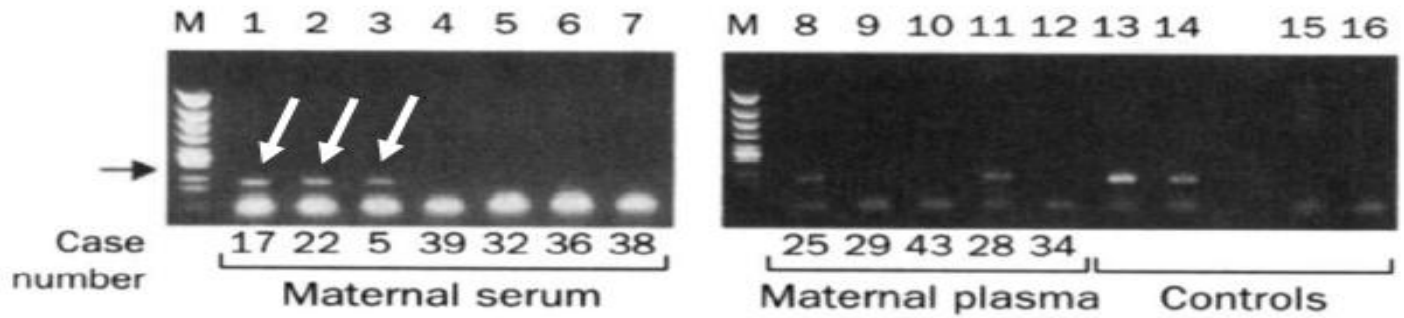
THE LANCET

Early report

Lancet 1997; **350:** 485–87

Presence of fetal DNA in maternal plasma and serum

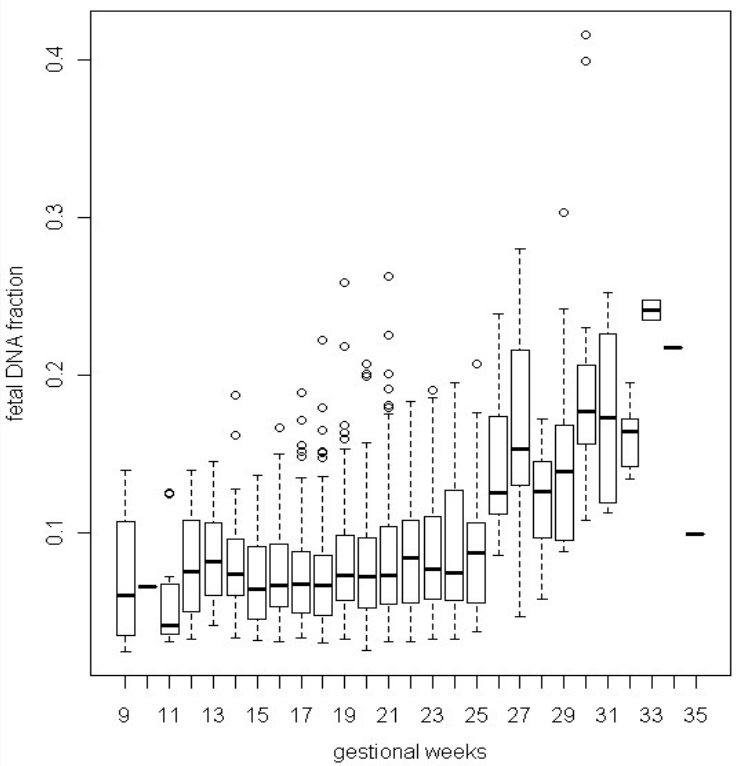
Y M Dennis Lo, Noemi Corbetta, Paul F Chamberlain, Vik Rai, Ian L Sargent, Christopher W G Redman, James S Wainscoat



Amplification of fetal Y-chromosomal sequences from maternal plasma and serum

游离DNA特点

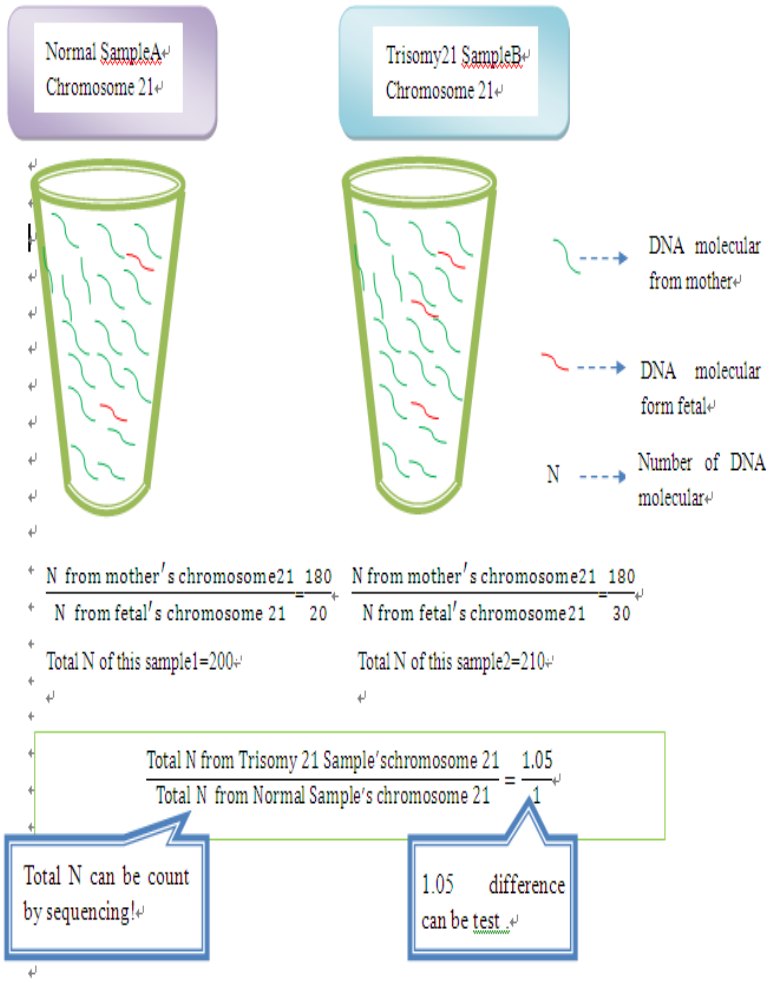
- 母血中存在胎儿游离DNA
- 胎儿游离DNA碎片与母亲DNA片段混杂，平均浓度在7%左右，难以分离。
- 胎儿游离DNA片段较小，平均在166bp，使用常规测序技术（Sanger等）很难检测。（数据量太小，无法区分噪音和信号）

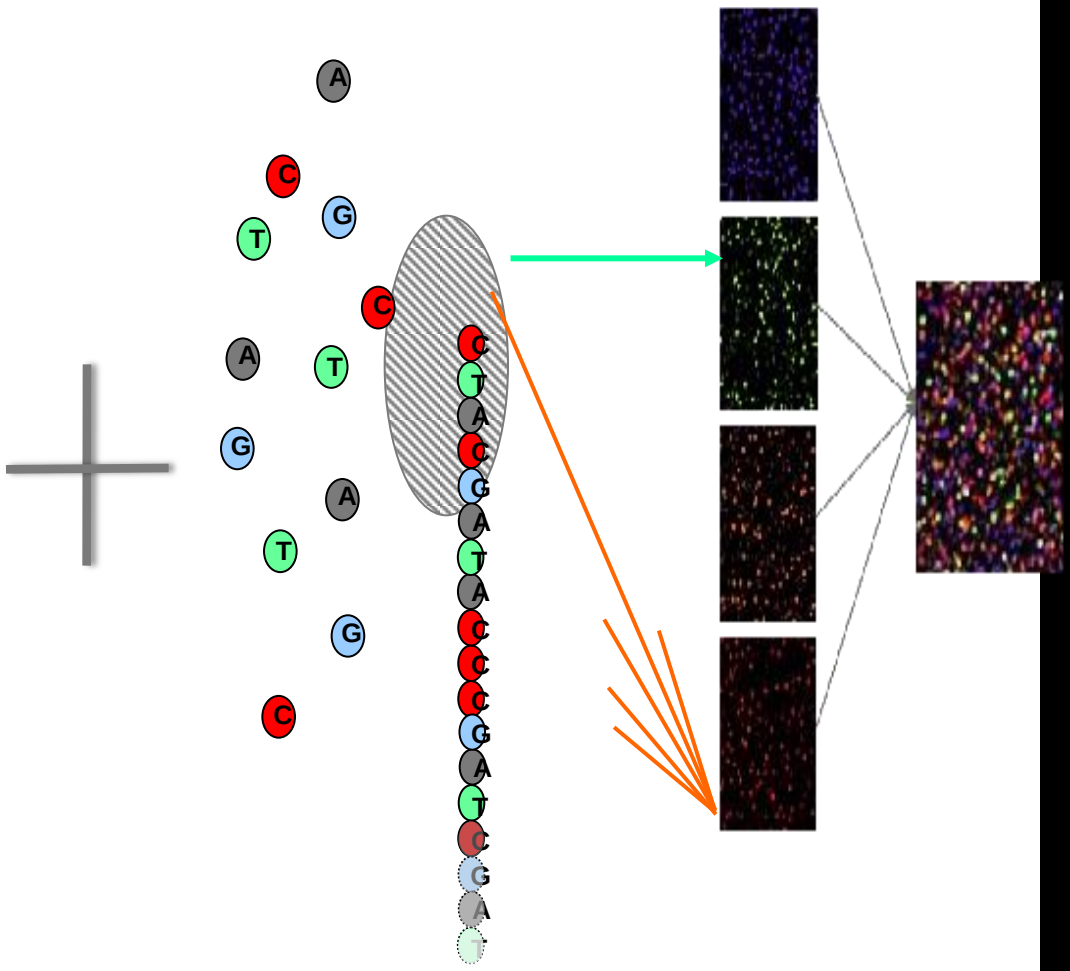


临床试验数据浓度随孕周变化趋势
（华大临检中心内部数据）

自1997年胎儿游离DNA发现以来，因测序技术瓶颈，此发现直至2007年都未能转变为可在临床使用的检测方案。

第二代高通量测序技术（生成大数据）





理论模型

100基因等份 /毫升血浆 （ 含 5等份的胎儿基因 ； 95等份的孕妇基因 ）

孕妇	正常胎儿	21-三体胎儿
11号染色体 :190 21号染色体: 190	11号染色体: 10 21号染色体: 10	11号染色体:10 21号染色体:15

