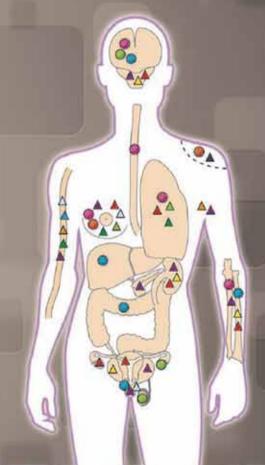


2013年 11月刊 总第60期



肿瘤基因组研究进展概述

CRISPR热潮 雄性哺乳动物为何选择一夫一妻制?



色进 经分

左命世界



# 目录 CONTENTS

# 专题

#### 肿瘤基因组研究进展概述

前言	Ī	. 01
-,	肿瘤基因组测序的诊断价值以及生殖系遗传变异的作用	02 05
$\overline{-}$ '	肿瘤药物基因组学:前途光明但道路曲折的新方向	
	1. 从基础研究发现到临床抗癌治疗工作 ····································	. 15
$\equiv$	肿瘤的表观遗传学重编程问题	. 19
	1. 诱导多潜能现象的启示	. 19
	2. 重编程作用与肿瘤表观遗传学的关系       3. 转录因子的作用	. 21
	3. 转录因于的TF用 ····································	. 23
	5. 染色质调控因子的作用 ····································	23
	6. 被抑制的染色质	. 23
	7. 活化的染色质	24
	8. 重编程与肿瘤干细胞的关系及其它	. 25
四、	肿瘤的异质性	27
	1. 细胞亚克隆间的战争	
	2. 更有意义的问题	
五、	Elaine Mardis,让肿瘤基因组学走上快车道的优秀女车手	. 34
	1. 学习测序专业知识	35
	2. 肿瘤基因组学	
	3. 让肿瘤基因组测序走进临床	. 38

#### 下一期(2013年12月刊)预告:巨噬细胞,人体免疫系统的精灵——巨噬细胞的过去、现在和将来

2013年4月28日,世界上规模最大的一项艾滋病疫苗研究宣布失败,这是人类在抗击艾滋病的战争中的一次滑铁卢。艾滋病毒能够破坏人体免疫系统,这次失败使科学家不得不重新思考对人体免疫系统的认识。巨噬细胞作为人体免疫系统的极其重要的组成部分,其重要性不言而喻。下一期生命奥秘将将从巨噬细胞的历史、研究现状、分类、功能等多个角度对其进行阐述,并展望其未来的发展趋势。

热	点、	CRISPR热潮 ····································	4
百	态	雄性哺乳动物为何选择一夫一妻制?	

封面故事图片为2013年3月29日*Science*杂志期刊封面图。图中显示人类结肠癌(在较低的中心)所示的 X-射线标记带有遗传序列的表达的区域。

本刊文章主要由国外网站文章编译而成,如有版权问题,请版权所有人与本刊联系。 凡本刊所载文章,版权归作者本人和本刊所有,如需转载,请注明作者及出处"生命奥秘"。 本刊提供的任何信息都不能作为医疗凭证和依据,仅供科研参考。



# 肿瘤基因组研究进展概述

# 前言

自2001年人类基因组测序完成起,研究人员们就将目光放到运用这些测序信息来更好地研究遗传学上。最近,越来越多的研究证实了在肿瘤发生、形成及发展过程中存在多种表观遗传学修饰机制。鉴定与遗传性肿瘤相关的单基因变异的前基因组时代(pregenome-era)已逐渐退出舞台,测序技术的进步使研究人员能够利用全基因组方法来检测肿瘤间的差异,以及肿瘤与患者DNA的表观遗传调控。本专题主要讲述这些进步如何在基因组水平上重塑了我们对于肿瘤病理机制的认识。

测序技术上的重大突破与测序成本的大幅下降,表明在不久的将来,常规癌症疗法 的重点将不再是肿瘤起源器官,而是肿瘤基因组特征。阅读深埋在我们体内的复杂的遗 传密码,从而在治疗过程中为每位患者制定最佳的方案,将对医学的发展有着重要的启 示作用。

# 一、肿瘤基因组测序的诊断价值 以及生殖系遗传变异的作用

研究人员希望能够为运用全基因组测序揭示个体患者相关突变的技术制定一套国际标准,让通过肿瘤基因组测序诊断发现的生殖系DNA变异信息更好地应用于临床。

现在,由于技术的进步,在科研机构里要完成从收集某个人的DNA样品到得到全基因组测序结果这一流程,不会超过一个星期。有了全基因组测序结果,我们就可以对一个人的遗传变异有一个全面的了解,不论是大范围的结构变异,比如基因拷贝数变异(copynumber change);还是小范围的突变,比如插入突变、缺失突变、单碱基突变等都逃不出我们的眼睛。随着测序技术的发展,全基因组测序的成本在逐步降低,目前进行一次商业化全基因组测序的价格已经跌破了1万美元(www.genome.gov/sequencing-costs, accessed 11 February 2013)。

在现代医学领域,DNA检测已经成为了一项诊疗常规。比如在人口只有540万的芬兰,去年(2012年)就有大约10万人接受了DNA检测。全基因组测序很快就会取代绝大部分医学上的遗传检测项目,因为从原理上来说,进行全基因组测序可以让我们认识人体基因组里的每一个碱基,而且目前的技术已经可以对90%以上的碱基进行测序。不过,单单依靠技术进步还不能让目前针对特定DNA进行检测的工作方式转换到全基因组测序检测上来。首先,数据存储就是一个大问题,这不仅涉及到存储容量的问题,还会涉及数据安全

和保密的问题。即便是在剔除误差(这也不是 一项简单的工作)、得到了确实可靠的突变数 据之后,如何解读、分析这份测序报告也不是 一件容易的事, 这需要专业的技能, 同时也需 要对各种突变的意义有着非常深刻的认识。 虽然已经有人在从事突变数据库建立这方面 的工作, 但要得到一个可靠的公共突变数据 库,我们还需要一段时间。另外,即便是携带 了易患肿瘤遗传因素的个体, 他们最终患上 癌症的几率也是不同的, 这还取决于外显率 (penetrance, 即某个基因型对个体表型的 影响水平)和遗传率(heritability,即表型受 遗传因素影响的程度)上的差异。进行全基因 组测序检测之后,客户还需要得到专业的遗传 咨询服务。在客户接受全基因组测序检测之 前,还需要提醒他们,可能会得到与检测目的 无关的、预期之外的结果。目前,有专业知 识、有能力承担这些任务的医疗工作者还比较 稀缺,而培养这方面的专业人才也需要一段比 较长的时间。另外, 由于全基因组测序技术还 属于是一项比较新的技术,所以政府在这方面 的规范也比较欠缺,因此,绝大部分医疗机构 都抱着一种"等等看"的态度,无形之中,这 也妨碍了全基因组测序技术在临床上的应用和 推广。

## 1. 发现生殖系突变与肿瘤之间的关系

对于癌症患者,进行全基因组测序检测就 是要发现他们体内肿瘤DNA和生殖系DNA之 间的差异(图1)。对这些数据进行分析就可以发现所谓的体细胞突变(somatic variant,

即只存在于肿瘤细胞当中,在正常细胞里不存在的突变)和生殖系突变(germline variant)。进行这种分析是为了发现抗癌药物的作用靶点,为个性化抗癌治疗提供必要的信息。不过虽然生殖系突变也能够提供很多与健康相关的、有价值的信息,比如与肿瘤易感性(cancer susceptibility)和药物代谢有关的信息,但是我们一直以来都比较忽视它们。

2009年出现了成本更低的外显子组测序技术,即对基因组里所有的编码区进行测序的技术。该技术已经被应用于判断多种肿瘤易感性的研究工作当中。首先采用这种技术开展研究工作的是对急性髓细胞样白血病(acute myeloid leukemia, AML)的研究。有人对一

名AML患者的肿瘤细胞和正常细胞的全外显子组序列进行了测序及比对,结果在肿瘤细胞里发现了8个非同义替换的体细胞单碱基突变,同时还发现了另外两个AML相关突变,它们分别是FLT3基因和NPM1基因的插入突变,不过这两个突变之前已经有人报道过了。最终,这位患者还是在确诊24个月之后不幸去世。不过这一次针对他开展的研究工作至少证明,全基因组测序能够发现致病的相关突变,为靶向治疗提供了指导信息。虽然之后开展的很多肿瘤全基因组测序工作关注的都是肿瘤细胞里的体细胞突变,但还是发现了一些与肿瘤易感性有关的生殖系突变成果,我们在后面会按照先后顺序陆续介绍这方面的信息。



图1 肿瘤患者全基因组测序检测工作流程图。整个过程都要以一个合适的遗传咨询为基础。检测结果应该包括三个部分,分别是在肿瘤细胞里发现的体细胞突变、与肿瘤相关的生殖系突变以及意外发现的其它生殖系突变。

#### 1.1 TP53基因

TP53突变基因携带者大约有90%的几率会患上肿瘤方面的疾病,尤其容易患上李弗劳明综合征(Li-Fraumeni syndrome),而且这些人往往都是在比较年轻的时候发病的。通过全基因组测序可以在已知的肿瘤易患基因(cancer susceptibility gene)当中发现新的突变,比如有一位37岁的患者,她之前患上了AML,在接受抗癌治疗之后又患上了乳腺癌和卵巢癌。对这位患者进行全基因组测序之后就发现,在她的生殖系基因组里,TP53基因中存在一段3kb的缺失。

#### 1.2 ATM基因

通过对两个遗传性胰腺癌家族的全基因组测序和全外显子组测序发现了ATM基因缺失突变。该研究发现,在所有156种之前没有被列入单碱基多态性(single-nucleotide polymorphism,SNP)数据库的失活突变当中,ATM基因缺失突变是与肿瘤关系最为密切的突变。这种突变呈杂合状态,在所有受影响的家族成员中都会分离。该研究发现,ATM基因是一种胰腺癌易患基因(predisposition gene),这也再一次展现出全基因组测序技术在发现家族肿瘤致病基因方面的强大威力。

在冰岛,有科研人员曾经对1795位居民进行过全基因组测序,结果在染色体8q24位置上发现了一个单碱基突变位点rs188140481,该位点与前列腺癌易感性有关。后来在其他几个欧洲人群当中也发现了rs188140481位点与前列腺癌高危患病风险之间存在密切的联系。与全基因组关联研究(GWAS)发现的突变相比,通过全基因组测序能够发现与患病风险有关的、更为罕见、更难以被发现的突变。

#### 1.3 POLE基因和POLD1基因

POLE基因和POLD1基因分别是编码 DNA聚合酶 ε 和 δ 亚单位的基因,通过对多 发腺瘤(multiple adenomas)和结直肠癌(colorectal cancer, CRC)患者进行全基因

组测序检测,发现这两个基因在这些患者体内都发生了生殖系突变。对患有这类疾病的多个家族成员进行研究发现,POLE基因和POLD1基因都是高外显率的结直肠癌易感基因,使用对照关联分析也验证了这一发现。

有人曾经将全基因组测序和外显子组测序技术结合起来,对240位成神经细胞瘤(neuroblastoma)高危个体进行了研究,结果发现了他们的遗传特点。发现其中有10人(在整个群体中占4%)携带了非常罕见的致病生殖系突变,其中包括ALK基因突变、CHEK2基因突变、PINK1基因突变、TP53基因突变、PALB2基因突变和BARD1基因突变。有1人携带了ALK突变基因,3人携带了CHEK2突变基因,1人携带了TP53突变基因,之前就已经发现这3种突变与肿瘤易感性有关。还发现有2人携带了PINK1突变基因,1人携带了PALB2突变基因,2人携带了BARD1突变基因,这3种突变都是新发现的肿瘤易感基因。

这些重点在于发现突变基因的研究工作在 拿到测序结果之后首先都需要和对照人群的 测序结果进行比对,对结果进行一番筛选,这 是一个必需的步骤, 只有这样才能发现特异性 与疾病相关的突变。不过这种筛选工作可能会 "筛掉"一些突变,而这些突变可能会和其它 遗传因素,或者环境因素一起共同发挥致病作 用。比如,随着被研究人群的规模不断增大 (这样就可以对患病人群和对照人群进行更细 致的分类),分析流程的不断优化,整个实验 设计也会更加合理, 更容易发现真正的致病等 位基因。对测序结果进行初筛之后就可以得到 候选的致病突变结果,然后再对这些结果一一 进行严谨的统计学验证, 就能够确定每一个突 变与疾病之间的关系。对于多发性原发肿瘤 (multiple primary cancer) 这种比较明显的 表型,采用全基因组测序更容易发现肿瘤易感 基因。不过对于乳腺癌等比较常见、或者不是 那么明显的表型,采用全基因组测序得到的结 果就不是那么直接,也不是那么容易分析的 了。