正常胎儿
Chr11:200
Chr21:200

21三体胎儿
Chr11:200
Chr21:205

数据比较: 21三体 > 正常胎儿

➢**检测样本越多，后台数据库越大，参照越可靠，分析的准确度和灵敏度越高。**

无创临床研究（双盲）

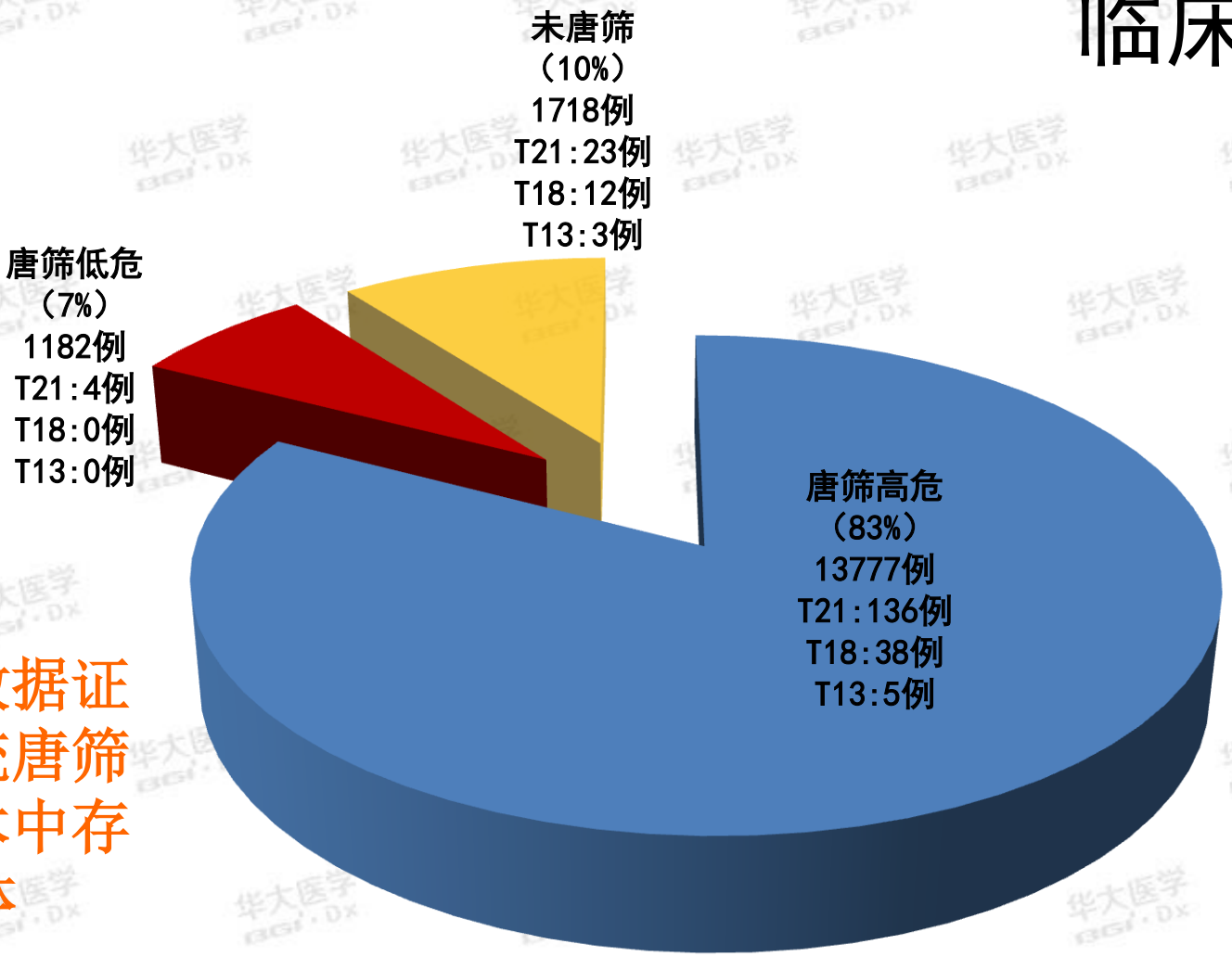
	总例数	检出高风险险样本数		
		T21	T18	T13
样本总数（例）	3000	150	51	9
核型分析（例）	3000	149	50	9
假阳性 （例）	–	1	1	0
假阴性 （例）	–	0	0	0
检出率	–	100%	100%	100%
假阳性率	–	0. 035%	0. 034%	0. 00%
阳性预测值	–	99. 33%	98. 04%	100%
假阴性率	–	0%	0%	0%
阴性预测值	–	100%	100%	100%

临床验证

16677例临床样本的基本情况（2010-2011）

	中位数或百分比	范围或检测例数
年龄(岁)	35	18-45
孕周(周)	20w+5	6w-40w+1
年龄≧35岁	36.72%	6124/16677
进行唐筛	65.43%	10912/16677
唐筛异常	80.63%	8798/10912
超声检测	98.13%	16367/16677
B超异常	3.39%	565/16677
胎儿细胞核型分析	11.06%	1844/16677
胎儿细胞FISH分析	1.76%	293/16677

临床验证

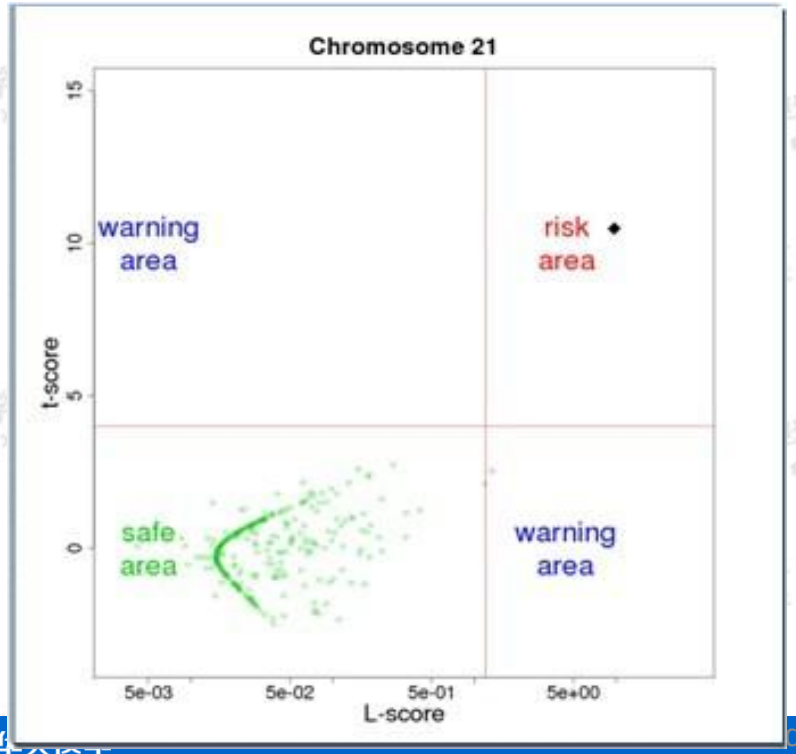


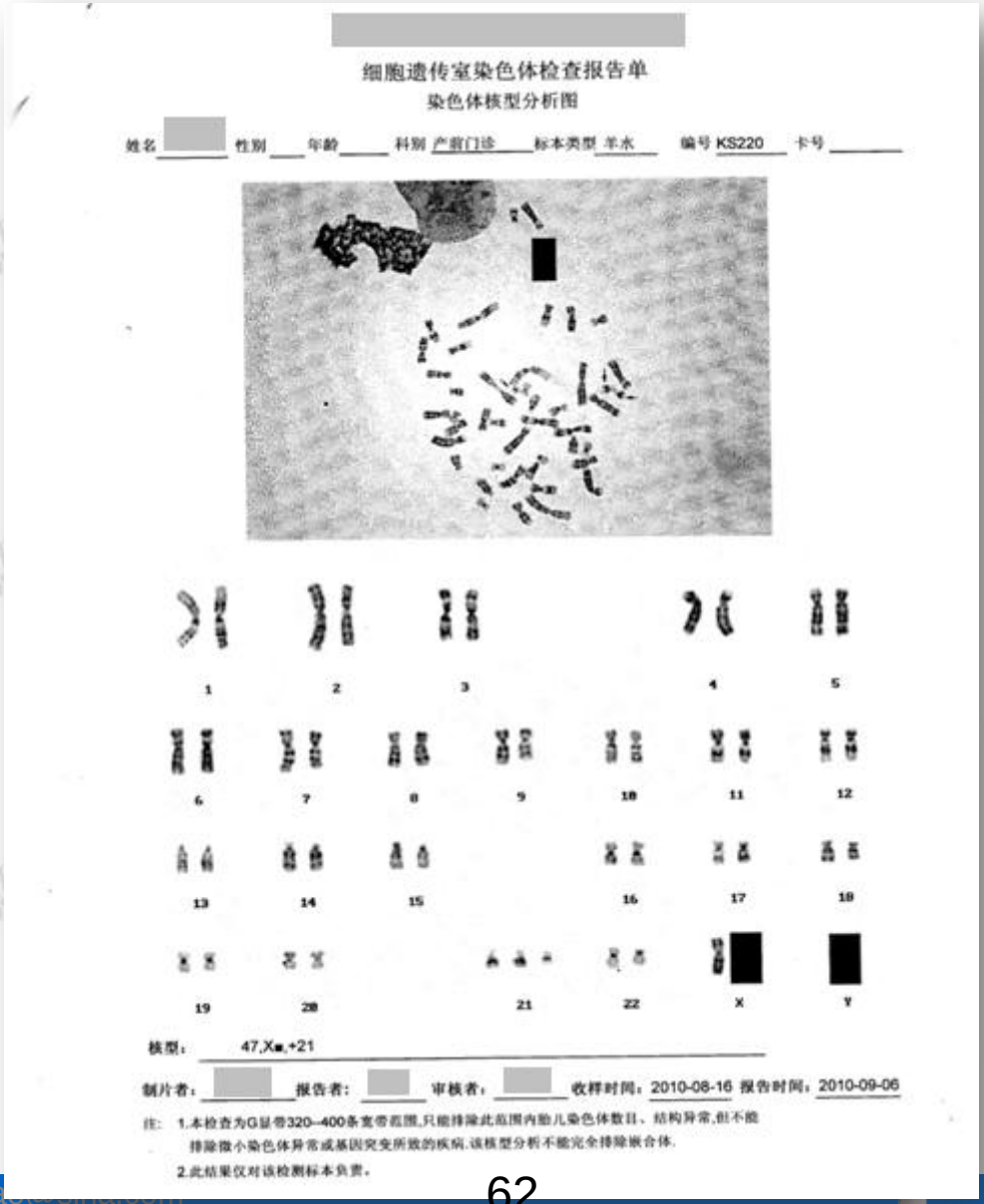
大样本数据证明在传统唐筛低危样本中存在21三体

案例分析

T21

编号:PDP10003761
年龄：31
唐筛结果：1/510（低风险）
NIFTY结果：T21
核型分析：47，XX，+21





细胞遗传室染色体检查报告单
染色体核型分析图

姓名 性别 年龄 科别 产前门诊 标本类型 羊水 编号 KS220 卡号

核型： 47,X■,+21

制片者： 报告者： 审核者： 收样时间：2010-08-16 报告时间：2010-09-06

注：1.本检查为G显带320--400条宽带范围,只能排除此范围内胎儿染色体数目、结构异常,但不能排除微小染色体异常或基因突变所致的疾病.该核型分析不能完全排除嵌合体.
2.此结果仅对该检测标本负责。

国际上无创产前基因检测情况

文献	21-三体综合征		18-三体综合征		13-三体综合征	
	灵敏度	假阳性率	灵敏度	假阳性率	灵敏度	假阳性率
Chiu et al.,2011(2-plex)	100% (89/89)	2.1% (3/146)	N/A	N/A	N/A	N/A
Chen et al., 2011	N/A	N/A	91.9% (34/37)	2.0% (5/252)	100% (25/25)	1.1% (3/264)
Ehrich et al., 2011	100% (39/39)	0.2% (1/410)	N/A	N/A	N/A	N/A
Palomaki et al.,2011	98.6% (209/212)	0.2% (3/1471)	N/A	N/A	N/A	N/A
Palomaki et al., 2012	N/A	N/A	100% (59/59)	0.28% (5/1971)	91.7% (11/12)	0.97% (19/1971)
Bianchi et al.,2012	100% (89/89)	0.0% (0/410)	97.2% (35/36)	0.0% (0/463)	78.6% (11/14)	0.0% (0/485)
Ashoor et al.,2012	100% (50/50)	0.0% (0/297)	98% (49/50)	0.0% (0/297)	N/A	N/A
Lau et al., 2012	100% (11/11)	0.0% (0/97)	100% (10/10)	0.0% (0/98)	100% (2/2)	0.0% (0/106)
合计	99.4 %(487/490)	0.2% (7/2831)	97.4% (187/192)	0.5% (15/3081)	92.5% (49/53)	0.8% (22/2826)

国际上，该技术的综合检出率远远高于现有各种唐氏筛查技术，且假阳性率极低

华大基因检测情况

截至2014年1月17日

- 共完成259836例
- 有核型验证的5254例
- 随访69700例

	阳性例数	假阳性率	特异性	检出率/灵敏度
T21	1874	0.014%	99.99%	99.65%
T18	578	0.020%	99.98%	99.66%
T13	195	0.018%	99.98%	100%

- 其他染色体：其它常染色体异常共168例，性染色体异常350例 ，染色体灵敏度 85.7%（见BMC方法学文章）

- 检测样本越多，后台数据库越大，参照越可靠，分析的准确度和灵敏度越高。

疾病研究策略

传统策略

The candidate gene approach—hypothesis-driven

大数据驱动的研究策略

The genome wide approach—data-driven, large scale analysis

-Slide from Francis Colins

目标序列捕获高通量测序技术

>1700种单基因病仍不知其致病基因

	Autosomal	X-Linked	Y-Linked	Mitochondrial	Total
*** Gene with known sequence**	12488	611	48	35	13182
+ Gene with known sequence and phenotype	317	19	0	2	338
# Phenotype description, molecular basis known	2560	226	4	28	2818
# Mendelian phenotype or locus,molecular basis unknown	**1631**	**136**	**5**	**0**	**1772**
Other, mainly phenotypes with suspected Mendelian basis	1857	135	2	0	1994
Total	18853	1127	59	65	20104

OMIM statistics, July 21, 2010.