要确定高外显率的突变,最好是选择家族 成员进行预测性的遗传检测,然后对已经明确 携带了突变基因的家族成员进行长期跟踪,这 样才能筛查出确定的疾病相关突变。还可以在 家族内对某个特定突变的基因分离情况进行检 测,以确定携带了该突变的高危人群和没有携 带该突变的"安全"人群。

## 2. 全基因组测序发现的生殖系突变: 结直肠癌易患基因

使用全基因组测序发现的生殖系突变结果,能够让临床医生更有针对性地帮助前来咨询的患者了解他们是否更容易患上某些特定的肿瘤。让我们来想象这样一个场景,一位患者被确诊患上了结肠癌,而且肿瘤已经浸润了局部淋巴结。这时候的病情已经危及了患者的生命,可是还是有很大一部分患者通过手术和其它辅助治疗手段能够康复,比如美国国立癌症研究院(NIH)的SEER数据库(http://seer.cancer.gov)的资料显示,IIIA级患者的5年生存率可以达到73%,IIIC级患者的5年生存率

可以达到28%。进行一次全基因组测序的费用和做两次结肠镜检的费用差不多,所以在获得了患者的知情同意之后,完全可以把他们的肿瘤组织和正常组织进行全基因组测序检查,这样就能够发现肿瘤基因组里的体细胞突变,对他们进行个性化的抗癌治疗。而且通过全基因组测序还可以发现生殖系突变信息。根据现有的结直肠癌易感位点资料我们知道,1/3的结直肠癌患者都与遗传因素有关系,大约有5%的结直肠癌患者都表现出了显性或隐性的 孟德尔式遗传特征。

#### 2.1 高外显率突变

如果在比较年轻的时候就发病,或者是身体多处出现原发性肿瘤,那么就要怀疑是不是患上了遗传性肿瘤易感综合症(Hereditary cancer predisposition syndrome),不过这些表现通常都不太特异。由于现代家庭的规模通常都比较小,所以不太容易搜集到家族的肿瘤病史信息,因此即便根据临床表现和家族史,大部分携带了结直肠癌易感基因的个体都没有被及时确诊也就不足为奇了。

最多见的家族性结直肠癌疾病是Lynch综合症(Lynch syndrome),大约有3%的结直肠癌患者患的都是这种疾病,它还有另外一个名字,那就是遗传性非息肉性结直肠癌(hereditary nonpolyposis colorectal cancer,HNPCC)。Lynch综合症是因为DNA错配修复基因*MSH2、MSH6、MLH1*或*PMS2*发生常染色体显性突变而导致的。携带了这些突变基因的个体一生中有超过一半的机会会患上结

直肠癌,他们患上子宫内膜癌(endometrial cancer)和其它腹腔内腺癌(intraabdominal adenocarcinomas)的机率也较高。其它 比较少见的家族性结直肠癌疾病还包括家 族性腺瘤性息肉病 (familial adenomatous polyposis)、Peutz-Jeghers综合症(Peutz-Jeghers syndrome)、幼年性息肉病 (juvenile polyposis)、MYH相关息肉病 (MYH-associated polyposis) 和聚合酶校对 功能相关息肉病(Polymerase proofreadingassociated polyposis) (表1)。对于这些高 外显率的结直肠癌肿瘤目前都有比较成熟的 手术和术后辅助疗法。尤其是对于Lynch综合 症, 更是可以进行早期处置, 通过常规的结肠 镜筛查就可以明显降低突变基因携带者的病死 率和发病率。现在还出现了可以防止Lynch综 合症发病的预防性药物。而且临床遗传检测机 构也已经能够确保为这些突变基因携带者及其 家人提供可靠的遗传咨询和防癌筛查服务。

生殖系遗传检测在某些方面与传统医学检测手段有所不同,因为生殖系遗传检测结果并不仅仅局限于某一个个人,全基因组测序结果可以一次性发现很多问题,也可能需要医疗介入,例如终生跟进寻找患癌诱因等。其它一些问题也会让遗传学检查变得更复杂,比如你的近亲是否有权了解你的检查结果。在一些家族中,如果发现指示病例(index patient,指示

病例指在一起暴发疫情中符合病例定义、最早 发现和报告的病例)携带了高外显率的结直肠 癌易患突变基因,那么就应该对他的直系亲属 进行遗传学咨询,必要的时候还应该对他们进 行遗传学筛查。虽然照道理这些遗传学检测和 防癌检查流程在很多国家里应该都已经建立起 来了,可还是有很多医疗保健系统没有很好地 做到这一点。

#### 表1根据结直肠癌患者的全基因组测序结果进行该肿瘤特异性易患基因的分析评估。

在检测前后,重点针对上面这些结直肠癌相关易患基因,对患者进行解释和遗传咨询并非是一项轻松的工作,如果要进行更广泛的、非特异性的(不针对某一种疾病的)全基因组测序及遗传咨询将更加困难。

高外显率的结肠癌综合症	易患基因	易患基因的功能	
	MLH1基因	DNA错配修复	
Lynch综合征 (Lynch syndrome)	MSH2基因	DNA错配修复	
	MSH6基因	DNA错配修复	
	PMS2基因	DNA错配修复	
家族性腺瘤性息肉病(Familial	4DCT E	Wnt信号通路	
adenomatous polyposis)	APC基因		
Peutz-Jeghers综合征(Peutz-Jeghers	LKB1/STK11基因	调控AMPK家族(adenosine monophosphate-	
syndrome)		activated protein kinase)蛋白活性	
幼年性息肉病(Juvenile polyposis)	SMAD4基因	TGF-β超家族和BMP蛋白信号传导作用	
MYH相关息肉病(MYH-associated	BMPR1A基因	TGF-β信号通路	
polyposis)	MUTYH基因	DNA碱基切除修复	
结直肠癌和家族性先天牙发育不全			
(Colorectal cancer and familial tooth	AXIN2基因	β-catenin/Wnt信号通路	
agenesis)			
聚合酶校对功能相关息肉病(Polymerase	POLD1基因	编码DNA聚合酶δ1催化及校对亚单位	
proofreading-associated polyposis)	POLE基因	编码DNA聚合酶ε催化亚单位	

#### 2.2 低外显率的突变

除了Lynch综合症患者之外,对结直肠癌患者及相关对照人群进行全基因组测序还能够发现一些低外显率的SNP位点(表2)。这些SNP在结直肠癌患者人群中出现的几率要高于在普通人群中出现的几率。在这些SNP位点中,绝大部分SNP的功能现在还都不清

楚,不过科学家认为这些SNP应该可以作为连锁致病突变(linked causative genetic variant)的标志物。比如在染色体8q24位置上就有一个致病突变SNP rs6983267,它可能就与致癌基因MYC的调控作用有着直接的联系,可能会通过影响增强子提高MYC基因的活性。

#### 表2 几个与结直肠癌相关的低外显率突变。

这些突变也可以应用于对结直肠癌患者的全基因组测序结果的分析工作当中。

SNP在染色体上的位置	SNP位点	受影响的基因
6p21	rs1321311*	CDKN1A基因
11q13. 4	rs3824999*	POLD3基因
Xp22. 2	rs5934683*	SHROOM2基因
1q41	rs6691170†	DUSP10基因
3q26. 2	rs10936599 †	MYNN基因
8q23. 3	rs16892766 †	EIF3H基因
8q24. 21	rs6983267†	MYC基因
10p14	rs10795668 †	位于基因间
11q23. 1	rs3802842†	C11orf93基因
10.19.19	11100550 4	DIP2B基因
12q13. 13	rs11169552 †	ATF1基因
14q22. 2	rs4444235†	BMP4基因
15.10.0	4770504#	GREM1基因
15013.3	5q13.3 rs4779584†	
16q22. 1	rs9929218†	CDH1基因
18q21. 1	rs4939827†	SMAD7基因
19q13. 11	rs10411210†	RHPN2基因
20p12. 3	rs961253†	位于基因间
20q13. 33	rs4925386†	LAMA5基因

<sup>\*</sup>数据来自M. G. Dunlop et al., Gut 10.1136/gutjnl-20 11-300537(2012).

现在,又发现了另外3个受到这些SNP影响的基因,它们分别是位于8q23.3上的*EIF3H* 基因、位于14q22.2上的*BMP4*基因和位于18q21.1上的*SMAD7*基因。虽然这些突变基因都比较常见,比如大部分人都会携带一至两个rs6983267高危等位基因,而且这些携带者患上结直肠癌的风险也只会比普通人高一点,因此基本上这些风险基因的作用可以被看作是中性的。通过对结直肠癌患者人群开展的GWAS检测,已经确定了20个结直肠癌高危基因组位点(genomic loci)。这些SNP位点

中每一个位点的作用都比较小,只能够将携带者的患病风险提高一点点,如果用优势比(odds ratio, OR)来表示只有1.07至1.28。对高外显率的结直肠癌易患基因进行的研究也陆续发现了一些新的、风险作用比较低的突变位点,比如APC基因I1307K突变位点和位于MLH1基因启动子区域里的SNP rs1800734等。

不过遗传因素也得综合起来发挥作用,它们彼此之间结合起来可以发挥更大的协同作用。比如携带了19个或者28个结直肠癌相关

<sup>†</sup>数据来自S. J. Lubbe, M. C. Di Bernardo, P. Bro derick, I. Chandler,R. S. Houlston, *Am. J. Epidemiol* . 175, 1 (2012 ).

等位基因的人的患病风险就会提高到2.3。虽然借助全基因组测序数据可以很清楚地发现患者是否携带了风险等位基因,可是以目前的临床水平来看,这些数据的临床意义还不是太大。不过随着对这些风险因素的了解的不断深入,这种情况一定会发生改变。如果要确定准确的患病风险因素,就需要对遗传、环境暴露因素、生活方式等各个方面的信息有一个完整的把握,同时也需要对相关分子机制进行功能方面的检测。

#### 2.3 作用未知的突变

作用未知、意义不明的突变(Variants of uncertain significance, VUS)指的是那些以目前掌握的信息几乎不可能认识其功能的突变,在给患者就高外显率的肿瘤易患基因进行遗传检测时,这些VUS会给我们带来很大的麻烦。曾经有人在开展*BRCA1和BRCA2*突变基因筛查时对这种VUS进行了广泛的研究,结果在30%的检测工作中都会发现至少一个

VUS。也有人正在想办法借助计算机或者临 床病理资料剔除掉这些VUS信息。在很多结 直肠癌高外显率易患基因中也会出现VUS。 与发现致病突变(causative mutations)相 比,结直肠癌易患基因更容易被发现,很多新 发现的截短突变(truncating variant)也很容 易(但不是总是)被认为是致病突变。不过 在很多被看作是家族性结直肠癌易患基因的 基因当中, 如果检测到意义未知的错义突变 (missense-type variant),就会给遗传咨询 工作带来麻烦。由于目前在临床诊断工作中对 这些突变进行功能验证还不太现实, 所以我们 很难确定这些突变的实际意义。在实际工作 中,发现VUS几乎不能给临床医生提供任何 有价值的信息。所以最好在检测工作开始之前 就给患者解释清楚这一点。由于大部分临床医 疗工作者的遗传学背景都比较弱, 所以常常冒 出来的VUS可能也会妨碍遗传筛查工作在临 床上的开展。

# 3. 全基因组生殖系测序结果的临床应用价值

使用全基因组测序结果进行肿瘤分析还存 在另外一个问题,那就是需要提前考虑清楚, 应该向患者介绍哪些测序结果和相关信息。由 于我们对人类基因组功能的认识还不够深刻, 所以在拿着一份目的尚不明确的全基因组测序 结果和患者进行沟通时就会碰到很多问题。虽 然这是需要专业人员才能够完成的工作,但是 我们也应该对非遗传咨询专业的其他医护人员 做好培训,让他们掌握应有的知识,了解整个 遗传咨询工作的流程,按照流程办事,这样才 能使整个遗传咨询工作能够顺利开展,这一点 非常重要。在遗传咨询工作中还有一个问题就 是客户,每一个人的受教育程度,以及对遗传 学的认识程度都是不一样的,他们可能是肿瘤 患者,也可能是其它疾病的患者。最理想的情 况应该是这样的,首先,遗传咨询师与客户沟通,充分了解客户的需求,然后根据客户的需求,然后根据客户的需求,一对一地向他们提供全基因组测序结果咨询服务。对于某些客户,我们只需要告诉他们简单明了、又有价值的突变信息,比如与药物代谢作用相关的遗传突变;或者与几种后果严重,但是完全可以预防的疾病相关的突变信息就可以了。但是还有一些客户,可能就会要求多了解一些检测信息。如果能够像图2展示的那样,有机会先对数据进行反复确认,然后再进行遗传咨询,那就可以更好地、更系统地利用全基因组测序结果。另外,制定临床遗传咨询管理办法,对青少年测序结果的隐私性加以保护也是很有必要的。

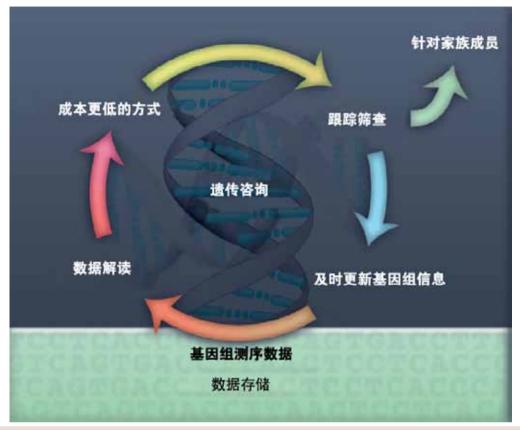


图2 全基因组测序结果长期应用的流程图。如果要充分利用全基因组测序结果,就需要定期更新到最新的基因组研究成果,为解读测序结果做好准备,同时管理模式方面的经验也要积累。

# 4. 未来的发展

进行全基因组测序的最终目的是为了及早发现患病风险,做出相应的干预措施。不过全面认识全基因组测序结果还需要综合考虑,因为表型并不是由基因型这一个因素决定的,表型还要受到环境等其它因素共同作用的影响。综合考虑所有因素才是个性化的风险评估,尤其是在肿瘤形成之前确定一个人的致瘤风险(单基因或多基因遗传风险、对环境因素的低耐受性等)时,更应该意识到这一点。在风险不太明显时,我们很难决定是否应该采取防癌措施,这与在不明确全基因组测序将有怎

样的好处时的心态是一样的。

全基因组测序技术对我们整个人类都是一项伟大的科技进步。可是我们还需要国际社会的共同努力,来推动这项技术在临床上的应用,尤其是如何提高对全基因组测序结果的解读能力,如何设计成本更低、更符合伦理道德规范的遗传咨询操作流程。各国也应该结合各自的实际情况,尽力推广全基因组测序技术,让全基因组测序技术更好地为防病、治病工作服务。

# 特约编辑招聘启事

为了及时收集生命科学最新资讯、提高《生命奥秘》办刊质量,现 面向从事生命科学或对这学科有浓厚兴趣的科研人员、学生诚聘特约编辑 (兼职)。

#### 职位职责:

选题要求内容新颖、评述精辟、注重时效和深入浅出。尤其欢迎以自身系统研究为基础的高水平译述与评论,结合所从事的科研工作提出自己的见解、今后设想或前瞻性展望。

#### 要求:

- 1.具备基因组学、蛋白组学、生物信息学、细胞生物学等生命科学学科背景;
- 2.具备良好的生命科学前沿触觉:
- 3.具备较高的外文文献翻译、编译水平;
- 4.具备较强的选题策划、资料搜集、组织能力,以及专业稿件撰写能力;
- 5.具有高级职称:或者拥有(正在攻读)该领域的最高学位。

有意者请将个人简历发送至 editor@lifeomics.com 联系人: 蔡小姐

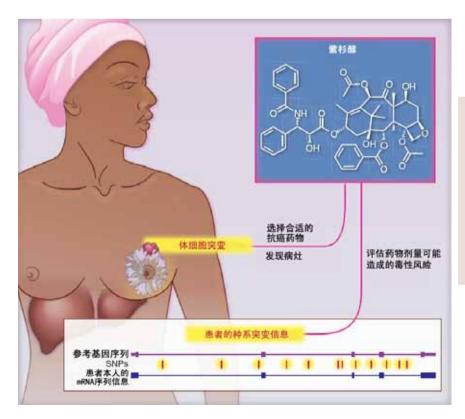


# 二、肿瘤药物基因组学: 前途光明但道路曲折的新方向

本文概述了如何利用肿瘤药物基因组学信息来为每一位患者提供最佳的个性化抗癌治疗服务。

一直以来,对人体肿瘤、以及肿瘤来源的正常细胞的相关调控信号通路的研究和认识都对临床肿瘤诊疗工作起到了非常大的影响和指导作用。因为这方面研究工作的进展,临床抗癌方案也从以前主要使用广谱的、细胞毒性抗癌治疗药物逐渐转换成了更加精确的针对某一条关键信号通路、或者特定基因的化疗方案或者生物治疗方案。因为有了这些进展,所以我们现在对乳腺癌、肺癌以及结直肠癌等常见肿瘤的治疗力度都有了大幅度的提升;对睾丸癌(testicular cancer)以及小儿急性淋巴细胞性白血病(acute lymphoblastic leukemia)等不那么常见的肿瘤也取得了不

错的治疗效果,极大地延长了患者的寿命,甚至能够治愈其中一部分患者。这种主要针对致病信号通路的治疗策略对慢性髓细胞性白血病(chronic myelogenous leukemia)和胃肠道间质瘤(gastrointestinal stromal tumor)等需要终身治疗的慢性疾病(在不复发的前提下)也同样有效,这极大地改善了这部分患者的预后,甚至对糖尿病和高血压病都同样有效。不过毫不夸张地说,这些研究进展可谓来之不易,开发任何一种新疗法或药物,平均每个月的开发成本都在数万美元以上,而且还可能带来毒性作用,给患者的生存质量带来负面的影响。



如图1所示,在肿瘤细胞里发现的体细胞突变以及在患者体内发现的可遗传的种系突变都能够对患者的预后以及抗癌的疗效造成影响。这些突变、也常常被叫做肿瘤生物标志物(cancer biomarker)可以被看作是广义的预后预测型分子标志物(predictive marker),所谓的预后预测型分子标志物既可以预测疾病的病程和发展,也可以预测疾病的预后,而前瞻性预测型分子标志物既可以预测疾病的病后,而前瞻性预测型分子标志物则可以用来预测某种治疗方案可能产生的治疗效果,筛选出最适的治疗人群。遗传信息既能够帮助临床医生们选择疗效最好的治疗方案,也可以指导临床医生们将治疗的副作用尽可能地降到最

低。

科学家们早在20世纪50年代就已经 发现了药物的治疗效果在人群中存在可遗 传的差异,当时认为这种差异主要是由于 每个人对药物的代谢能力(metabolism) 不同而造成的,所以才出现了药物遗传学 (pharmacogenetic)这个研究方向。现在 这个研究方向已经涉及到了与药物相关的各 个领域,比如药物的吸收(absorption)、 在体内的分布(distribution)以及代谢 (excretion)等,还包括了药物的靶标信 息和下游的效应因子(downstream effect mediator)等。表1列出了几个最常见的用基 因型信息来选择抗癌治疗药物的案例。

表1 有助于指导临床选择抗癌化疗药物,或者对患者进行相应支持治疗的药物基因组学DNA分子标志物。

种系突变	体细胞突变	相关的抗癌药物或者支持药物	突变的影响作用
硫代嘌呤甲基转移酶 (Thiopurine methyltransferase)	-	6 - 巯 基 嘌 呤 (Mercaptopurine)、 2-氨基-6-巯基嘌呤 (thioguanine)	有使中性粒细胞减少 的风险
U D P 葡糖醛酰基转移酶 I A 1 ( U D P - glucuronosyltransferase 1A1)	-	伊立替康(Irinotecan)、尼 罗替尼(nilotinib)	有使中性粒细胞减少 的风险、有用药量不 足的风险
6-磷酸葡萄糖脱氢酶 (Glucose-6-phosphate dehydrogenase)	-	拉布立酶 (Rasburicase)	贫血
细胞色素P450 2D6 (Cytochrome P450 2D6)	-	可待因(Codeine)、羟氢可 待酮(oxycodone)、他莫昔 芬(tamoxifen)	影响镇痛剂的疗效、 需要调整用药剂量
-	Janus激酶2(Janus kinase 2,JAK2)	Ruxolitinib	改变药物活性
-	人表皮生长因子受体 1 (Human epiderma growth factor receptor 1, EGFR)	Cetuximab Erlotinib Gefitinib Panitumumab	改变药物活性
-	KRAS (Kirsten rat sarcoma viral oncogene homolog)	Cetuximab Panitumumab	药物活性降低
-	ABL (Abelson murine leukemia viral oncogene homolog 1)	Imatinib, dasatinib, nilotin ib	改变药物活性

(接下表)

KIT (v-kit Zuckerman 4 sarcoma oncogene hor	feline viral Imatinib	改变药物活性
HER2 (Hepidermal factor re2)	growth Lanatinih Trastuzumah	增强药物活性
BRAF (v-Raf sarcoma oncogene h B1)	viral Vemurafenih	增强药物活性
- ALK (Anap lymphoma re tyrosine kin	eceptor Crizotinib	改变药物活性

# 1. 从基础研究发现到临床抗癌治疗工作

通过对肿瘤DNA的分析来帮助临床医生 开展临床抗癌诊疗工作,这种工作模式已经 在临床上存在了20多年。比如以费城染色体 阳性,即9号染色和22号染色体发生了转位 (9:22 chromosomal translocation) 的急 性淋巴细胞性白血病 (acute lymphoblastic leukemia)患者为例,在过去往往会对他们 进行骨髓移植治疗, 而不会选择细胞毒性的 抗癌药物进行治疗。现在, 我们还可以选择 用伊马替尼(imatinib)或者其它ABL酪氨酸 激酶抑制剂(ABL tyrosine kinase inhibitor) 类药物对这类患者进行治疗。对于HER2基因 发生大量扩增的乳腺癌患者来说, 使用赫赛汀 (trastuzumab) 这种抗HER2酪氨酸激酶蛋 白的单抗类抗体药物,或者HER2酪氨酸激酶 抑制剂拉帕替尼(lapatinib)也都能够取得比 较好的疗效。因此,表1中列举的这些遗传信 息已经成为了指导临床医生们选择抗癌药物的 必备工具,随着DNA测序技术的成本的不断 降低,我们将发现更多的体细胞突变位点,在 这些信息的指引下,可以为每一位患者制定出 成本最低、效果最好的最佳抗癌治疗方案。

就某些种类的肿瘤而言,体细胞突变评 估工作已经对临床抗癌工作起到了非常积极

的作用。比如在大约30%的结肠癌患者中都 发现KRAS基因12或13密码子发生了突变, 这些信息提示我们,如果对这些患者使用昂 贵的表皮生长因子受体(epidermal growth factor receptor, EGFR) 抗体类药物进行治 疗,不但不会取得比较理想的治疗效果, 反而还会带来很多的毒性反应。EGFR、 BRAF和JAK2基因分别发生突变的肺癌、黑 色素瘤(melanoma)以及骨髓增生性疾病 (myeloproliferative disorder) 等肿瘤往往对 酪氨酸激酶抑制剂类抗癌药物比较敏感。最 近,美国食品与药品监督管理局(FDA)又 刚刚批准了一种预测抗癌疗效的分子标志物 (molecular predictor),它可以应用于临床 疗效预测工作,这种分子标志物甚至可以预测 出疗效不到10%的抗癌药物的治疗作用。

除此之外,肿瘤细胞可能继续发生突变,产生抗药性,这样就会形成耐药的肿瘤组织。除了最初级的治疗措施之外,目前还缺少个性化的肿瘤医学试验数据来指导患者的临床治疗工作。在选择药物的时候,最常用的方式还是根据整个人群的平均试验数据为基础进行选择,这样就很难确定这种药物对每一个患者个体的具体疗效。所以我们还需要进行大量的

工作,确定、并且验证每一种抗癌药物的具体治疗效果。在使用比较传统的细胞毒性抗癌药物(cytotoxic agent)时更加要注意这一点,因为这类抗癌药物只会对部分癌症患者起效,然而,无论从科学的角度还是经济的角度,都不一定能保证在临床上可以开展基因组学研究,并取得研究成果,进而将基因组学该成果应用于临床。

随着测序技术的不断成熟, 以及测序成 本的不断下降, 这项技术也越来越多的被应用 于肿瘤研究工作当中。虽然在目前的研究工作 里,最常用的方式还是利用测序技术来发现 肿瘤细胞DNA里的体细胞突变,但是肿瘤还 可能与其它因素有关, 比如表观遗传学性状 (epigenetic trait),以及microRNA、RNA 表达水平、甲基化修饰情况以及染色质标志 物(chromatin marks)等。目前,最流行的 遗传学筛查工作是针对某几个相关基因进行 DNA特定捕获,然后对DNA片段进行测序, 了解该基因的突变情况。使用这种测序结果可 以筛选出针对某一条特定信号通路的靶向抗癌 药物,从而进行有针对性的治疗。相反,如果 使用常规的肿瘤组织学检测手段,可能就难以 获得这方面的信息,体现不出这条信号通路的 重要性。