运用高通量测序
寻找FREEMAN–SHELDON综合征致病基因

4个Freeman–Sheldon综合征个体全外显子组测序即找到致病基因

Number of genes in which each affected has at least one…	FSS24895	FSS24895 FSS10208	FSS24895 FSS10208 FSS10066	FSS24895 FSS10208 FSS10066 FSS22194	Any 3 of 4: FSS24895 FSS10208 FSS10066 FSS22194
Non-synonymous cSNP, splice site variant or coding indel (NS/SS/I)	4,510	3,284	2,765	2,479	3,768
NS/SS/I not in dbSNP	513	128	71	53	119
NS/SS/I not in eight HapMap exomes	799	168	53	21	160
NS/SS/I neither in dbSNP nor eight HapMap exomes	360	38	8	1 (*MYH3*)	22
…And predicted to be damaging	160	10	2	1 (*MYH3*)	3

Ng, S.B., et al., Targeted capture and massively parallel sequencing of 12 human exomes. *Nature*, 2009. 4

定位米勒综合征致病基因

4 Miller 综合征个体全外显子组测序

致病基因：DHODH, which encodes a key enzyme in the pyrimidine de novo biosynthesis pathway

在另外3个家系得到证实

Table 3 Number of candidate genes identified based on different filtering strategies

	Number of affected exomes			Subsets of 3 exomes		Subsets of all 4 exomes		
Dominant model	1	2	3	Any 1	Any 2	Any 1	Any 2	Any 3
NS/SS/I	4,645–4,687	3,358–3,940	2,850–3,099	6,658	4,489	6,943	5,167	3,920
Not in dbSNP129	634–695	136–369	72–105	1,617	274	1,829	553	172
Not in HapMap 8	898–979	161–506	55–117	2,336	409	2,628	835	222
Not in either	453–528	40–228	10–26	1,317	109	1,516	333	44
Predicted damaging	204–284	10–83	3–6	682	37	787	126	11
Recessive model								
NS/SS/I	2,780–2,863	1,993–2,362	1,646–1,810	4,097	2,713	4,293	3,172	2,329
Not in dbSNP129	92–115	30–53	22–31	226	61	270	90	42
Not in HapMap 8	111–133	13–46	5–13	329	32	397	75	19
Not in either	31–45	2–9	2–3	100	6	121	14	4
Predicted damaging	6–16	0–2	0–1	35	2	44	4	1

Ng, S.B., et al., Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet,* 2010. **42**(1): p. 30-5.

大数据带来新发现（先天性黑蒙）

- Home
- Sponsors
- Gavin's Groupies Gives Back
- Who is Gavin?
- Learning with Gavin
- What is LCA?
- Is there a cure for LCA?
- What is gene therapy?
- Valentine Cookies for a Cause
- What is Gavin's Groupies?
- Gavin's Golf Classic - 2011
- Tour de Sight
- Contact us
- Events/Fundraisers
- A Night Without Light
- Bunco!
- Alternative Gift Giving

Who is Gavin?

Gavin Robert Stevens was born on October 2, 2008, to proud parents, Troy and Jennifer. It was also on this day, Landon (age 6), became a big brother. Gavin was born healthy and with no complications. But it was shortly after his birth, Jennifer had a mother's intuition something was not right. Routine pediatrician appointments did not show any concern with his vision. It was still very unsettling and as the days went by our worry grew. We finally saw a pediatric ophthalmologist four months later after asking for a request to a specialist.It was on that day, our suspicions were confirmed. Within moments of looking into Gavin's dilated eyes, the doctor slumped back into his chair. He looked at us, and very honestly informed us that Gavin's condition was "devastating to his vision", and that we would have to see a pediatric retina specialist in the weeks to follow. It was at this moment our lives were turned upside down. We walked out of the doctors office, with our little 4 month old baby feeling brokenhearted and overwhelmed with sadness and confusion. We still didn't have all the answers, and certainly even more unanswered questions. Months later, he was diagnosed with LCA by Dr. Lee at Children's Hospital Los Angeles. A disease so unfamiliar, we have never heard of it. The words "there is no cure" rang loud in our ears. The grieving process started. We mourned for his vision, but we didn't lose focus of the big picture. Gavin is happy and healthy. We will fight for him to visually enjoy the world as you and I do. At this point, Gavin has one of the severest forms of LCA, as it can present itself differently with each individual affected with LCA. At best, he has light perception, with no visual acuity or central vision. You can say Gavin is completely blind. Our hope is that Gavin and many others who live their lives in darkness due to retinal diseases, may have an option to receive treatment to correct their vision, when it is available.

LCA:先天性利伯氏黑蒙

- 常染色体隐性遗传
- 1:80000
- 眼球震颤、呆滞、无瞳孔反射、失明
- 基因治疗

Gavin家兄弟姐妹4人，其中2个孩子是LCA患者

筛查LCA 17个已知相关基因，未见致病突变

全外显子组基因检测

↓

2460 未经报道的变异

↓

10 个无义突变改变了蛋白质的表达

↓

5 个突变表达在视网膜上

↓

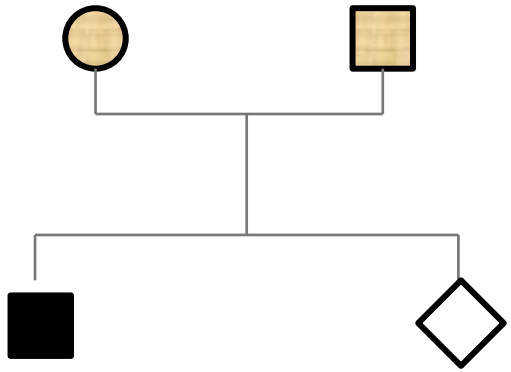
NMNAT1 ,LCA9 ,1p36.22

验证

Sanger法验证

c.769G>A, p.Glu257Lys
错义突变

c.507G>A, p.Trp169*
无义突变



769G>A p.Glu257Lys
507G>A p.Trp169*

+/+

P W Chiang etal. Exome sequencing identifies *NMNAT1* mutations as a cause of Leber congenital amaurosis. ***Nature Genetics*** 44, 972–974 (2012)

在9个无亲属关系的家庭中的10个病例中，均发现了NMNAT1突变

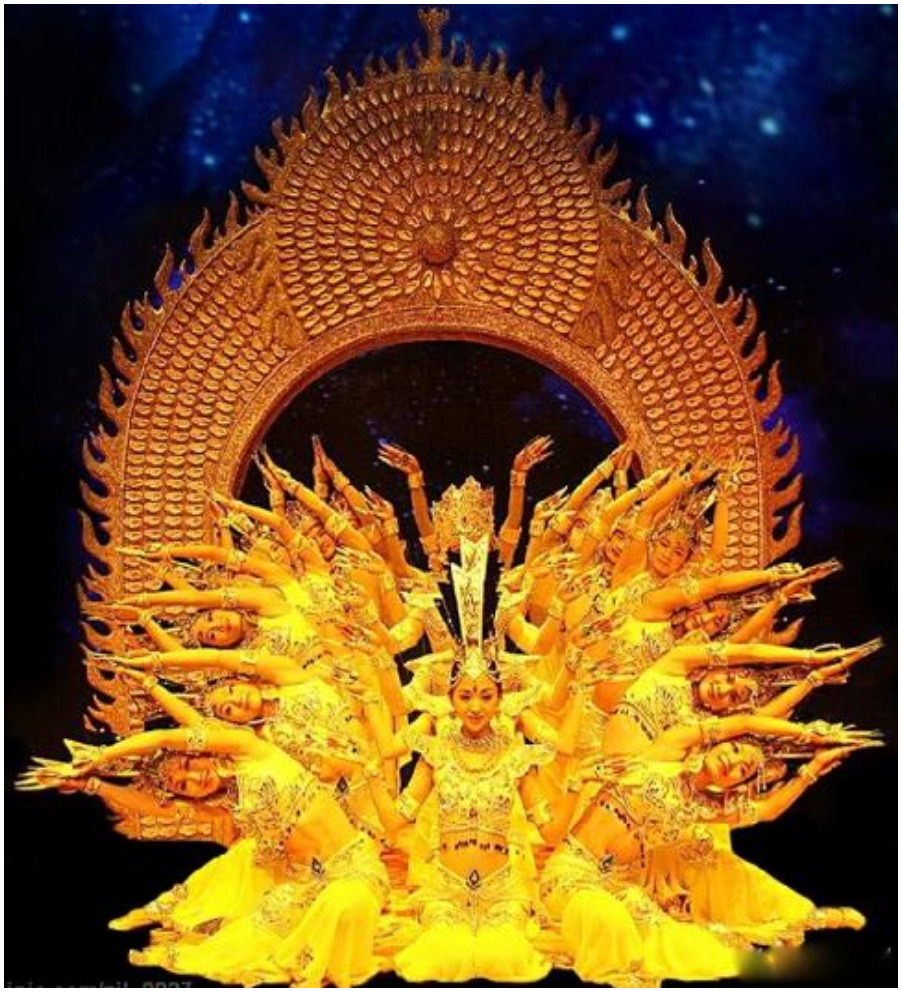
世界首例外显子检测结果对临床治疗方案提供依据

Bainbridge, J. W. B., Smith, A. J., Barker, S. S. et al. (2008).
Effect of gene therapy on visual function in Leber's congenital amaurosis.
The New England journal of medicine, 358(21), 2231–2239.

现在： 饮食正常，体重恢复，炎症消失，并发症消失，自主产生正常血细胞。
6岁，和正常人一样健康地生活、学习。

GAVIN的故事告诉我们什么？

样本的收集和存储很重要（探索及验证）
大数据库很重要（排除噪音：不致病的基因突变）
算法很重要



•我国现有听力言语残疾者2780万人，居各项残疾之首，每年大约3万例重度耳聋患者出生。

•遗传因素占耳聋发病因素的60%左右。

•遗传性耳聋目前没有根治的办法，只能预防。国内外学者建议新生儿及准备结婚、生育及产前的夫妇双方，都进行耳聋基因检测。

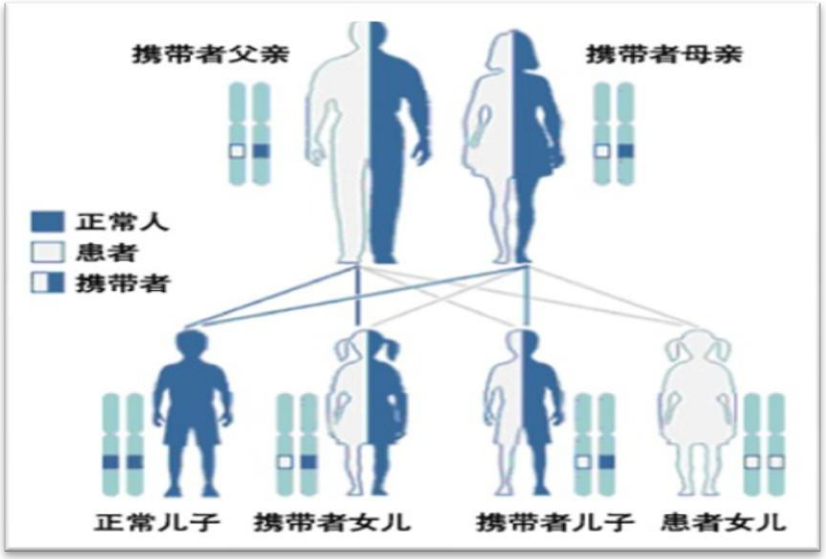
遗传性耳聋

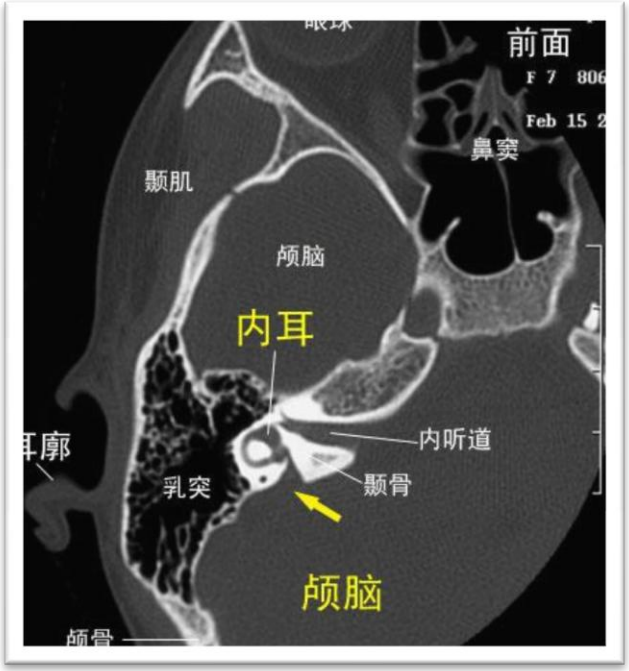
◆儿童期耳聋：

26%-33%为GJB2突变

20%-28%为SLC26A4突变

常染色体隐性遗传





耳聋基因检测

耳聋检测类型	检测意义
耳聋基因筛查（4个基因20个位点）	涵盖了携带率最高的20个位点
遗传性耳聋基因全测序（51个基因+2个突变位点）	针对有家族遗传史的，常规筛查方法未能发现突变的人群

- 有助于第一时间发现耳聋基因携带者或迟发型听力损失高危儿，可以早发现，早干预，不丧失语言能力。
- 针对携带12S rRNA突变的新生儿，需要指导抗生素的应用，可以避免药物性耳聋产生。

- 耳聋基因携带率
  - 3个基因8个位点筛查：~3.4%（佛山1万人数据）
  - 4个基因20个位点筛查：5.6%（天津7万人数据）
  - 仍然有很多不知道致病位点

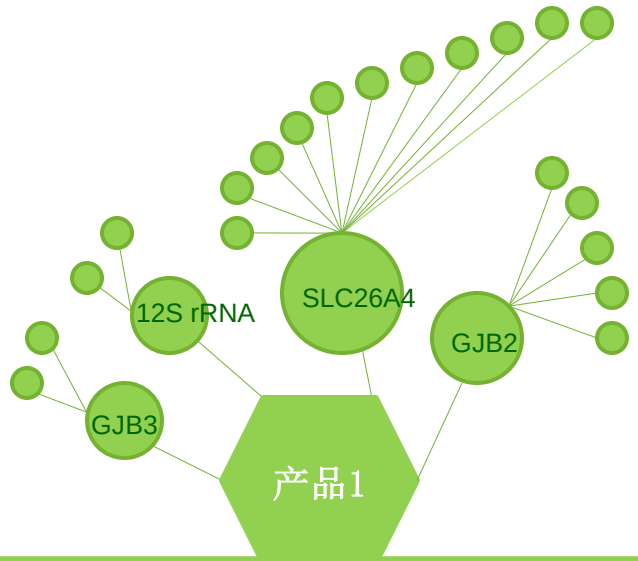
一个真实的故事

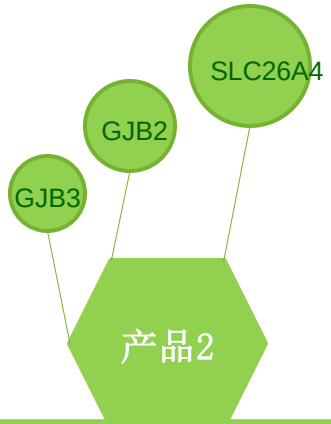
林先生，清远人，夫妻俩听力正常，生育两个女儿均为极重度耳聋患者。林先生为生育一个听力正常的孩子，去全国几家听力学检测的专家医院挂号就诊，希望能检测出两孩子均为耳聋的原因，均未检测出患病原因。

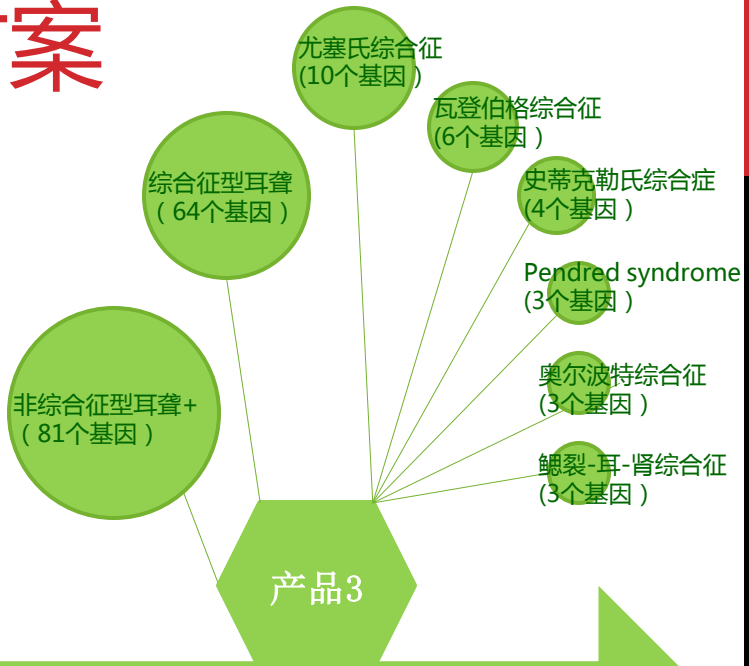
一个真实的故事

通过我们的科学设计，发现两孩子均为TMC1基因上新发的突变，且两孩子均为纯合子，父母为此突变的杂合子，此突变为常染色体隐性遗传，用目前临床上利用的基因检测方法尚不能检测。

耳聋基因检测完全解决方案







遗传性耳聋
4个常见基因检测

1）新生儿筛查
2）婚前、孕前、产前筛查

遗传性耳聋
3个基因全检测

1）新生儿后续检测
2）婚前、孕前、产前后续检测

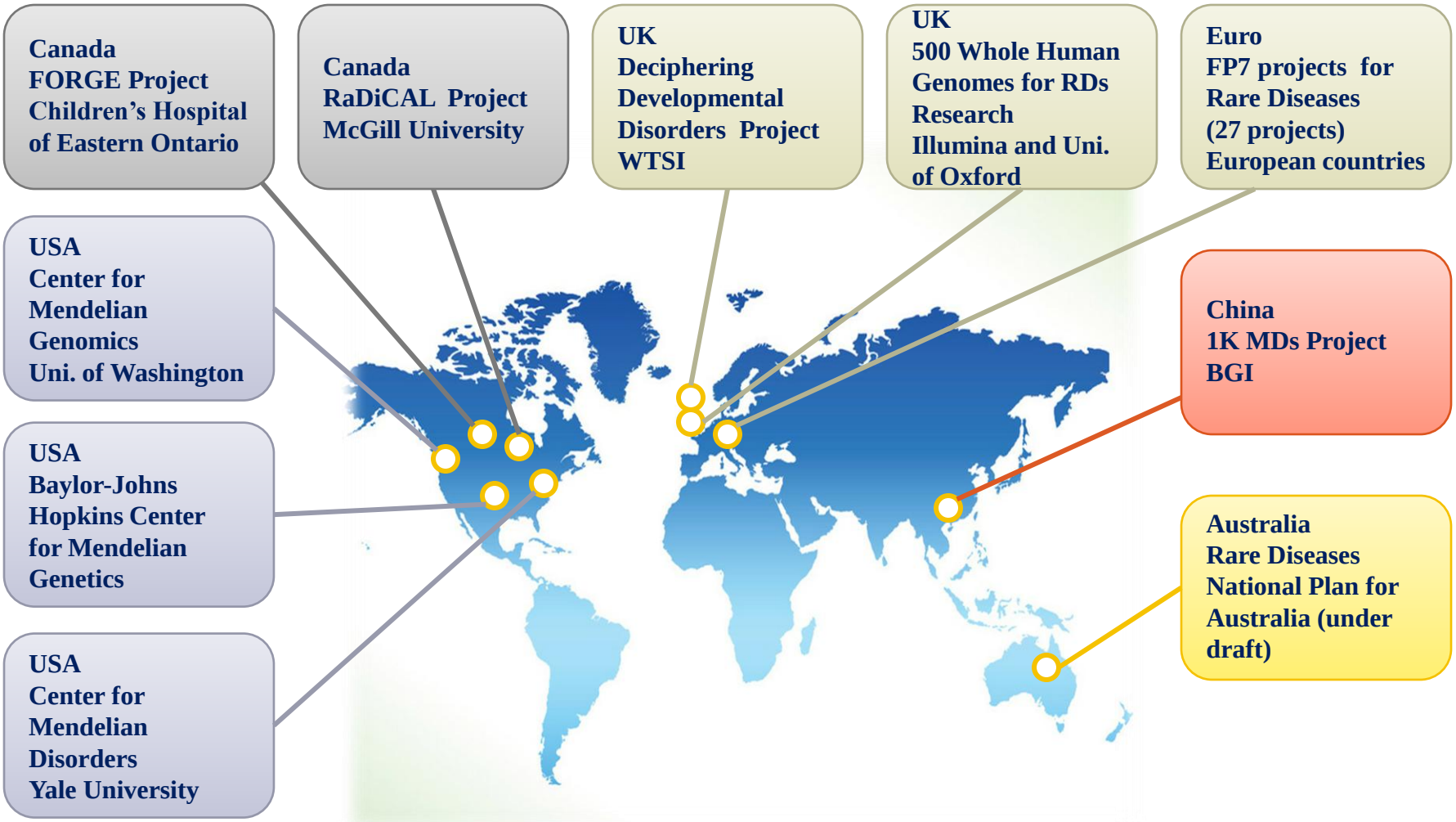
遗传性耳聋
127个基因检测

1）非综合征型及综合征型耳聋患者诊断
2）耳聋遗传史家系成员检测

百万新生儿听力及耳聋基因筛查计划

301医院与华大基因牵头，启动百万新生儿听力及基因联合筛查计划

国际罕见病研究联盟


Canada
FORGE Project
Children's Hospital of Eastern Ontario
Canada
RaDiCAL Project
McGill University
UK
Deciphering Developmental Disorders Project
WTSI
UK
500 Whole Human Genomes for RDs Research
Illumina and Uni. of Oxford
Euro
FP7 projects for Rare Diseases
(27 projects)
European countries
USA
Center for Mendelian Genomics
Uni. of Washington
USA
Baylor-Johns Hopkins Center for Mendelian Genetics
USA
Center for Mendelian Disorders
Yale University
China
1K MDs Project
BGI
Australia
Rare Diseases National Plan for Australia (under draft)