所以现在急需能够发现突变、并且充分认识其临床意义的所谓基因组医学(Genomic medicine)来推动测序技术在临床工作中的应用,证明这些突变的临床价值。有研究人员曾经选择结直肠癌和非小细胞肺癌患者进行过研究,针对145个基因进行了深度测序(deep sequencing),结果分别发现有98%的肿瘤(40个中有39个)和83%的肿瘤(24个中有20个)发生了体细胞突变。有52.5%的结直肠癌肿瘤和72%的非小细胞肺癌肿瘤至少含有1个与某种化疗药物相关的突变。美国国家癌症研究所和国立人类基因组研究所(National Cancer Institute –National Human Genome Research Institute, NCI-NHGRI)联合开展的肿瘤基因组图谱项目(Cancer Genome

Atlas)对取自身体各个解剖部位的各种肿瘤组织进行研究之后也得到了类似的结论。

由于我们对大部分抗癌药物的耐药机制 都不清楚, 所以还需要开展大量的临床药物基 因组学研究工作,利用深度测序技术来发现与 药物敏感性和耐药性相关的作用机制。如有一 些非小细胞肺癌患者最开始使用EGFR酪氨酸 激酶抑制剂类抗癌药物可以取得非常好的治疗 效果, 可是后来就会出现耐药性, 对这部分患 者进行测序研究发现了问题所在。所以目前临 床上已经形成了这样一套常规, 先对EGFR基 因进行测序, 然后再根据测序结果选择是否使 用EGFR酪氨酸激酶抑制剂类抗癌药物。对使 用依维莫司(everolimus)进行治疗的晚期膀 胱癌患者进行全外显子组测序发现, TSC1基 因突变与患者对依维莫司的敏感程度有关。 携带突变TSC1基因的患者在使用依维莫司进 行治疗之后肿瘤复发的平均时间大约为4.1个 月,而TSC1基因没有发生突变的患者在使用 依维莫司进行治疗之后平均1.8个月之后就会 复发。之后又在5.2%(96人中有5人)的晚期 膀胱癌患者中发现了这种TSC1基因的功能缺 失突变, 因此这部分患者可以选择使用依维莫 司进行抗癌治疗, 他们会取得比较好的治疗效 果。

在开展临床活检工作时,能够取得的组织量是有一个限度的。由于不能取得足够多的高质量肿瘤DNA标本,所以现在的工作方式开始从单基因检测逐渐转向了多基因检测,这样就可以利用同样多的标本获取更多的信息。另外,在使用基因检测板(gene panel)或全基因组检测(whole- genome assessment)时,因为不太确定的结果,或者错误的结果所带来的质控问题也会对各个临床实验室得出的分子诊断结果的判读带来一定的影响。另外,我们还需要对25~80%的病例进行预测性的分析,因为对这部分病例的测序工作会发现一些意义不明、但是又与所患肿瘤相关的突变。这种情况在乳腺癌相关的BRCA1/2基因检测工作中表现得尤为突出。所以还需要建立一套实

验室和临床共用的方法,来判断哪些基因中出现的哪些突变具备临床应用价值。

对于某些肿瘤,现在在进行临床试验时选择志愿者的标准已经开始从解剖学分类(即是乳腺癌还是胃癌)逐渐转向了肿瘤的体细胞突变分类(即是EGFR基因突变还是BRCA1基因突变)。现在对能够决定肿瘤侵袭能力的所谓"司机突变(driver mutation)",以及相关治疗反应性的关注也越来越多,这也需要对大量的患者进行测序筛查,发现可靠的治疗靶点,并且进行临床试验验证。有大量的患者都是在没有进行分子诊断、确定遗传背景的情况下就接受了实验性治疗。要在学术科研单位开展这方面的研究难度比较大,因为进行测序发现体细胞突变也需要一定的时间和成本。

除此之外,还会因为一些其它的不曾预料的原因,比如目前对新鲜组织样本的偏好,即采集活检标本进行测序了解体细胞突变信息。很多临床活检操作都没有在介入放射科医生(interventional radiologists)的影像指引下开展,因此不能够对肿瘤组织进行足够安全的活检,而且很多医生也没有受过非常专业的

培训,不一定能够妥善的处理组织标本,最 大限度的呈现出肿瘤的分子特征(大部分社 区肿瘤医生采集完组织标本之后都会将组织 放在福尔马林里固定, 再交给其他机构的病 理医生检查)。如果开发出可以对福尔马林 固定以及石腊包埋处理过的组织,或者对血 浆和血液循环系统里的肿瘤细胞(circulating tumor cell) 进行测序的技术就能够解决这些 问题。如果要开展个性化肿瘤诊疗工作,我们 还需要考虑周全的整体基础方案,确保每一位 从事相关工作的医疗人员都能够接收到良好 的、足够的培训, 能够为患者提供相应的医疗 服务。诸如国家全面癌症覆盖网络(National Comprehensive Cancer Network) 或基础医 学(Foundation Medicine)这样一些专注于 疾病相关基因靶标的学术机构或者商业机构也 都在大力地推动个性化医疗在癌症临床诊疗工 作中的应用。随着测序成本的不断降低,以 及对测序结果解读能力的不断增强, 体细胞 DNA检测工作一定会在肿瘤临床医疗活动中 成为一项常规。

# 2. 种系DNA信息在优化抗癌药物的用药剂量、 判断用药风险方面的作用和价值

虽然在肿瘤基因组学研究工作中最开始 关注的主要是促进肿瘤生长的体细胞突变,但 是后来慢慢发现,患者之间的种系差异也会影 响临床抗癌治疗工作。由于在抗癌治疗工作中 开展的支持治疗的主要目的是减少放化疗的 副作用,所以在如何选择、使用抗癌药物这方 面,种系遗传突变也可能会起到非常重要的影 响作用。

一直以来,我们都忽视了消化道在药物 转运过程中的作用以及肝脏在药物代谢方面的 作用。这些作用都会对药物的使用剂量、使用 方法以及给药间隔时间等诸多方面造成影响。 比如在治疗慢性髓细胞性白血病(chronic myelogenous leukemia)时,如果高度特异性特效药在作用于肿瘤细胞,或者细胞内的分子靶标之前就被代谢,或者清除掉,那么给药量就会不够,效果当然就会大打折扣,同时还可能会促进耐药性产生。这种药物基因组学突变对药物代谢的影响作用已经在某些抗癌药物中有所体现,比如他莫昔芬(tamoxifen)这种用于治疗雌激素受体阳性乳腺癌患者的抗雌激素特效药就是一个典型的例子。人体通过对他莫昔芬的生物转化作用可以将药物转化成endoxifen(这是含量最丰富的一种)等多种

活性物质,细胞色素P450 2D6(CYP2D6)是这个生物转化过程里最重要的一种酶。以往临床上总是将他莫昔芬与抗抑郁药物联用,用来治疗因为抗雌激素治疗导致的潮热(hot flash)等副作用,不过后来研究发现,抗抑郁药物能够抑制细胞色素P450 2D6的酶活性,也就抑制了endoxifen这种活性物质的产生,这样他莫昔芬就完全失去了治疗功效,等于是捡了芝麻丢了西瓜,于是临床上马上停止了将他莫昔芬与抗抑郁药物联用的作法。现在已经有人在开展直接用endoxifen来治疗乳腺癌的临床实验工作。

在美国人当中,大约有7%的人存在 CYP2D6缺失或者多态性问题,从而导致 体内蛋白的丰度与其他人不一样, 或者干 脆就丧失了酶活性。已经有20多篇论文 报道过CYP2D6多态性与乳腺癌用他莫昔 芬治疗效果之间的关系,最近也有好几个 研究指出,完全缺失CYP2D6蛋白,或者 该蛋白的酶活性很低的纯合型突变基因携 带者用他莫昔芬的治疗效果最差。有人把 CYP2D6野生型基因携带者称作"有效代谢者 (extensive metabolizers) ",将CYP2D6 杂合型突变基因携带者称作"中等效率代谢者 (intermediate metabolizers)",有研究人 员对这两种乳腺癌患者进行过研究, 想看看 每一种患者的他莫昔芬最佳用量是多少,结 果发现对于中等效率代谢者需要将用药剂量 增加一倍才能够让血液里endoxifen的含量达 到有效代谢者的水平。这个研究成果也符合 美国FDA推荐的剂量调整方案,在器官功能 异常、药物相互作用、不同年龄等情况下, FDA都会给出推荐的剂量调整方案,这也说 明根据CYP2D6的情况来决定他莫昔芬的用量 应该成为乳腺癌诊疗工作中的常规,以免出现 用药量不足的问题。

在决定临床抗癌诊疗方案、评估某种方 案的风险收益比时,最主要的考量点是控制 肿瘤的效果。缺少客观的、可供评估药物副 作用的数据,是导致这种局面的原因之一。 为了解决这个问题,科研人员又启动了旨在 发现抗癌药物副作用的基因组学研究工作, 这些副作用包括外周感觉神经病变、心血管 毒性、听力损伤以及其它毒性作用。比如微 管抑制剂(microtubule inhibitor)紫杉醇 (paclitaxel)就被广泛用于治疗乳腺癌、肺 癌、卵巢癌等多种肿瘤,但是紫杉醇经常会诱 发神经病变,这种副作用往往会使治疗不得不 终止。可是我们现在完全不能判断患者用药之 后是否会出现这种副作用,所以也无法指导临 床用药选择,无法指导用药方案,无法改善患 者的生存质量。

现在已经有人使用全基因组关联研究(genome-wide association study, GWAS)方法和针对特定基因的检测来帮助预测患者用药之后出现神经病变的风险。比如最近有人选择了855名用紫杉醇治疗过的,尚未发生淋巴结转移的欧洲裔乳腺癌患者进行了一次GWAS研究,结果在FGD4基因(该基因编码FGD1相关F-actin结合蛋白)里发现了一个与发生外周感觉神经病变相关的单核苷酸多态性位点,危害比(hazard ratio)为1.57,95%置信区间为1.30–1.91;在欧洲和非洲裔美国人中也得到了同样的结论。由于FGD4基因是一种外周神经病变基因,所以应该可以用该基因的突变情况来预测使用紫杉醇之后发生外周神经病变副作用的可能性。

还有一项研究发现了另外一个与紫杉醇治疗乳腺癌之后出现神经病变副作用有关的危险因素,那就是CYP2C8\*3的状态,这可能使出现神经病变副作用的机率增加一倍。虽然CYP2C8\*3在非洲裔美国人群中不是那么常见,但是在他们当中还是发现了同样的规律,而且影响的强度都差不多。这种跨种族的现象表明,不论是CYP2C8这种影响药代动力学的突变,还是FGD4这种影响生物转化的突变都可能使患者出现神经病变副作用。

如何避免化疗后遗症也是化疗研究中的 重要课题,这里指的化疗是要杀死所有的肿瘤 细胞,包括没有被检测到的肿瘤细胞。这一点 对于治疗儿童肿瘤尤其重要,因为很多患者都需要常年忍受化疗带来的毒副作用。使用蒽环类抗生素化疗药(anthracycline)导致的心肌病(Cardiomyopathy)就是一种非常严重的后遗症,会对患者的生活质量带来长期的影响。科研人员们通过对特定基因的研究已经发现了与导致心肌病风险相关的种系突变。也得到了与顺铂(cisplatin)导致听力损伤风险相关的药物基因组学种系突变的资料,这为帮助我们预测这些后遗症出现的机率打下了基础。

不过目前将药物基因组学研究成果应用于抗癌治疗还存在比较大的困难,比如很难搜集到足够多的病例,来发现有价值的突变,并且对其进行验证。组织、开展这样一个临床试验一般都需要7至10年的时间才能完成,才能发现一个有价值的突变,后续的验证工作也同样困难,所以目前在文献中有报道的成果还非

常少。我们还缺少与抗癌药物疗效有关的遗传 学资料, 因此很难证明药物基因组学发现与疗 效的关系。不过使用大规模、多代家族成员的 细胞系可以解决这个问题, 研究发现, 29种 常用化疗药物的细胞毒性作用的遗传力差异非 常大,跨度在10%至70%之间,其中66%的 细胞毒性遗传力超过了30%。这一发现既有 助于优化药物评估方案, 也可以指导组织研 究,将之前搜集的临床组织标本用于验证研究 工作。条形码技术(bar coding)和机器人技 术 (robotics) 的应用可以使细胞系表型研究 的处理能力大大增加, 一个项目就可以处理 500至1000个细胞系,同时也可以进行组织全 基因组学分析。如图2所示,创新的药物基因 组学研究策略可以使我们更快发现与合理用药 有关的基因组学信息。

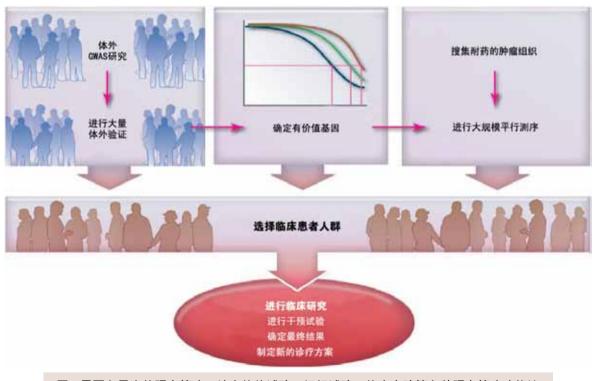


图2 需要多层次的研究策略。综合体外试验、组织试验、体内实验等多种研究策略才能让药物基因组学研究成果更快地走进临床。

# 3. 让药物基因组学研究成果更快地走进临床

对药物基因组学研究成果的重复和验证工作也是一大工作难点。因为往往很难对大量的肿瘤患者进行统一的治疗,也很难对他们进行全面的评估,客观地评价治疗效果。操作规程应该是先采集所有参加临床试验的患者的基因组数据,然后在取得知情同意的前提下进行药物基因组学研究。目前制药公司开展的绝大部分大规模临床试验都是按照这套流程进行的,这在一些美国国立癌症研究所的临床试验工作中也已经成为了一项常规,不过这套流程在学术科研单位或者某些基金资助的科研工作中还没有系统地建立起来。

一方面是想让最新的研究成果尽快应 用于临床,另一方面是需要足够的数据来尽 量确保这些发现的真正价值, 在这两种倾向 之间如何保持平衡是一个很大的难题。认为 前瞻性的随机对照试验是唯一的解决方案, 这给我们带来了很大的麻烦, 因为这会让临 床与科研之间出现一个5至10年的滞后。从 资金、临床试验基础设施和快速接受最新的 研究策略的角度来看,科研资助机构也不能 满足对这类研究工作的优化需要。有几项工 作正在努力尽快地接受药物基因组学研究 数据,这是建立在与遗传学和药物治疗学有 关的各研究单位的意见达成一致的基础之 上的。其中一项工作就是建立临床药物遗 传学促进协会(Clinical Pharmacogenetics Implementation Consortium, CPIC), 该组 织有来自4大洲的80多家研究机构参与。对于 CPIC这样的组织,最关键的一点就是要意识 到在医疗活动中要考虑很多方面, 比如用药剂 量等,要有足够的证据证明这些决定是对患者 有利的,即使是在大家都在等待一个最佳的研 究、从更高的层面来指导临床治疗的时候。

不过在制定临床诊疗方案时还是存在双 重标准,比如承认药物与突变之间的相互作 用,也能够快速地接受相关的作用理论,但是由于缺乏大量的前瞩性研究数据,所以不能将这些研究成果尽快地应用到实际工作当中。CYP2D6和他莫昔芬、CYP3A4和紫杉烷(taxane)等都是这方面的案例。药物与突变相互作用的标准会受到熟悉程度的影响,这往往需要好几年的时间,而且还需要确保患者服药时间与其它药物等变量的控制,不过这种相互作用的预测能力比同一条生理通路里的基因缺失证据还要小。所以需要开发一套新的工作流程,在已经被临床验证过的生理病理通路中发现的任何突变都应该尽快地应用于临床。

药物基因组学研究的目的就是要对传统的生物标志物进行补充、解释意外发生的事件、剔除疗效差的药物、确定用药剂量、预测药物副作用等。这些作用都非常重要,在调查工作中一定不要忽视。此外还有另外一方面的作用,那就是促进新型医疗模式的转变,比如降低30天内的再次入院率、改善集束化护理(bundled care)的效益、优化社区医疗服务等,不过这都属于远期目标,还没有被纳入NIH的资助范围。通过对患者的观察(observational cohort)和对电子病历的研究也可以实现这些目标,这也可能会促进药物基因组学走进临床。

现在是时候往前更进一步了。虽然科研人员们发现了一些药物基因组学表型,但是这方面研究成果的转化效率还是非常低。原因有很多,比如缺乏足够的数据来证明其可靠性,干预式药物遗传学研究还开展得比较少,如何建立一套清晰、规范的临床抗癌诊疗常规等其它实际操作方面的问题都是阻碍药物基因组学研究成果尽快向临床转化的原因。另外还存在社会方面的因素,比如对遗传检查的接受程度、保险的报销问题等。要使个性化医疗更快地走向临床,这些问题都需要得到解决。

# 三、肿瘤的表观遗传学重编程问题

本文介绍了表观遗传调控控制细胞命运的机制。对这些作用机制的研究将有助于我们更好地认识肿瘤的转化、异质性和肿瘤干细胞模型等方面的问题。

在发育过程中,细胞最终的分化方向该 如何决定是一个动态的过程, 在多种时空因 素的精确控制下,细胞最终会表现出各种不 同的表型(phenotype)。从一个全能细胞 (totipotent cell) 开始,经过一系列的自我增 殖(self-renewal)和分化(differentiation) 作用,这些全能细胞最终会演变成各种各样的 细胞、组织和器官, 所有这些组成在一起才 能够形成一个有血有肉的生命体。DNA序列 特异性的转录因子(DNA sequence-specific transcription factor) 在决定细胞最终分化命 运的过程中起到了非常重要的作用,这一点在 MyoD(该转录因子能够决定肌细胞的分化) 以及和多能性 (pluripotency) 相关的几个核 心转录因子(core TF)的多个相关研究中都 已经得到了证实。经典理论认为, 转录因子主 要与启动子(promoter)附近区域发生相互 作用,继而启动基因的转录作用。之后的研究 又在人类基因组中陆续发现了多个远端的增强 子样元件(distal enhancer-like element), 这些元件也能够被转录因子激活, 而且是以一

种细胞种类特异性的方式(cell type –specific pattern)被激活的。

转录因子发挥上述这些激活作用就必须 要先解决染色质的(折叠)结构问题,我们都 知道染色质是由DNA、RNA、组蛋白和多种 调控蛋白结合在一起形成的紧密的高级结构。 首先, 转录因子要招募染色质调控因子, 这些 染色质调控因子能够帮助转录因子接近DNA 靶标序列,同时也能够识别染色质上的组蛋白 修饰信号,了解染色质当时所处的状态信息和 功能信息。不过转录因子是否能够起到激活作 用首先取决于染色质的可接近性, 所以染色质 调控因子和染色质的状态都是决定转录因子的 转录活性的因素之一。细胞分化主要依赖启动 子和增强子,不过这些启动子和增强子总是被 各种转录因子和染色质修饰物占据着, 所以由 一系列转录因子、相关的染色质调控因子,以 及染色质状态构成的层级鲜明的调控网络才能 确保细胞分化过程有条不紊、井然有序地开展 下去(图1)。

# 1.诱导多潜能现象的启示

2006年,日本科学家Shinya Yamanaka 发现了诱导多潜能现象,即一个已经分化的细胞可以在几个转录因子的作用下直接被重编程(reprogrammed)为一个具有多向分化潜能的诱导多潜能干细胞(induced pluripotent stem cell, iPS cell)的现象,他因此获得了诺

贝尔奖。Yamanaka的研究成果对于整个干细胞研究领域和再生医学研究领域都是一项领先的、了不起的发现。这项研究工作证明了,如果细胞分化、发育过程中的基因调控机制出现了异常,将会带来多么大的影响。这也让我们对转录因子、染色质调控因子和染色质的状态

等直接决定细胞最终状态的因素的作用机制又有了更加深刻的认识。现在,我们已经发现大量的因子和作用机制,都与细胞的恶性转化(malignant transformation)有着密切的关

系。接下来,我们将就这些与细胞重编程和癌 变过程有关的遗传和表观遗传学因子,及其作 用机制进行一番介绍和讨论。

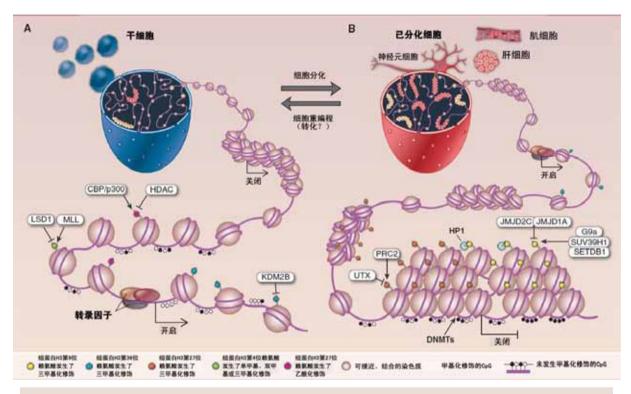


图1 细胞的特异性分化过程取决于细胞染色质结构的整体改变。(A)在多潜能细胞里,染色质是高度动态变化的,而且整个染色质都是容易被各种转录因子接近、结合的。(B)在细胞分化时,处于未活化状态的基因组区域会被各种抑制性组蛋白修饰物"掩盖"起来,形成被抑制的染色质(repressive chromatin),不表现出调控活性。DNA甲基化修饰、组蛋白修饰和各种染色质调控因子(这些调控因子的表达水平会随着分化过程的进展不停地做出相应的调整)都能够对染色质的整体结构起到调控作用。另外,某些基因位点发生局部的染色质结构改变也会影响转录活性。

科学家们最初发现的诱导多潜能现象是通过在成纤维细胞(fibroblast)里表达了Oct-4、Sox2、Klf4和c-Myc这四种转录因子之后发现的。这其中又只有Oct-4和Sox2这两种核心转录因子是必需的,另外两种转录因子只是起到了增强重编程效率的作用,可以被其它的转录因子,比如Nanog和Lin28等替换。这种细胞能够在不同类别之间直接转换的现象更进一步证明了转录因子对细胞状态的决定性作用。正确的转录因子组合可以让各转录因子互相发挥协同作用,与启动子和增强子结合,

通过激活基因网络的的转录和表达,确保细胞的状态发生转变。在重编程过程中还包括染色质局部和总体的结构发生改变,同时表观遗传学修饰作用也会发生相应的改变。在iPS的重编程过程中,转录因子通过招募染色质调控因子使染色质重新激活,启动基因转录这一系列反应都是在重编程过程的最初阶段发生的。相比之下,二价结构域(bivalent domain)的形成,以及多向分化细胞标志性的染色质整体解压缩过程则都发生在比较靠后的阶段。其中还包括染色质修饰(chromatin

modification)和重构(remodeling)的过程,这些都离不开染色质调控因子的催化活性作用,这说明染色质调控因子对于重编程反应至关重要。另外,染色质之前所处的状态以

及DNA甲基化修饰的情况也都会起到阻碍作用,影响转录因子与染色质的结合,抑制基因表达,干扰细胞状态发生改变的过程。

# 2. 重编程作用与肿瘤表观遗传学的关系

在恶性转化过程中通常会包括重新启动的细胞发育进程,该过程和细胞的重编程过程有点类似,也会让肿瘤细胞像iPS细胞或者其它直接重编程干细胞(directly reprogrammed stem cell)那样,重新获得无限的自我增殖能力。之所以说细胞恶性转化过程和重编程过程是类似的,是因为在这两种作用中都有一套共同的促进因子和抑制因子。比如有好几种重编程转录因子同时是癌基因(oncogene)的转录因子;好几种能够抑

制重编程作用的转录因子也都是有名的抑癌因子,其中就包括p53和Ink4A/Arf等,这些转录因子在细胞增殖和凋亡过程中既能够抑制重编程机制,同时也能够抑制细胞恶性转化机制。染色质调控因子也具有同样的特点,也是在重编程过程中起到了效应因子和调控因子的作用,也一样对癌基因具有影响作用。这些发现都表明,表观遗传学对于细胞重编程的重要意义可能也同样体现在细胞恶性转化的过程中。

# 3. 转录因子的作用

如图2所示,每一种重编程因子在癌基因的作用中都占有一席之地。Oct-4能够促进生殖系细胞肿瘤的形成,在临床诊断工作中它也是一种非常重要的分子标志物。在肺鳞状细胞癌(squamous cell carcinoma)、食管鳞状细胞癌以及小细胞肺癌(small-cell lungcarcinoma)中,Sox2也都会大量表达,起到一个谱系存活癌基因(lineage-survival oncogene)的作用。与此同时,对于恶性胶质瘤(glioblastoma)、乳腺癌和尤因肉瘤(Ewing sarcoma)等肿瘤的肿瘤干细胞而言,Sox2还是一个必须的启动因子(图2),这一点与Sox2对于多向分化潜能干细胞和组织干细胞的意义一模一样。

c-Myc可见于多种人类恶性肿瘤。c-Myc的表达可能也是为什么iPS克隆小鼠会自发形成肿瘤的原因之一。c-Myc能够诱导胚胎干细

胞(embryonic stem cell, ES cell)及多种恶性肿瘤相关基因的表达,从而促进细胞的增殖。另外,c-Myc还可以通过对RNA聚合酶II延伸因子——P-TEFb的作用,提高细胞内基因表达的整体水平。这种作用机制可能就是为什么仅一种癌基因就能够增强多向分化潜能网络的活性,同时也能够启动多种恶性转化进程的原因。