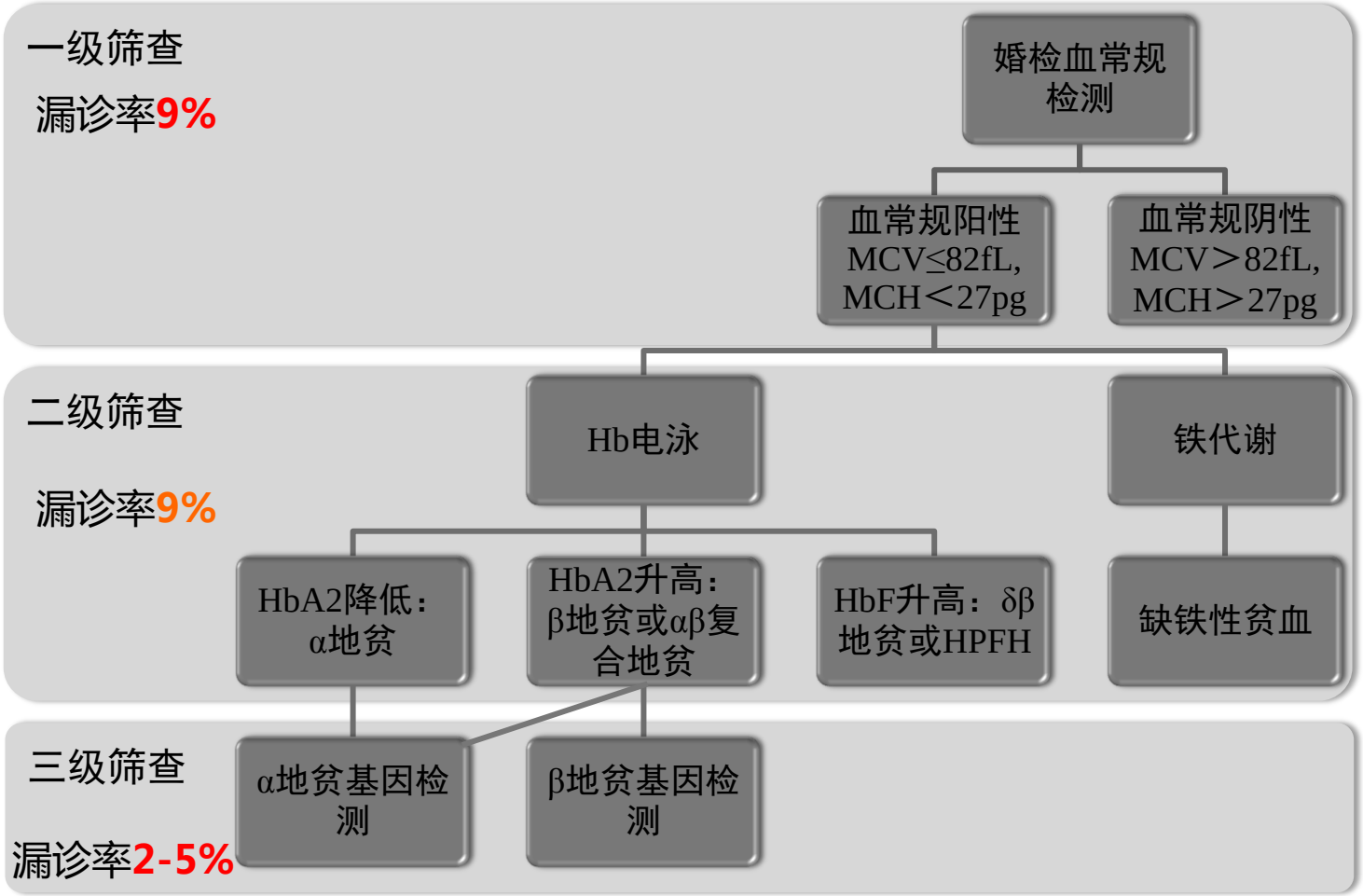
千种单基因病计划

- 收集合作样本3545份
- 合作单位108家
- 涉及疾病近200种

华大健康
BGI health

什么是地贫？
遗传病，贫血，严重的致死或致残，治疗费用高昂

现有地贫筛查流程



为什么不从根本（基因）出发去解决问题？

一、二级筛查漏诊率文献： MCV、MCH、RBC 计数和血红蛋白电泳在地中海贫血筛查中的诊断价值（刘丹，广东佛山市妇幼保健院，中国妇幼保健，2010年25卷）
三级筛查漏诊率出处：广州金域地贫基因检测报告

地中海贫血的筛查告诉我们什么？

父母表型正常

小孩可以是致死或致残的出生缺陷！

隐性遗传疾病
（单个疾病的发病率都很低，
但综合发病率>1%）

携带突变并不可怕，可怕的是携带方式致病，而本人毫不知情。

产前/孕前 是预防遗传病的最佳时机

进行性假肥大性肌营养不良

110 杜氏(型)肌营养不良的腓肠肌肥大。

- **X连锁隐性遗传，致病基因DMD**
- **其中2/3是由于母亲遗传，1/3是自身突变**
- **翼状肩胛、假肥大、Gower征**
- **肌酶、肌电图、肌活检**
- **2-8岁起病，12岁左右不能走动，20岁左右死亡**

血友病

- X连锁隐性遗传
- 甲型（F8）77%, 乙型（F9）19.6%
- 关节腔出血，疼痛难忍
- 避免外伤、手术出血，凝血因子替代疗法

肾上腺皮质增生

- 常染色体隐性遗传，90%为21羟化酶缺乏型（CYP21）
- 女孩男性化，男女发病比例1:2
- 失盐型醛固酮↓，K+ ↑　Na+↓

脊肌萎缩症

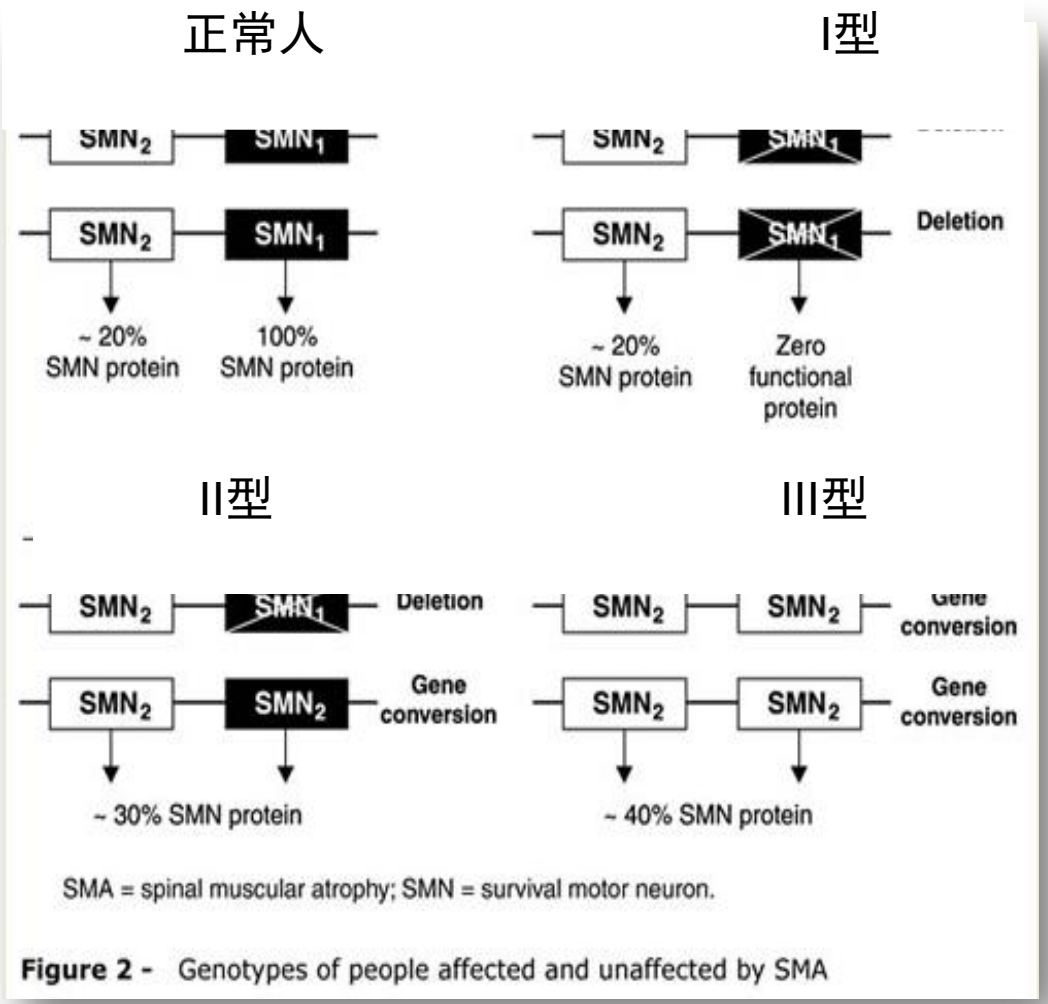


Figure 2 - Genotypes of people affected and unaffected by SMA

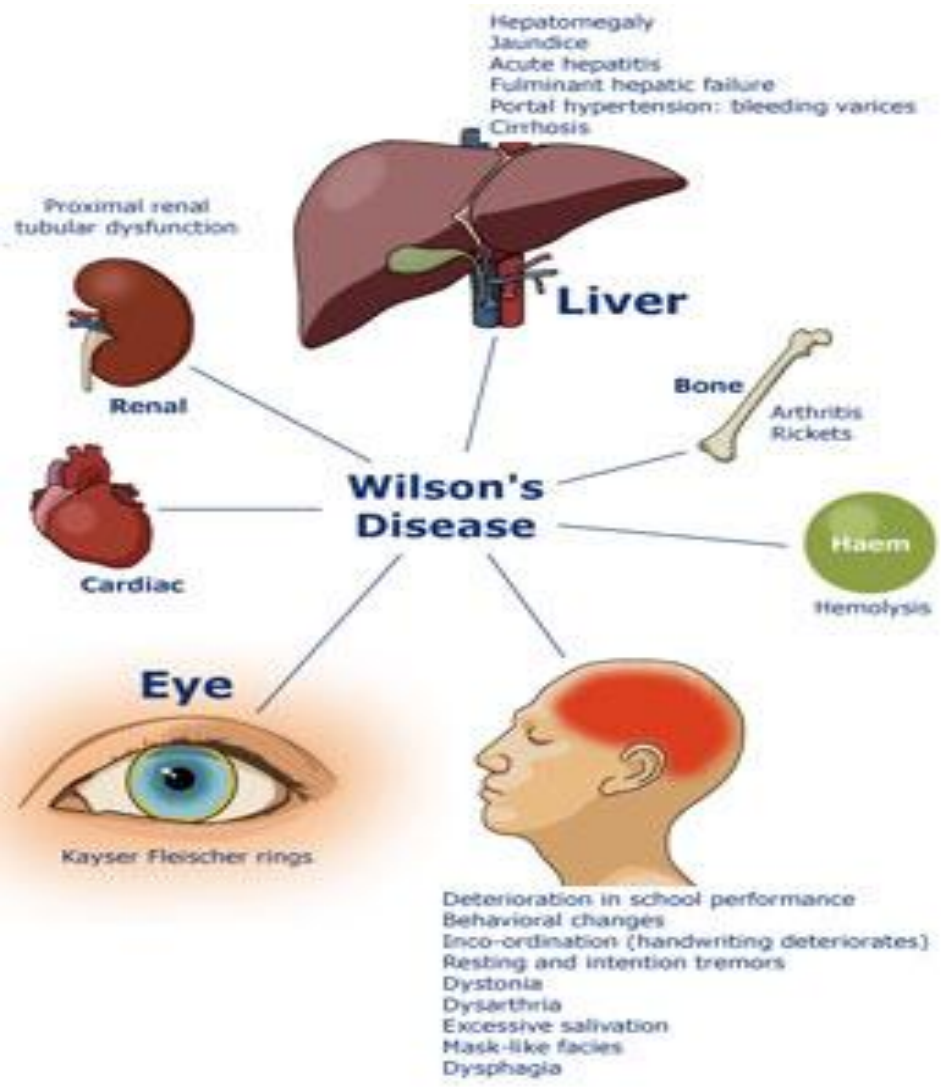
- **常染色体隐性遗传，致病基因为SMN1**
- **表现为肌肉无力，严重程度与近似基因SMN2的拷贝数有关。I型：2-4月发病；II型：6-12月发病；III型 2-30岁发病。**
- **失盐型：醛固酮↓，K+↑ Na+↓**
- **肌电图：脊髓源性损害**