KIf4在癌变过程中则起到了双重作用,它能够促进乳腺癌、皮肤癌的形成,同时对于胃癌、结直肠癌和膀胱癌却又能够起到抑制作用。KIf4之所以能够起到促癌作用,这可能与它具有的核重编程作用有关。KIf4除了具有各种重编程作用之外,还具有p53抑制作用,所以它能够极大地提升细胞的自我增殖能力。在胚胎干细胞和肿瘤细胞中,KIf4还具有端粒酶活性诱导功能,这是因为KIf4能够促进

β-catenin蛋白与*TERT*基因的启动子区域结合。Klf4的抑癌功能则可能与它对促凋亡基因 CKDN1A的诱导作用有关。

Nanog在多种恶变细胞中都有表达,该转录因子在肿瘤干细胞中可能起到了非常关键的作用。在肝癌细胞中,Nanog是CD23分子阳性的干细胞样肝癌细胞(CD24-positive stem like cell)的必需因子。在该细胞中,Nanog的表达主要由STAT3信号通路来维持。在恶性胶质瘤细胞里,Nanog能够诱导Gli1因子表达,这也提高了CD133分子阳性的肿瘤干细胞的克隆能力(clonogenic potential)和致瘤能力(tumorigenic potential)。在结肠癌、前列腺癌和尤因肉瘤(Ewing sarcoma)等恶性肿瘤的干细胞中也都有Nanog转录因子的表达(图2A)。

Lin28也在干细胞生物学中具有非常重要的地位,它能够加快iPS重编程的速度,提高胚胎细胞的增殖能力。Lin28是一种RNA结合蛋白,该因子因在多种人类肿瘤,尤其是在未分化肿瘤和晚期肿瘤中都有表达而闻名。Lin28能够抑制Let-7 microRNA家族成员分子

的成熟,所以对包括c-Myc、K-Ras、Sall4和 Hmga2等多种癌基因在内的基因都起到了去 抑制的作用。最近的科学研究又发现了好几项 Lin28因子的新作用,其中就包括了对Oct-4因 子转位(translation)的直接作用。Lin28因 子可能就是依靠这些作用机制,才能够在重编程过程中对c-Myc因子起到辅助作用的,也可能是因为这些作用机制所以才具有促癌作用。

除了前面介绍的这些多向分化潜能转录因子之外,Olig2、Foxg1等其它能够促进细胞类别发生直接转变(direct lineage conversion)的转录因子也都与人类的肿瘤相关。这些研究成果都表明,重编程过程与恶性转化过程之间一定存在某种关联。不过值得注意的是,这两者之间的关系远不止于此,在这两条途径中,任何一个因子的作用都不一定是确定的,同一个因子在不同的情况下可能会具有不同的作用。另外,我们不能武断地认为很多与癌变相关的转录因子就一定具有重编程功能,最近在肝癌细胞和多种实体瘤细胞里发现的转化融合蛋白(transforming fusion protein)就是一个例子。



图2 与iPS细胞核重编程及细胞癌变过程都有关的基因简介。图A和图B分别展示的是与iPS细胞核重编程及细胞癌变过程都有关的转录因子和染色质调控因子。这些因子都与癌基因或抑癌基因有关,也与通过其它机制参与细胞癌变过程的基因有关。

# 4. DNA甲基化修饰的作用

DNA甲基化修饰(DNA methylation)是一种相对稳定的表观遗传学修饰作用(epigenetic modification),它能够促使某些基因的启动子或重复元件(repetitive element)沉默。在细胞的重编程过程中,甲基化修饰模式(Methylation pattern)必须被重置,而这个过程正是细胞重编程操作里的一大难点。对于肿瘤细胞而言,CpG岛(CpG island)的超甲基化修饰(hypermethylation)是一个非常明显的特征性标志。DNA甲基转移酶(DNA methyltransferase)Dnmt3a遗传失活也是白

血病和淋巴瘤这两种恶性肿瘤的一大特征。 虽然其中的致病机制还不是十分清楚,但是 DNA甲基化修饰一定对细胞转化起到了抑制 作用,因为DNA甲基化修饰可以抑制表观遗 传学重编程基因的活化,同样也能够抑制细胞 分化基因或者细胞凋亡基因的活化,从而促进 细胞癌变。因此,细胞的表观遗传学特征对于 不同的恶性肿瘤而言,可能会表现出完全不同 的作用。实际上,我们在科研工作中也的确观 察到了很多染色质调控因子在不同的肿瘤中表 现出了看似"自相矛盾"的作用,接下来将对 此进行介绍。

### 5.染色质调控因子的作用

如图2B所示,染色质调控因子也与细胞重编程和癌变作用有关。这些染色质调控因子可以通过几个途径促进细胞恶变。很多染色质调控因子都和转录因子一样,也与癌变或者抑癌作用有关,而且它们的作用可以通过功能获得突变或者功能缺失突变,以及转位等突变方式受到影响。在没有发生突变的情况下,染色质调控因子也可以借助融合蛋白,或者其它致癌因子(oncogenic factor)对基因的表达发挥调控作用。染色质调控因子还可以对

其它肿瘤相关生理进程施加影响,比如上皮细胞向间充质细胞的转变过程(epithelial-to-mesenchymal transition, EMT)、细胞老化过程(senescence)、基因组的稳定性以及肿瘤细胞的转移(metastasis)过程等。这些不同的作用,以及染色质调控因子对多种肿瘤表现出积极促进作用的事实都给我们出了一个大难题,因为这会让我们难以预测某种染色质调控因子突变、染色质状态的改变或者表观遗传学治疗措施对肿瘤会带来什么样的影响。

# 6. 被抑制的染色质

在已分化细胞中,基因组里未活化的部分都存在于各个被抑制的染色质(repressive chromatin)里。这些完全不具备调控活性的染色质结构紧密(图1),它们对基因组里大部分的区域都有影响。典型的被抑制染色质结构包括有特征性标志H3K9me3(组蛋白H3第9位赖氨酸发生了三甲基化修饰)的异染

色质(heterochromatin)、富含H3K27me3(组蛋白H3第27位赖氨酸发生了三甲基化修饰)的多梳抑制区域(Polycomb-repressed regions)以及富含核纤层蛋白(nuclear lamina)和H3K9me2(组蛋白H3第9位赖氨酸发生了二甲基化修饰)的结构高度紧密的区域。

H3K9位点的甲基化修饰与相应的染色质调控因子对于染色质结构的稳定性、EMT和细胞老化等进程都至关重要。存在于某些已分化细胞中的H3K9me3结构域似乎能够阻止重编程转录因子之间的结合;在iPS过程中,这种甲基化修饰作用似乎也没有得到充分的重置。因此,抑制Suv39h1、Setdb1和G9a等能够促进H3K9位点发生甲基化修饰的染色质调控因子应该就可以提高细胞重编程的效率。H3K27me3和相关的多梳蛋白抑制子(Polycomb repressor)与组织特异性的基因表达调控作用有关。Ezh2和Utx分别能够催化H3K27me3位点的甲基化修饰作用和去甲基化修饰作用,这两种因子也都是确保iPS细胞重编程过程顺利进行的必需因子。

Suv39h1因子是癌基因诱导的细胞老化机制里的必需因子,在小鼠动物研究中发现,缺失了Suv39h1因子之后会使细胞的活力降低、基因组失稳、小鼠患肿瘤的风险增高。在存在异常基因靶标的情况下,通过融合蛋白回复Suv39h1因子的表达能够促使小鼠患上急性髓细胞样白血病(acute myeloid leukemia)。如图2B所示,Setdb1和G9a也与多种恶性肿瘤有关,其中G9a还与肿瘤转移和EMT进程有关。多梳蛋白调控因子在多种恶性肿瘤中也是通过突变或过表达发挥作用的因子之一。尤其是Ezh2和Bmi1都可以通过表观遗传学机制介导CDKN2A基因沉默,这两种因子也都是肿瘤干细胞的必需因子。

# 7. 活化的染色质

基因和调控元件的活化都与染色质的状 态有关。增强子和启动子最明显的标志就是 H3K4位点的甲基化修饰程度以及组蛋白的 乙酰化修饰情况(acetylation)(图1)。 甲基化修饰作用是由含有混合细胞性白血病 (mixed-lineage leukemia, MLL) 同系物 (果蝇trithorax蛋白的同系物)和Wdr5等辅 助因子共同组成的复合体催化完成的。这些染 色质调控因子在细胞重编程过程和癌变过程中 也都发挥了重要的作用。Wdr5蛋白能够与多 能转录因子Oct-4直接相互作用,是产生iPS 细胞的必需因子之一。虽然MLL融合蛋白中通 常缺少催化结构域,但是这种蛋白却是一种非 常强的促白血病因子,它可以招募其它辅助因 子来发挥作用。如图2B所示,科学家们已经 在多种肿瘤细胞里发现了失活的MLL突变蛋白 同系物。H3K4去甲基化酶Lsd1也具有促癌作 用,如图2B所示,该因子在不同的环境下既 可能起到辅助激活因子的作用,也可以起到辅 助抑制因子的作用。

基因转录从起始转换到延伸过程的这个步骤是非常关键的一个转录调控步骤。 H3K79甲基转移酶Dot1l能够促进这种转换过程进行,从而抑制iPS细胞的重编程作用。 Dot1l蛋白也是促进白血病发生的必需因子之一。MLL融合蛋白能够通过招募Dot1l蛋白起到介导异常调控机制的作用。因此,用Dot1l蛋白小分子抑制剂治疗由MLL介导的白血病会取得非常好的疗效。多发性骨髓瘤融合蛋白Mmset和Kdm2b分别能够催化H3K36位点的甲基化修饰反应和去甲基化反应,这也是转录延伸阶段的特征性标志之一。Kdm2b能够促进iPS细胞生成,促使细胞永生化,也有助于白血病干细胞的形成和维持。

还有很多其它染色质调控因子也都与染色质活化或染色质重构(chromatin remodeling)作用有关,这些染色质调控因子还与细胞的转化作用相关。CREB结合蛋白CBP和p300组蛋白乙酰转移酶组成的复合体是某些转录因子的辅助激活因子,可以激活

与细胞增殖、凋亡、分化以及DNA稳定性等 作用相关的表观遗传学程序。这种CBP/p300 复合体还可以对p53、Rb以及c-Mvc等非蛋白 因子起到乙酰化修饰作用,增强这些蛋白的 转录因子活性。CBP/p300复合体还参加了染 色质重排(Chromosomal rearrangement) 过程,促进了肝癌的形成。在结肠癌和乳腺 癌等多种实体瘤中也发现了因为突变而缺失 了功能的CBP/p300复合体。染色质的动态调 控作用还需要核小体重构机制(nucleosome remodeling)的参与。BAF重构复合体(BAF remodeling com plex) 能够极大地增强重编 程作用的效率,因为该复合体可以促进转录因 子Oct-4与靶位点之间的结合。如图2B所示, 在多种肿瘤中,包括ARID1A在内的SWI/SNF 同系物都是遗传失活的主要因子, 不过其中的 机制现在还不甚明了。

上述这些研究以及其他一些研究都表明,染色质调控因子除了像我们之前了解的那

样,能够在发育和重编程过程中发挥作用之 外,还在人类肿瘤的发生发展过程中发挥了各 种各样的作用。很多染色质调控因子都具备相 似的催化活性和发育表型,这种情况在不同的 恶性疾病的具体情况下往往会产生互相矛盾的 结论, 因为这些染色质调控因子可能会在基因 表达调控、分化、老化、端粒调控以及其它生 理病理过程中发挥不同的作用。同一个染色 质调控因子在多种不同的情况下可以发挥不 同的作用,比如可以促进表观遗传学重编程 (epigenetic reprogramming) 过程,也可以 抑制细胞的分化,维持细胞的表观遗传学现 状,还可以起到抑制凋亡或者老化的作用。恶 变前细胞 (premalignant target cell) 的表观 遗传学状态会发生改变,或者维持某种肿瘤细 胞特定表观遗传学状态,这都是突变的转录因 子和染色质调控因子的作用, 所以这些因子在 不同的情况下可能会起到不同的作用。

# 8. 重编程与肿瘤干细胞的关系及其它

本文探讨了重编程以及肿瘤相关概念中的异同点,重点突出了通过肿瘤基因组测序研究发现的,在这两种不同生理过程中都能发挥作用的转录因子和染色质调控因子。与重编程机制有相似作用的事实表明,有一些染色质调控因子在肿瘤的发生过程中出现了变化,这可能是某些原本就容易发生表观遗传学状态重建改造作用的细胞里最开始出现的变化,这有利于这些细胞发生转化、去分化、增殖不受控制、或者难以凋亡。后面陆续发生的其它一些事件则能够进一步巩固转化的结果,促使肿瘤组织里不断出现各种细胞(图3)。在某些肿瘤中,干细胞样的肿瘤细胞是肿瘤形成的关键,它们也成为了肿瘤组织里的主体。从事重编程研究的科研人员正在试图利用各种转录因

子来认识细胞,与此同时,从事肿瘤研究的科研人员对肿瘤组织里各种各样的表观遗传学状态的重视程度也在不断增加。肿瘤细胞通过激活某些转录因子以及对相关染色质调控因子来进行调控的方式,能够对自身的表观遗传学状态进行动态的调控,使已分化的肿瘤细胞重新转化成干细胞样的肿瘤干细胞,促进肿瘤组织的生长。肿瘤细胞这种动态的双向转变理论虽然目前还只是一种假说,但是却可以让我们用一种观点来看待肿瘤组织里的所有细胞,所以这种理论完美地结合了肿瘤干细胞理论和随机突变理论。不论用哪一种理论来解释细胞恶变,细胞重编程机制和恶性转化机制的共同点都有助于我们更深入地认识肿瘤的生物学本质,也有利于指导对肿瘤的临床诊治工作。

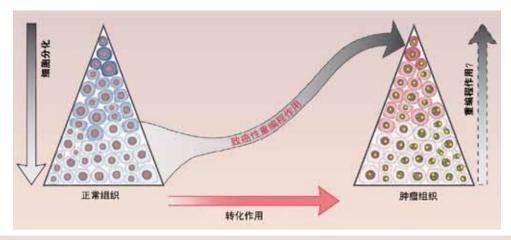
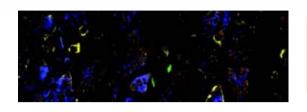


图3 正常组织和肿瘤组织里的细胞等级层次。图左和图右分别展示的是正常组织和恶性肿瘤组织里的细胞等级层次情况,这些细胞的表观遗传学修饰情况不尽相同,其中有极少数细胞是干细胞,它们能够按照图中颜色渐变的箭头方向所示,经过一系列的中间步骤产生出已分化的子代细胞。细胞重编程实验已经发现,细胞的分化过程是可逆的,如图左和图右的箭头所示。图中红色箭头表示的细胞转化过程是一个不断的积累遗传突变和表观遗传学事件的过程。细胞转化过程一旦启动,就会陆续发生其他的、不同的事件,最终会形成一个遗传背景多种多样(图中黄色和绿色星形图案表示)的肿瘤组织,在每一种遗传背景相同的亚克隆细胞里,还会出现表观遗传学背景不同(图中渐变的颜色表示)的情况。染色质调控因子和转录因子等关键调控因子的活性如果发生改变,也会对肿瘤起到双重作用,可能会促进细胞转化,也可能会促进表观遗传学状态转变,即起到致癌的重编程作用。我们推测,这些调控因子还会在肿瘤形成之后继续发挥作用,促使已分化的肿瘤细胞重编程为干细胞样的肿瘤干细胞,在肿瘤组织里细胞分化与重编程这两种作用最后会打成一个平衡的状态。



# 四、肿瘤的异质性



混杂的肿瘤组织。图中的乳腺癌组织切片里呈现出 多种颜色的荧光,这些不同颜色的荧光代表了多种 不同的分子水平上的异常,这也说明在同一个肿瘤 组织里,肿瘤细胞并不是全都一样的。

越来越多的遗传学研究结果表明,肿瘤是由一群异质性细胞组成的混合体,这可能就是让临床医生的抗癌治疗屡次失败的原因。

作为一名肿瘤学家,Charles Swanton也经常会无奈地告诉前来找他的晚期肺癌和乳腺癌患者,他也束手无策。尽管临床上有好几套抗癌治疗方案,每一种方案的疗效在最开始也都还算不错,可是最后都会失效,而且劫后余生的肿瘤会展开更加疯狂的反扑,生长得比以前还要迅速。"肿瘤细胞就好像知道我们下一步要干什么似的,它们早早地就做好了准备,我们只能迎来失败的结果。" Swanton无可奈何地说道。

不过Swanton现在似乎已经找到了失败的原因。在英国伦敦癌症研究所(Cancer Research UK's London Research Institute)里,Swanton领导着一个科研小组

正在对肿瘤展开研究,他们最近对一位肾癌患者不同部位的肿瘤组织进行了DNA测序,结果发现这些同属于一个肾脏的肾癌组织在DNA水平上是有差异的。虽然原发灶和其它部位,包括转移灶的肿瘤组织的确存在一些共同的遗传变异,但是很多变异都不是普遍存在的,只会在某些部位特异性地出现,这表明肿瘤组织里存在多种不同的肿瘤细胞成份。所以Swanton认为,其中有一些细胞可能会耐受某种抗癌疗法,还会因为采用了这种抗癌疗法而获得生长优势。"对我而言,这个研究成果终于让我们发现了一点线索,能够解释为什么在抗癌治疗工作中总是败下阵来。" Swanton最后总结道。



"对我而言,这个研究成果终于让我们发现了一点线索,能够解释为什么在抗癌治疗工作中总是败下阵来。"

——英国肿瘤学家CHARLES Swanton

Swanton的研究结果不仅以确凿的证据验证了很多肿瘤临床医生都非常熟悉的事实,同时也给他在科学界带来了意想不到的名声。Swanton团队去年3月在大名鼎鼎的《新英格兰医学杂志》(The New England Journal of Medicine, NEJM)上发表了他们的研究成果。到去年底,他们的这篇文章就已经被引用了143次,据Thomson Reuters公司(让我们又爱又恨的SCI就是他们搞的)的统计,这一引用率远远超过了2012年发表的其它原创性生物医学研究论文。各种演讲邀请也雪片般地飞来,Swanton的邮箱都要被挤爆了。"我这辈子从来都没这么忙过。"今年41岁的Swanton说道,他在5年前才有了自己的实验室。

Swanton开展的这种基因组学研究让科研人员对一个旧观点又重新拾起了兴趣,早就有人认为,肿瘤就是由多种不同细胞组成的混合体,现在Swanton等人又用现代科技技术证实了这种假说,否定了我们一度当作正确的,认为肿瘤组织主要是由一群遗传背景一模一样的肿瘤细胞组成的观点。现在发现,肿瘤组织里往往会包含有多种遗传背景不同、但是彼此之间又有所关联的亚群肿瘤细胞。随着肿瘤的进化,同样的肿瘤细胞会不断分裂、分化,出现新的突变和各种遗传错误,就好像一棵大树在不断生长的过程中会伸出新的枝桠一样。所以,同一个肿瘤组织里的两个部位,或者同一个肿瘤原发灶形成的多个转移灶,它们彼此之间都会出现很大的遗传学差异。

肿瘤内部这种遗传异质性所达到的程度 让肿瘤生物学家和肿瘤临床医生都吓了一跳。 他们不得不重新修改预定的,比如针对某种肿 瘤突变制定的研究和治疗方案,因为这些突变 信息基本上都是对一个肿瘤样本进行研究得来 的结果。临床医生也担心,根据这项最新的研 究结果,他们在临床上对一次肿瘤活检样品进 行检测得来的遗传学和分子生物学的结果可能 也不太可靠,以此结果指导临床治疗不一定会 带来比较好的治疗效果。可是目前这种治疗 略在临床上正有大肆推广的趋势。很多针对某 种突变靶点起作用、非常有潜力的新药往往也 是在最开始表现出一定的"奇效"之后很快失 去了功效,这可能也是因为肿瘤具有的这种异 质性特点在作怪。由于这些特效药杀死了一部 分敏感的肿瘤细胞,结果却让耐药的细胞获得 了生长优势,最后耐药细胞成为了肿瘤组织里 的主力军。

"随着研究的不断深入,我们发现肿瘤是一只狡猾的狐狸。在肿瘤组织内部也充满了竞争和进化,所以不论我们用什么样的抗癌手段,它们最后都能躲开。"英国Wellcome基金会Sanger研究所(Wellcome Trust Sanger Institute in Hinxton, U.K)的肿瘤遗传学家Peter Campbell这样评论道。

不过也有一些研究人员比较乐观,认为塞翁失马焉知非福。既然我们已经认识到肿瘤的异质性问题,那么就可以有的放失地进行处置,达到更好的抗癌治疗效果。"这一发现虽然让我们大吃一惊,但同时也让我们非常兴奋,因为我们总算知道是什么让肿瘤表现出这种异质性了。"美国波士顿Dana-Farber肿瘤研究所(Dana-Farber Cancer Institute in Boston)的Kornelia Polyak这样说道,Polyak主要对乳腺癌组织里的异质性问题开展研究。



"将来在这个研究领域还会开展更多的试验研究工作,但是我认为我们也需要静下来好好想一想,解决肿瘤异质性问题究竟对临床抗癌工作有什么样的帮助。"

——Dana-Farber肿瘤研究所Kornelia Polyak

# 1. 细胞亚克隆间的战争

病理学家在19世纪中末期就已经开始在显微镜下观察肿瘤组织了,从那个时候起,他们也已经开始认识到肿瘤组织并不是由单一的细胞组成的,这是最早观察到的肿瘤异质性现象。在20世纪70年代至80年代初期,肿瘤生物学家又发现从同一个肿瘤组织不同部位取到的肿瘤细胞在小鼠内的成瘤能力,以及转移能力是有非常大的差异的,这些细胞对抗癌药物的反应也是不一样的。

认为肿瘤是由多种不同细胞混合而成的提法也非常符合肿瘤的克隆进化理论(clonal evolution theory of cancer)。1976年,肿瘤生物学家Peter Nowell在《科学》(Science)杂志上提出了这种理论,他认为一个细胞随机获得了一系列突变,获得了超强的增殖能力,并且在众多细胞(其它的、不同的突变细胞或克隆)中取得了竞争优势(达尔文式的生存竞争)之后就会形成肿瘤。不过Nowell也认为适应性不佳的克隆会死亡,被淘汰出局,最终只有一个最有竞争力的克隆会胜出,成为肿瘤组织里的绝对主力。肿瘤组织里其它细胞的突变程度都不如这种最多数的细胞。这一观点后来又得到了更多的支持,

因为美国约翰霍普金斯大学(Johns Hopkins University)的研究人员Bert Vogelstein和他的研究团队成功地将这一理论应用于克隆性肿瘤(colon cancer),描述了一个缓慢的、特定突变线性积累,将正常的上皮细胞转化成息肉(polyp),进而形成主要由一种细胞克隆构成的肿瘤的全过程。

可是在近十年里,Polyak的实验室和其他一些实验室对很多与肿瘤相关的异常病理情况,比如特定基因发生突变、基因拷贝数异常、蛋白表达量异常等进行了研究,结果发现在肿瘤组织里存在多种不同的细胞亚群,或者叫亚克隆(subclone)。只有当DNA测序的成本降到一定程度,同时科研人员能够取到肿瘤组织深部的细胞进行测序的时候,我们才有可能发现罕见的突变,不过现在已经有人开始从事这方面的工作,掌握肿瘤组织的全貌了。

比如在近4年时间里,世界上多个实验室,包括一个由美国华盛顿大学医学院(Washington University School of Medicine in St. Louis)的Elaine Mardis领导的课题组,对乳腺癌细胞的基因组进行测序研究,结果发现一位乳腺癌患者体内的转移灶与原发灶

之间在遗传水平上就有非常明显的差异,而且这些转移灶应该都是由原发灶里某些亚克隆细胞演变而来的。Vogelstein的团队对取自多个前列腺癌组织不同部位的组织样品进行了基因组测序,结果发现了多个存在遗传差异的细胞亚克隆,而且可以看出这些亚克隆之间随时间演进的过程。其中有一些亚克隆源自原发灶,然后在前列腺外组织形成了转移灶。包括Mardis团队进行的基因组测序工作在内的多项对白血病进行的研究都发现,这些血癌细胞有时也会"分裂"成多个不同的亚克隆。