肝豆状核变性

◆常染色体隐性遗传

◆ ATP7B基因突变引起的铜代谢障碍

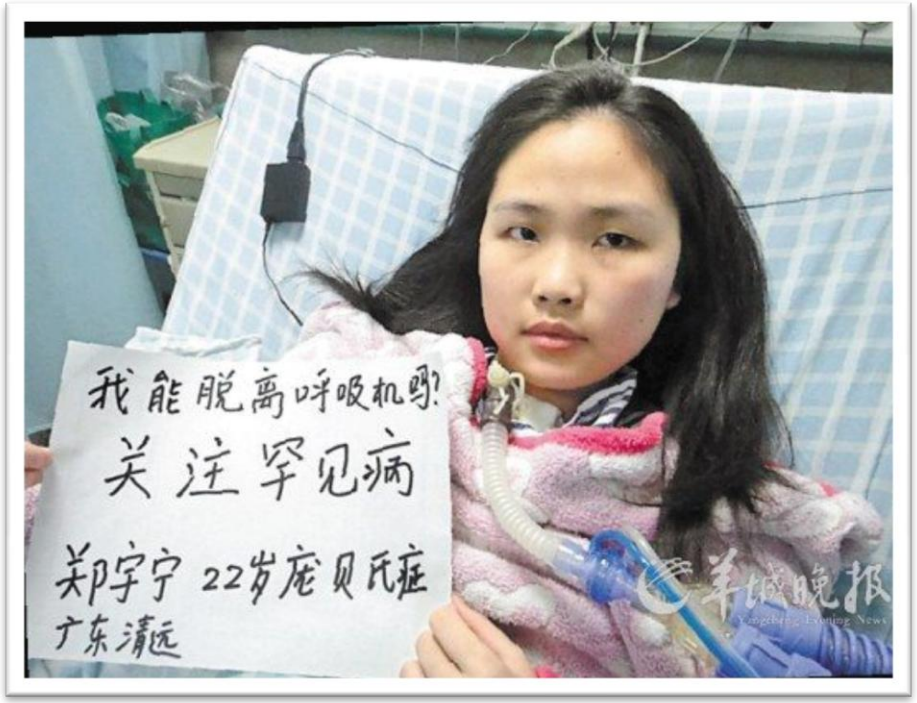
◆ 进行性加重的肝损害、神经系统症状、角膜外缘色素环



糖原累积病II型（POMPE）

◆ 常染色体隐性遗传

◆ 酸性α糖苷酶（GAA）缺乏，溶酶体糖原堆积，导致全身肌张力低，运动后肌力快速下降，肌电图呈肌源性损害。



隐性遗传病孕前筛查

Carrier Testing for Severe Childhood Recessive Diseases by Next-Generation

Bell *et al*. Science Translational Medicine,2011

104名个体

448种遗传病

437个基因

每个个体携带3～4个隐性杂合突变基因

多样本、多病种、多基因同时筛查

中国12种常见隐性遗传病（综合携带率~28%）

序号	疾病	发病率	携带率	基因	遗传方式
1	进行性假肥大性肌营养不良	1/3500 男性		DMD	X 连 锁 隐性遗传
2	血友病（甲型、乙型）	1/4000-1/5000 男性		F8、F9	
3	遗传性耳聋	1/1000新生儿	1/16	GJB2、SLC26A4	常染色体隐性遗传
4	肾上腺皮质增生	1/5000	1/35	CYP21A2	
5	脊肌萎缩症	1/6000	1/40	SMN1	
6	苯丙酮尿症	1/5,000-1/10,000	1/55	PAH	
7	肝豆状核变性	1/30000-1/100 000	1/87-1/150	ATP7B	
8	糖原累积病II型	1/40000新生儿	1/100	GAA	
9	半乳糖血症I型	1/62000新生儿	1/125	GALT	
10	地中海贫血（α、β）	1-5/10000、0.66/100新生儿	南方1/10	HBA1、HBA2 HBB	
11	镰形细胞贫血症	1-5/10000		HBB	
12	隐性多囊肾			PKHD1	

隐性遗传病筛查

- **华大一次性对12种隐性遗传疾病进行临床孕前筛查**

随机抽取185个表型正常育龄男女（**无遗传病家族史**）

结果如何？ → 32人是携带者

携带率：17.3%

华大进行12种隐性遗传病的携带者情况

疾病	基因	总（232）	without 12种遗传病家族史（185）	with 12种遗传病家族史(21)	不良妊娠史or家族其他成员不良妊娠史(21)	近亲结婚（5）
β-地中海贫血	*HBB*	2/232	0/185	1/21	1/21	0/5
遗传性耳聋	*GJB2*	5/232	4/185	1/21	0/21	0/5
	SLC26A4	5/232	4/185	0/21	1/21	0/5
进行性肌营养不良	*DMD*	2/232	0/185	2/21	0/21	0/5
半乳糖血症	*GALT*	0/232	0/185	0/21	0/21	0/5
苯丙酮尿症	*PAH*	12/232	8/185	1/21	2/21	1/5
肝豆状核变性	*ATP7B*	6/232	5/185	1/21	0/21	0/5
糖原累积病II型	*GAA*	1/232	1/185	0/21	0/21	0/5
隐性多囊肾	*PKHD1*	1/232	1/185	0/21	0/21	0/5
乙型血友病	*F9*	0/232	0/185	0/21	0/21	0/5
甲型血友病	*F8*	1/232	0/185	1/21	0/21	0/5
α-地中海贫血	*HBA1*、*HBA2*	4/232	2/185	2/21	0/21	0/5
肾上腺皮质增生	*CYP21A2*	1/232	1/185	0/21	0/21	0/5
脊肌萎缩症	*SMN1*	7/232	6/185	1/21	0/21	0/5
总计	——	**47/232（20.3%）**	**32/185（17.3%）**	**10/21（47.6%）**	**4/21（19.1%）**	**1/5（20%）**

解决方案：

一次性对多种已知治病基因的遗传性疾病做基因筛查（降低成本）
而不是单个疾病单个疾病地筛查

基因与肿瘤

宫颈癌：
唯一病因明确且可预防可根治的癌症

- 宫颈癌：女性恶性肿瘤中第二位，生殖道恶性肿瘤第一位
- 我国宫颈癌：发病率逐年上升，年轻化趋势明显
- 2008年全球估计新发宫颈癌病例52.98万，死亡病例25.51万人，其中85%新发病例在发展中国家。

李媛媛

梅艳芳

宫颈癌常规检测

1、TCT（Thinprep cytologic test）：新柏氏液基细胞学技术。采用液基薄层检测系统检测宫颈细胞并进行细胞学分类诊断，是目前国际上比较成熟的宫颈癌细胞学检查技术。用于检测宫颈脱落细胞，采用细胞学的方式检测宫颈细胞是否发生病变。

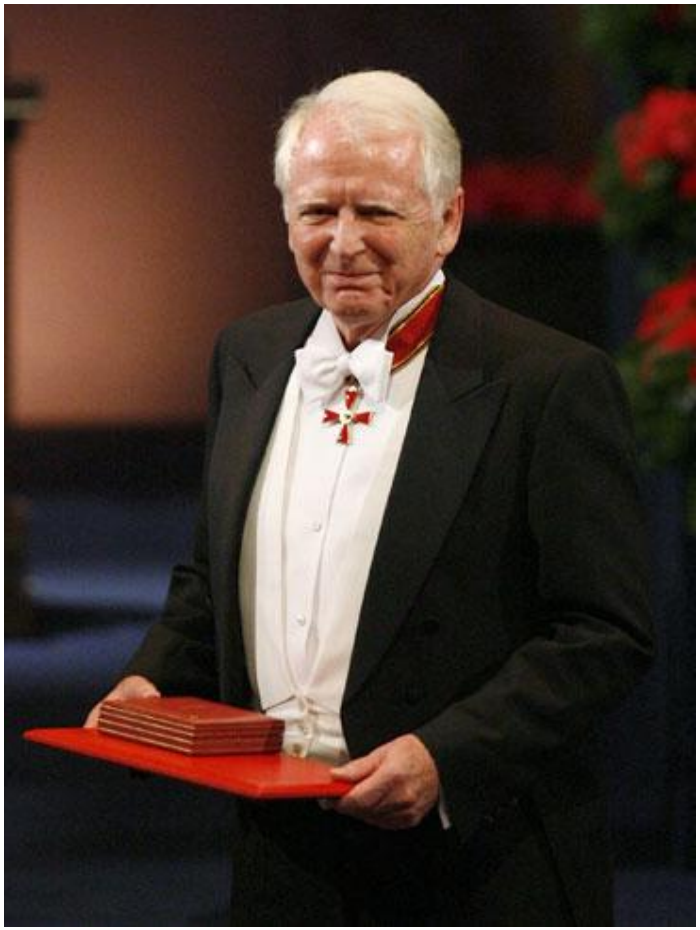
2、阴道镜检查：TCT结果显示为可疑或阳性而肉眼未见明显癌灶着，阴道镜可将病变放大6~40倍，直接观察宫颈上皮及血管的细微形态变化。

3、宫颈和颈管活组织检查：确诊宫颈癌前病变和宫颈癌的最可靠和不可缺少的方法。一般在阴道镜指导下，在可疑癌变部位取多点活检，送病理检查。

可检测的前提是细胞/组织已经发生病变

HPV与宫颈癌的关系

1995年，国际癌症研究机构（IARC)专题讨论会通过——高危型HPV感染是宫颈癌的主要原因



2008年HPV发现者获诺贝尔奖

HPV检测技术

1	21	41	61	81
2	22	42	62	82
3	23	43	63	83
4	24	44	64	84
5	25	45	65	85
6	26	46	66	86
7	27	47	67	87
8	28	48	68	88
9	29	49	69	89
10	30	50	70	90
11	31	51	71	91
12	32	52	72	92
13	33	53	73	93
14	34	54	74	94
15	35	55	75	95
16	36	56	76	96
17	37	57	77	97
18	38	58	78	98
19	39	59	79	99
20	40	60	80	100

生殖器疣　　致癌型

- ✓ 覆盖所有致癌型别
- ✓ 精确检测到样本中低至１－３拷贝的病原体；
- ✓ 血液，尿液，与组织样本中的任何病毒；
- ✓ 一次性病毒基因分型；
- ✓ 超低成本
- ✓ 微测序原理，数字信号

HPV与宫颈癌 –
白种人的疫苗,我们适合么?

美国《人乳头状瘤病毒疫苗政策和免疫计划指南》全面展开。目前宫颈癌疫苗已经有3.3万人份的临床试验报告送交中国政府，等待审批。宫颈癌疫苗在美国的价格是360美元。含6、11、16、18共四型。

中国超过20万份HPV普查数据显示，发病率最高的是：16,52,18,58四种型别

对症下药 → 对因下药

易瑞沙（吉非替尼）是一种新型的小分子量肿瘤治疗药物，临床主要用于治疗既往接受过化学治疗的局部晚期或转移性非小细胞肺癌（NSCLC），尤其对肺腺癌疗效确切。

易瑞沙（吉非替尼）

其作用机制主要是通过抑制EGFR自身磷酸化而阻滞传导，抑制肿瘤细胞的增殖，实现靶向治疗。

易瑞沙——EGFR基因敏感突变的晚期NSCLC患者一线治疗里程碑

自IPASS研究公布至今已有3年，期间IPASS的结果一次次地被证实，从WJOG3405 到NEJ002 ，从OPTIMAL到EURATC，我们都看到了EGFR基因敏感突变的NSCLC患者一线使用TKI治疗的客观缓解率可以达到约70%，中位PFS可以达到10个月左右，中位OS超过20个月，这都是在之前单纯化疗时代从未出现过的。EGFR基因突变真正登上了NSCLC治疗的舞台。虽然由于后续大量交叉治疗的影响，这些研究并没有看到突变患者一线使用TKI能够带来显著的生存获益，但考虑到TKI带来的卓越的缓解率、PFS和生活质量的改善，ASCO、ESMO指南仍然推荐EGFR-TKI作为突变阳性患者的一线治疗。

目前，整个NSCLC的分类基础已经由之前的腺癌鳞癌转变为EGFR突变型和EGFR野生型。整个治疗策略也随之发生了变化，对于今后的NSCLC一线患者，我们都应尽快明确其EGFR基因突变状态，对于EGFR基因敏感突变的患者首选采用TKI的治疗。

在中国，易瑞沙是目前唯一拥有NSCLC一线突变适应症的EGFR-TKI ，同时它也被全球66个国家批准用于EGFR基因敏感突变的晚期NSCLC治疗。

ONC-1207-IR-0165

EGFR基因敏感突变的NSCLC患者一线使用TKI治疗的客观缓解率可以达到约70%

案例 1– EGFR基因存在敏感突变但易瑞沙无效怎么办？

实例说明

临床背景：女，48岁，非小细胞肺癌，左肺腺癌并左肺门及纵隔淋巴结转移

常规检测：检测*EGFR*和*KRAS*的热点突变，发现*EGFR*存在L858R突变（NM_005228），*KRAS*基因2号外显子12,13号密码子无变异。**预示对EGFR的抑制剂易瑞沙（也称吉非替尼）敏感。**

临床方案：给药易瑞沙口服2个月，病灶无明显变化，提示可能药物无疗效。

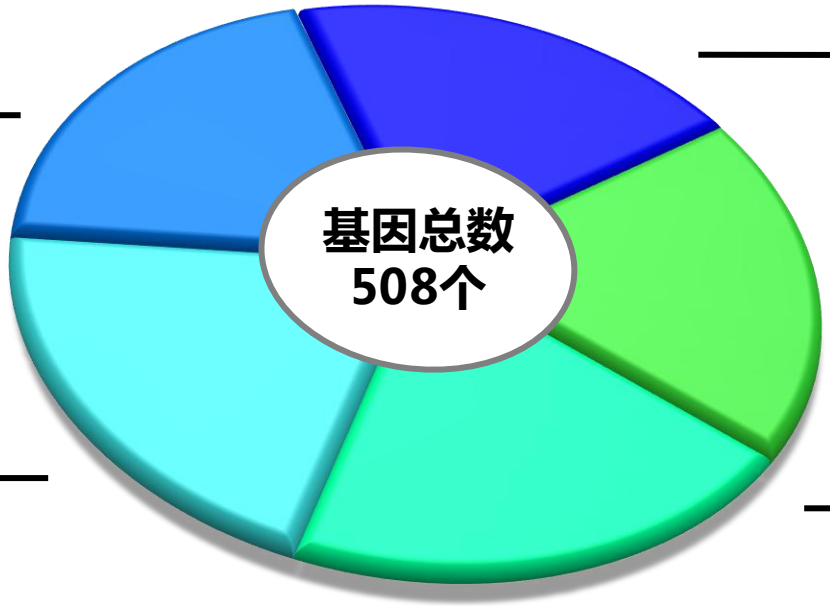
华大检测：对该患者原发癌组织及淋巴转移癌组织样本进行508个肿瘤相关基因测序，**发现患者除存在EGFR的L858R药物敏感突变，ERBB2（Her2）亦存在Ala733Val激酶区突变**(PMID:16397024)。**据研究报道，ERBB激酶区突变会导致患者对可逆的EGFR-TKIs如特罗凯（也称厄罗替尼）、易瑞沙等耐药**。但是对不可逆的EGFR-TKIs，如**Afatinib**敏感(2013年FDA批准的一线治疗药物)。

总结：解释了患者服用易瑞沙无效原因，并为临床新方案制定提供参考。

临床意义：
A.避免盲目用药，耽误治疗，造成经济负担；
B. 给患者更多的选择，找到最佳治疗方案；

检测内容

- **共检测508个基因：**



遗传性肿瘤易感基因 63个

药物靶点基因 119个

信号通路基因 345个

肿瘤高频突变基因 127个

原癌/抑癌基因 129个

- **临床意义：**

1. 个体化用药指导
2. 查找病因

基因组医学发展的挑战

科学发现
大数据积累（建立可靠的分析标准）

我们有太多的未知

基因与疾病，基因与健康的关系

- 自闭症
- 肿瘤
- 糖尿病
- …

科学发现是前提

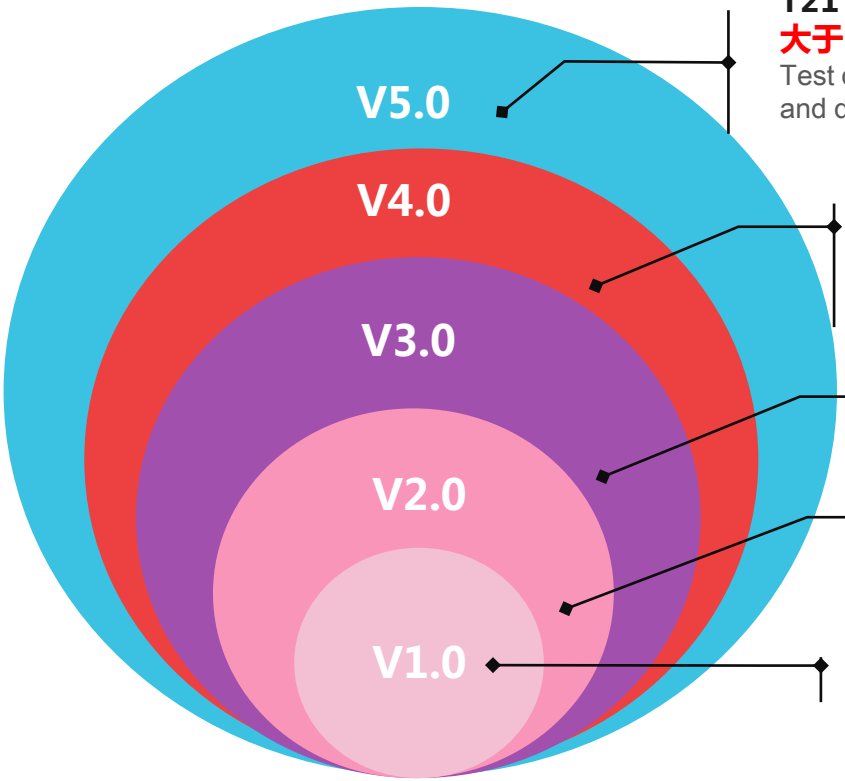
基因组医学的发展需要**积累大数据，不断完善临床应用标准**

- **唐氏**综合征（>20**万人**临床数据）
- **耳**聋（10**万人、百万人**临床数据；多基因，多位点）
- **眼病** （**千人基因**组数据库作为参照）
- **各种**单基因疾病（**千人、百万人基因**组数据库作为参照）

- …

从大数据积累中
建立可靠的临床应用标准

无创产前染色体异常数据库升级模式



V5.15
91214例

2013年12月

T21，T18，T13，性染色体,其它常染色体,大于10M缺失重复的检测数据库；
Test database of T21, T18, T13, Heterosome, other Autosome and deficient-duplication (>10M);

2013年01月

T21，T18，T13，性染色体和其它常染色体的检测数据库；
Test database of T21, T18, T13, Heterosome, and other Autosome；

2012年06月

T21，T18，T13，性染色体检测数据库；
Test database of T21, T18, T13 and Heterosome ；

2012年04月

T21，T18，T13检测数据库；
Test database of T21, T18, and T13；

2011年12月

T21检测数据库；
Test database of T21;

2010年04月

T21 → T18 → T13 → 性染色体 → 其他常染色体 → 大于10M缺失重复 → ……

遗传病数据库

- HGMD
- MORL
- Deafness Variation Database
- LOVD

- **更新功能注释**
- **实现自动化解读**

- 突变频率数据库
- 突变位点数据库

公共数据库

- 定期对公共数据库更新

BGI in-house数据库

积累900万条变异位点信息

临床应用案例积累

- 定期对新发突变数据更新

染色体病数据库

- DGV、
- ISCA、
- HGMD、
- SCAN、
- The Wellcome Trust Sanger Institute

- **更新功能注释**
- **实现自动化解读**

- 致病CNV数据库

公共数据库

- 定期对公共数据库更新

BGI in-house数据库

积累近200个致病性CNV信息

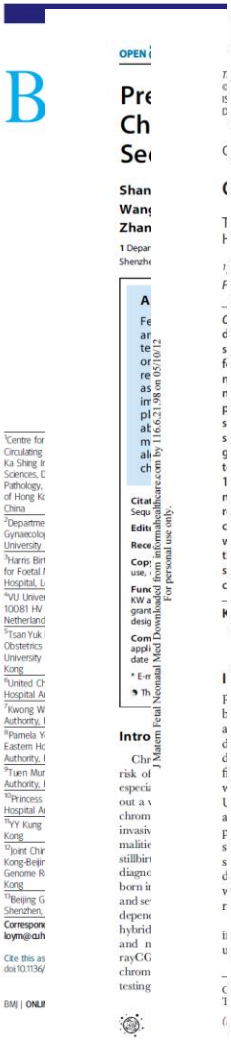
临床应用案例积累

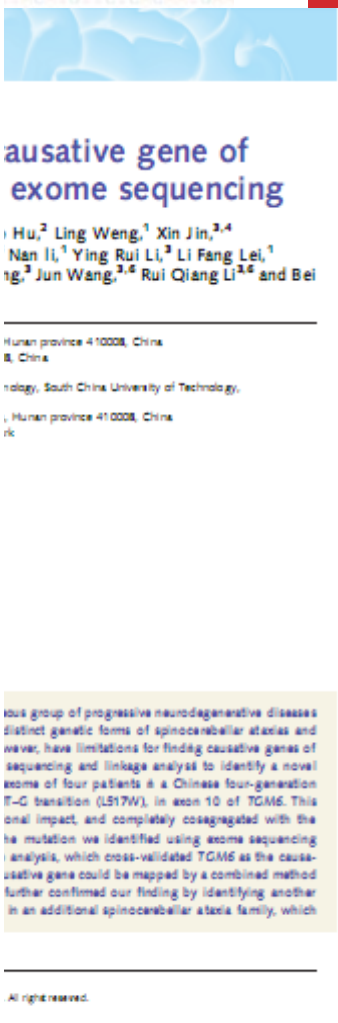
- 定期对新发CNV数据更新

依托国家基因库积累大数据

2011年1月30日，国家发改委批复同意深圳依托华大基因研究院组建国家基因库

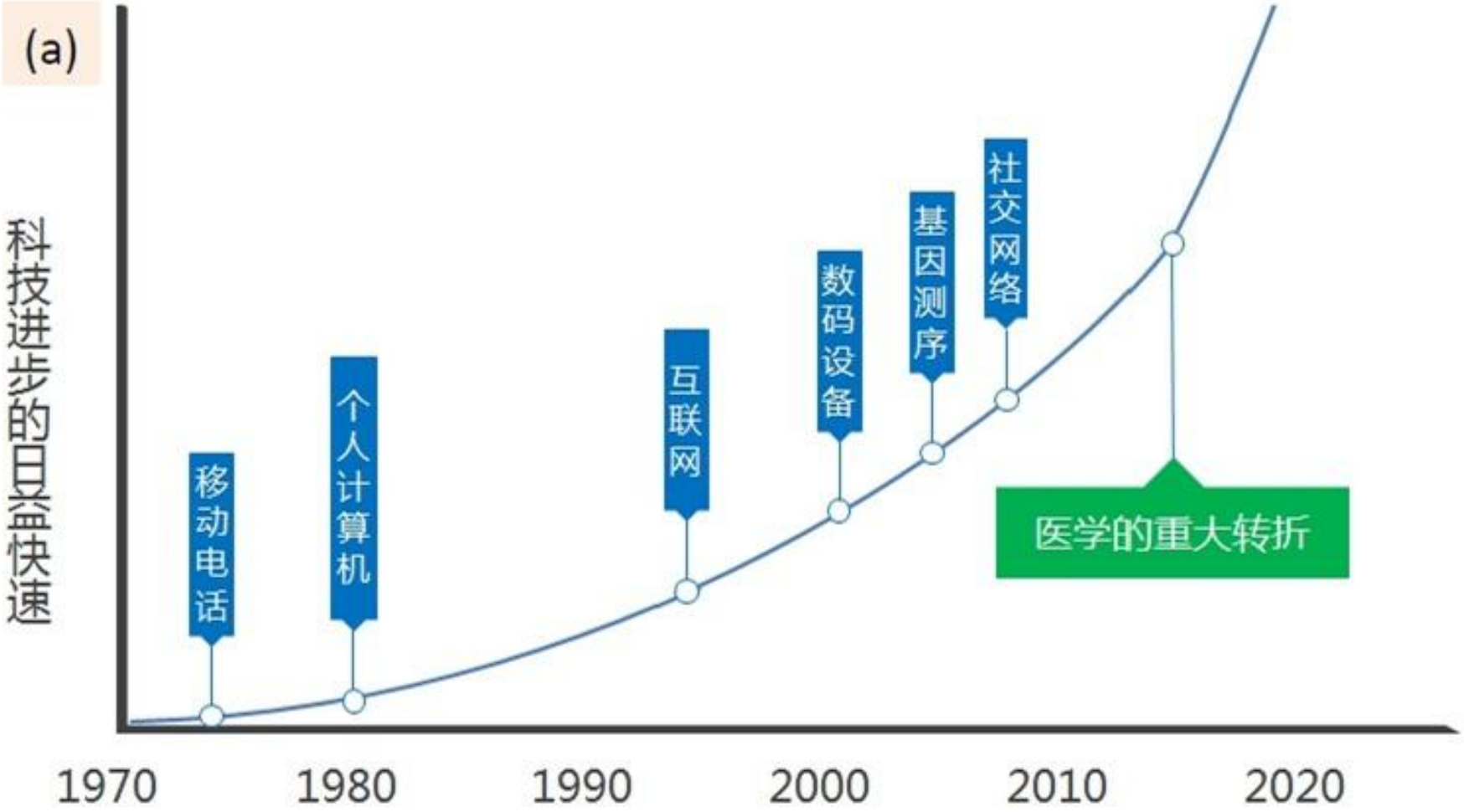
发表医学论文和申请的专利



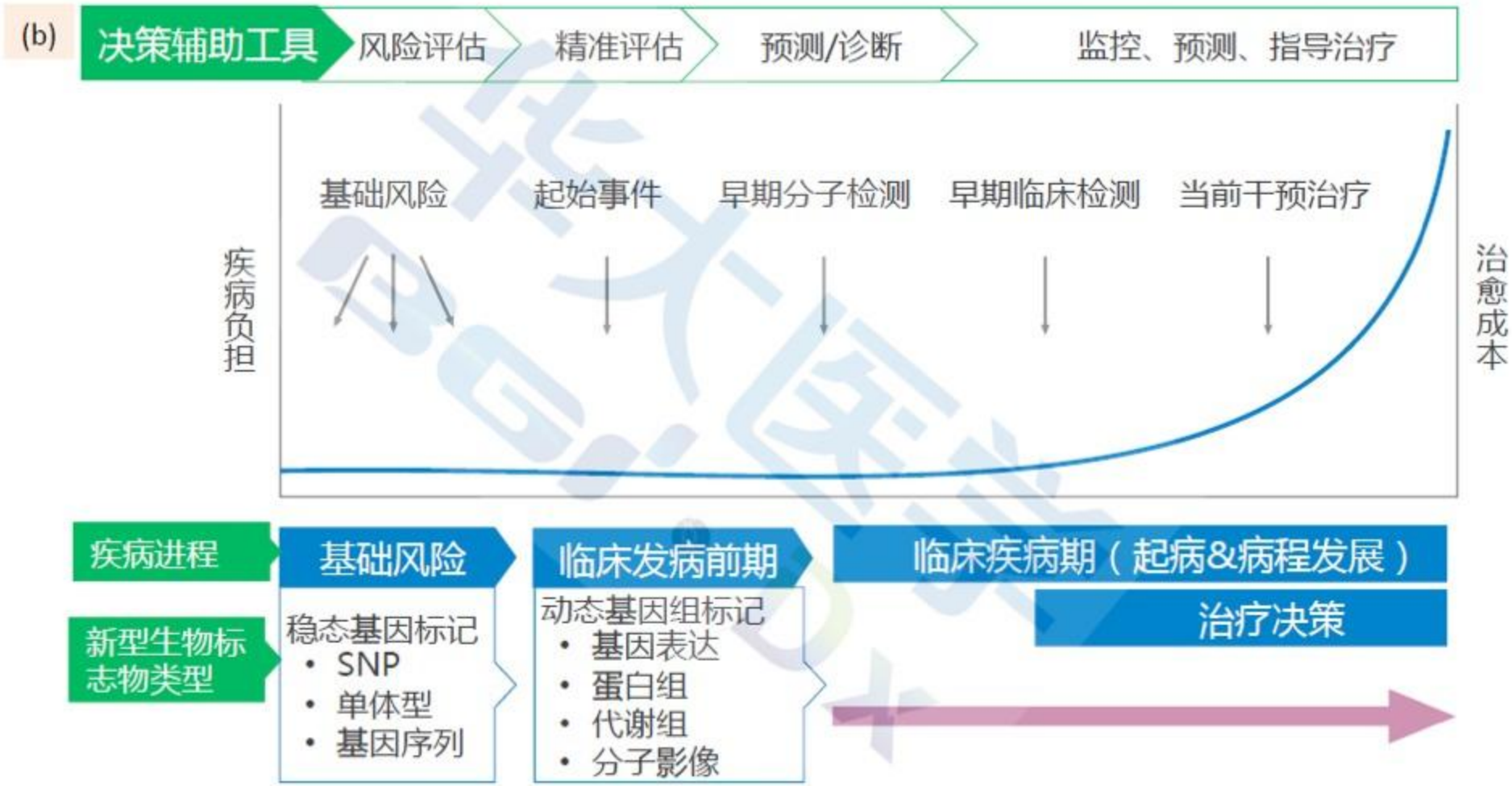


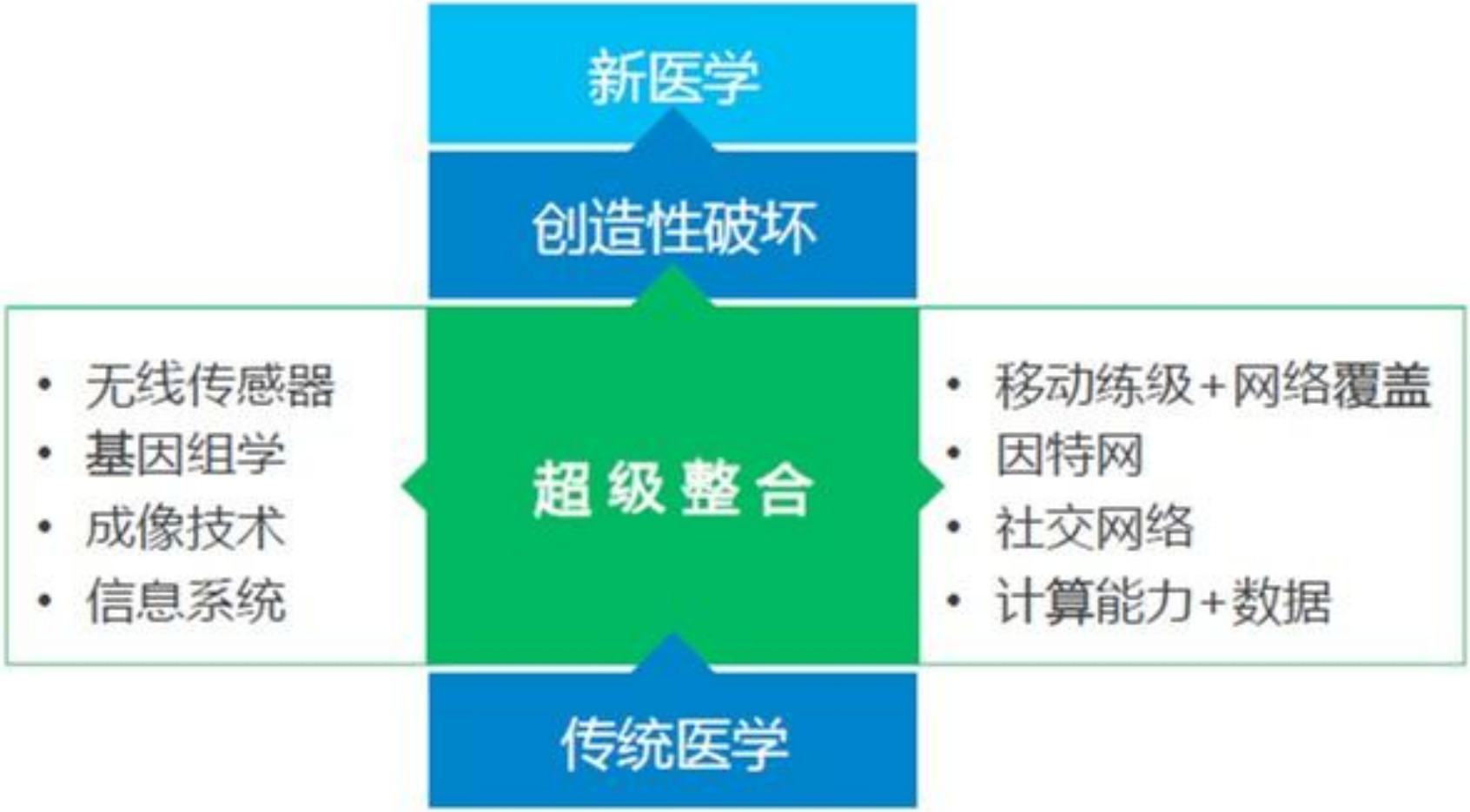
类别	发明名称	申请号	申请日	法律状态
国内申请（5 个）				
发明	一种细胞染色体分析方法	201010252849.X	2010-8-13	已公开（2010-12-29）
发明	胎儿染色体非整倍性的检测方法	201080032859.0	2012-1-20	受理
发明	用于确定人体具有异常状态的系统和方法	201110174686.2	2011-6-24	受理
实用新型	一次性羊膜腔穿刺抽取微量羊水装置	ZL201120196288.6	2011-6-13	授权（2012-5-30）
发明	基于 HiSeq 测序技术检测乙型肝炎病毒分型和耐药基因的方法	201210222621.5	2012-6-29	受理
发明	确定胎儿健康状况的方法	201210271283.4	2012-7-31	受理
PCT 申请（18 个）				
发明	胎儿染色体非整倍性的检测方法	PCT/CN2010/000568	2010-4-23	初审补正（2010. 5. 13）
发明	一种细胞染色体分析方法	PCT/CN2010/001230	2010-8-13	受理
发明	NONINVASIVE DETECTION OF FETAL GENETIC ABNORMALITY（核心算法）	PCT/CN2011/001070	2011-6-29	受理
发明	一种检测单细胞染色体非整倍性的方法	PCT/CN2011/079972	2011-9-21	受理
发明	一种染色体微缺失综合征的检测方法	PCT/CN2011/001805	2011-10-28	受理
发明	一种制备单细胞及微量核酸文库的方法	201110316066.8	2011-10-18	受理
发明	确定基因组是否存在异常的方法及系统	PCT/CN2011/084165	2011-12-17	受理
发明	一种检测病毒在待测样本中整合方式的探针及其制备方法和应用	PCT/CN2011/082855	2011-11-24	受理
发明	检测核酸样本中预定事件的方法和系统以及捕获芯片	PCT/CN2011/084380	2011-12-21	受理
发明	对核酸样本中预定区域进行基因分型的方法和系统	PCT/CN2011/084395	2011-12-21	受理
发明	一种遗传变异的检测方法	PCT/CN2011/002244	2011-12-31	受理
发明	确定样本基因组中是否存在拷贝数变异的方法、系统和计算机可读介质	PCT/CN2012/070680	2012-1-20	受理
发明	一种拷贝数变异检测方法和系统	PCT/CN2012/073545	2012-4-5	受理
发明	确定胎儿基因组中预定区域碱基信息的方法、系统和计算机可读介质（单倍体）	PCT/CN2012/075478	2012-5-14	受理
发明	鉴定双胞胎类型的方法和系统	PCT/CN2012/075969	2012-5-23	受理
发明	用于确定人体具有异常状态的系统和方法	PCT/CN2012/077175	2012-6-19	受理
发明	确定外源基因在人类基因组中整合方式的方法、系统和计算机可读介质	PCT/CN2012/078311	2012-7-6	受理
发明	确定个体是否患有异常状态的方法及系统	PCT/CN2012/080500	2012-8-23	受理
HK 申请（1 个）				
发明	一种细胞染色体分析方法	11105645.5	2011-6-7	受理
国外申请（3 个）				
发明	NONINVASIVE DETECTION OF FETAL GENETIC ABNORMALITY	EP11863253.8	2012-10-16	受理
发明	NONINVASIVE DETECTION OF FETAL GENETIC ABNORMALITY	US13/641,080	2012-10-12	受理
发明	NONINVASIVE DETECTION OF FETAL GENETIC ABNORMALITY		2012-9-20	受理

未来医学的重大转折


(a)
科技进步的日益快速
移动电话
个人计算机
互联网
数码设备
基因测序
社交网络
医学的重大转折
1970
1980
1990
2000
2010
2020

Cell 2014 157, 241-253


(b)
决策辅助工具
风险评估
精准评估
预测/诊断
监控、预测、指导治疗
基础风险
起始事件
早期分子检测
早期临床检测
当前干预治疗
疾病负担
治愈成本
疾病进程
新型生物标志物类型
基础风险
稳态基因标记
• SNP
• 单体型
• 基因序列
临床发病前期
动态基因组标记
• 基因表达
• 蛋白组
• 代谢组
• 分子影像
临床疾病期（起病&病程发展）
治疗决策


新医学
创造性破坏
• 无线传感器
• 基因组学
• 成像技术
• 信息系统
超级整合
• 移动练级+网络覆盖
• 因特网
• 社交网络
• 计算能力+数据
传统医学

GENOMIC MEDICINE

THE FUTURE IS NOW

THANKS