这些试验(得到进化分支结论,而不是 线性演讲结论的实验) 大部分检测的都是血液 标本或组织活检标本, 所以这种试验得到的结 果都是多种细胞混合体的检测结果。Nicholas Navin以及美国纽约冷泉港实验室(Cold Spring Harbor Laboratory in New York) 的 遗传学家Michael Wigler实验室的同仁采用了 另外一种方法对肿瘤的异质性问题进行了研 究,他们的方法是从乳腺癌肿瘤活检组织中 分离出单个的细胞,对其进行DNA测序。通 过这种方法发现了很多种截然不同的亚群细 胞,其中就包括爆发式获得新生突变的细胞, 这很符合达尔文的间断式进化(punctuated evolution) 理论,可是却不符合我们一直 认为的循序渐进式的突变方式,美国休斯 敦市德克萨斯大学MD Anderson癌症中心 (University of Texas MD Anderson Cancer Center in Houston)的Navin介绍道。

在所有对实体瘤的遗传性异质性问题开展的研究工作中,最引人注目的可能就是由Swanton的实验室进行的工作了。除了进行其它分子学检测之外,Swanton等人还对取自一个肾癌肿瘤标本的不同位置,以及同一名患者体内转移灶标本的外显子组进行了测序。他们共计发现了128个突变位点,其中只有1/3的突变是在所有被测标本中都能检出的共有突变。对另外3名患者进行同样的测序研究,也得到了同样的异质性结果。即便是同一个肿瘤

基因,在不同部位的肿瘤组织中,该基因的突变情况都会不一样,这说明即便针对同一个基因,各部分细胞的进化历程也会不一样。

据Swanton介绍,有一些能够促进肿瘤细胞生长的、在肿瘤发病初期就已经出现的致病突变(driver genetic change)在肿瘤的发展、转移和扩增过程中会一直存在,这可是一个天大的好消息,因为这样一来,以这种突变为靶点的抗癌药物就可以一直使用下去,而且对绝大部分肿瘤细胞都会起到杀伤作用。可是Swanton自己进行的研究,以及其他一些人开展的研究工作却发现,如果某种药物在使用之初对转移灶有抑制作用,那么在经过一段时间的治疗之后,它十有八九会失效,这是因为这种药物对携带其它罕见突变的肿瘤细胞不起作用,敏感细胞都被杀死之后,肿瘤组织里基本上就只剩下这些不敏感的细胞了,所以治疗会失败。

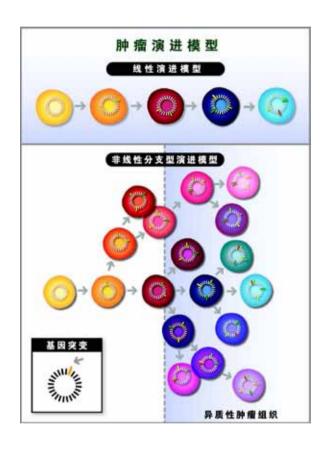
还有一些科研人员认为,肿瘤的耐药性可能来自其它方面的因素导致的差异,而不是不同的基因突变。他们发现,对DNA的化学修饰作用,或细胞内能够决定基因表达开启功能的其它分子波动都会使肿瘤细胞具备抗药性。"这无疑使问题变得更加复杂了。"加拿大多伦多大学(University of Toronto in Canada)的John Dick这样说道。John Dick和Shibata在今年2月1号出版的《科学》(Science)杂志上发表了一篇论文,他们发现结肠癌组织中遗传背景相似的肿瘤细胞会对抗癌药物表现出不一样的耐药性。

Polyak认为,肿瘤组织的异质性,不论是源自DNA突变,还是细胞内的其它因素,都可以解释抗癌治疗为什么会失败。Polyak还认为这种异质性导致的耐药机制要比肿瘤干细胞理论(cancer stem cell)更加合理,因为这种肿瘤干细胞理论认为,肿瘤之所以会产生耐药性是因为少数携带耐药突变的肿瘤细胞在肿瘤组织中取得了绝对的数量优势。

## 2. 更有意义的问题

研究肿瘤进化问题的科研人员也很欢迎这种遗传异质性理论,不过他们对此并不感到惊奇,而且他们担心这一研究成果对临床治疗的意义不大。美国洛杉矶南加州大学(University of Southern California in Los

Angeles)的Darryl Shibata表示,所有这些研究成果都是描述性的结果,可最关键的问题是,这种异质性有重要的科学意义和实际应用价值吗? Shibata认为它的实际意义现在还不太明朗。



不过这并没有吓住受美国马里兰州贝塞斯达美国国立癌症研究院(U.S. National Cancer Institute in Bethesda, Maryland)资助开展癌症基因组计划(Cancer Genome Atlas)研究工作的科研人员,他们还是一如既往地继续探究肿瘤的异质性问题。据美国国立癌症研究院的Stephen Chanock介绍,他们有一部分工作是反过头来,对原发肿瘤灶的剩余标本进行测序,如果有可能的话,也要对转移灶标本进行测序,进一步了解肿

瘤的演进历程。英国Sanger研究所(Sanger Institute)的Campbell则希望在英国和其他地区开展的、另外的癌症基因组研究项目能够同时对原发灶和转移灶进行测序。有一些研究人员甚至认为现在开展的这些研究工作漏掉了最重要的、最致命的突变,因为他们只对一次活检取得的标本进行测序。Shibata就认为,这实在是有点说不过去,他们应该在最开始就进行测序检测的。

根据这些对肿瘤异质性开展的研究工

作,Campbell认为在临床工作中至少应该做出一次调整,他认为临床医生如果要为患者制定抗癌治疗方案,那么在条件允许的情况下,应该对肿瘤原发灶和转移灶进行基因组测序。而且在整个病程过程中,也需要进行多次组织活检检查,随时掌握肿瘤的演进情况。

据Swanton介绍,围绕肿瘤的异质性特 点还有一个非常重要的问题一直没能得到解 决, 那就是遗传背景更加复杂多样的患者预 后是否就更差。根据Dana-Farber癌症研究 所和美国博大研究所(Broad Institute)之前 取得的研究成果,Swanton等人认为应该是 这样的。今年2月在《细胞》(Cell)杂志上 发表了一篇文章, 该文指出, 在接受化疗的 慢性淋巴细胞白血病 (chronic lymphocytic leukemia)患者人群中,那些"原发灶"肿 瘤细胞中含有一个或多个致癌基因的人接受治 疗的频率要高一些,他们的死亡率也会高一 些。Campbell的课题组也发现,在脊髓发育 不良 (myelodysplasia) 这种白血病前期疾病 (preleukemia condition) 患者人群中,体内 存在更具侵袭性亚克隆细胞的人更容易发展成 为白血病患者,而且死亡率也更高。

很多有抱负的科研人员都希望能够更仔细地研究肿瘤的异质性问题,他们希望能够制定出一套分级标准,帮助指导临床治疗工作。比如美国加州大学旧金山分校进化与癌症中心(Center for Evolution and Cancer at the University of California, San Francisco)的主任Carlo Maley就针对Barrett食管(Barrett's esophagus)这种癌前病变(这是一种食道癌的癌前病变,具体表现为食道下部细胞出现异常)的演进过程建造了一个模型,该模型认为,在数年前尚未发展成真正的食道癌的时候,如果癌前病变细胞的遗传背景越丰富,即异质性越高(在该模型中表现为数值越高),那么今后癌变的机率就越大。

还有一些研究人员比较乐观,他们认为 肿瘤的异质性特点不会影响靶向抗癌药物的功效。Vogelstein认为,由于绝大部分癌症患者 体内的原发病灶都会被医生通过手术的方式解决掉,所以真正妨碍治疗工作的只是转移灶里的那一小撮顽固分子。比如有一些转移灶组织里会有两条非常重要、而且互不重叠的信号通路发生了突变,但是我们完全可以使用针对这两条信号通路的特异性靶向药物进行治疗,Vogelstein认为在这种情况下进行联合用药治疗是会收到很好的疗效的,而且能够极大地降低耐药情况出现的机率。

Swanton则希望通过阻止异质性现象出现的方法解决异质性问题。他的课题组最近就通过对结肠癌的研究发现,结肠癌细胞往往会缺失一些能够确保DNA进行高保真复制的基因,所以经常会出现各种染色体异常(chromosomal abnormality)。因此Swanton认为,如果将来出现了一种可以确保在细胞分裂时DNA复制不会出错的药物,那就能够极大地减少异质性发生的机率,从理论上来说也就可以减少肿瘤细胞耐药的机率了。

美国佛罗里达Moffitt癌症中心(Moffitt Cancer Center in Tampa, Florida)的放射科 医师Robert Gatenby则认为,更根本的方法 就是不要去想如何才能将肿瘤细胞赶尽杀绝, 让它们长好了, 那些耐药的肿瘤细胞会受到肿 瘤组织里主流细胞的压制, 所以让它们共存, 互相制约才是最好的方法。Gatenby用小鼠进 行了试验, 他给小鼠移植了高度恶性的卵巢癌 组织,然后给小鼠使用了低剂量的化疗药物 (这种剂量的药物不能清除肿瘤细胞, 但是能 够让肿瘤的生长速度减慢, 让肿块的大小稳定 在一个固定的水平上),结果发现这些小鼠和 对照组相比寿命至少会延长一倍, 能够活到 4个多月。Gatenby将他的这种疗法称作"适 应性抗癌疗法(adaptive therapy)",他指 出,我们应该对万物的进化抱有一种敬畏之 心,也应该接受(肿瘤)进化给我们的抗癌治 疗带来的影响, 要学会利用这种进化机制, 而 不要被它所击败。

不过据Gatenby介绍,他在2009年曾经申请进一步扩大他的研究工作,结果他的标书

迎来了史上最差评语。但是在去年秋天,他终于获得了美国国立癌症研究院创新课题研究项目(Provocative Questions program)提供的高达210万美元的科研基金。这个创新课题研究项目就是要鼓励科研人员开展新颖的、异想天开的研究工作。Gatenby计划将立即启动一个小型的临床研究项目,招募一批前列腺癌患者,用低于常规剂量的抗雄性激素药物

(antiandrogen therapy) 对他们进行治疗,研究者通过利用影像学方法和数学模型来预计肿瘤的生长情况,从而定期调整药物剂量。

Polyak认为这种策略值得一试,他指出,将来这个研究领域还会得到更多资助,乃至开展更多的临床试验,但是,还需要一段时间才能知道究竟是什么会对临床抗癌工作有帮助。



# 五、Elaine Mardis, 让肿瘤基因组学 走上快车道的优秀女车手



图中从左至右分别是Timothy Ley、Elaine Mardis和Richard Wilson,他们正在努力将全 基因组测序技术应用到白血病 (leukemia) 的临床诊疗工作当 中。

对DNA测序技术有着一股子热情的Elaine Mardis现在又想利用她的专业知识为癌症患者排忧解难了。

在ELAINE MARDIS第一次向跆拳道黑带段位发起冲击的时候,她的教练Richard Wilson还是有一点担心的。因为在考试中需要完成用无支撑侧踢(side kick through a free-standing)踢碎5厘米厚铺路砖的测试项目。作为一项沿用多年的跆拳道黑带考试项目,现在的考试内容已经发生了一点小小的改变,铺路砖已经变成了环氧材料(epoxy)制成的砖块,这种砖比以前的铺路砖更难被踢碎。那天有很多参加考试的学生在第一次尝试时都失败了。"不过Mardis是个例外,她一击就成功了。"Wilson自豪地回忆道。

Mardis成功的秘诀就是专注以及一往 无前的动力,这也是支撑她能够在基因组 测序的道路上走过整整30年的力量源泉。 现在已经是美国华盛顿大学医学院基因组 研究所(Genome Institute at Washington University School of Medicine in St. Louis) 副所长的ELAINE MARDIS也是基因组测序领 域里为数不多的几位女性佼佼者之一。早在 DNA测序技术刚刚诞生之初,MARDIS就投 身到这个行业之中了,她现在正在致力于以更快的速度解开人类基因组密码的工作,同时也在积极运用基因组测序技术破解肿瘤难题。将来某一天,对肿瘤患者体内的肿瘤组织进行全基因组测序将成为临床工作中的一项常规检测。对于这种提法,现如今的肿瘤学家早就已经习以为常了,可是在Mardis等人首次破解肿瘤基因组密码之前,可没人敢这么想。过去大家都认为这项工作的难度之大简直无法想象,而且即便能够实施,成本也是贵得惊人。美国俄勒冈健康与科学大学(Oregon Health & Science University in Portland)的肿瘤系统生物学家Paul Spellman认为,这是一项基础性的研究工作,可不是随便一个人就能够轻而易举完成的。

Mardis等人的兴趣可不只是对一位患者,或者一种肿瘤进行全基因组测序。她们承担的工作是由美国国立健康研究院(NIH)发起的一个大型科研项目——肿瘤基因组图谱计划(Cancer Genome Atlas),该项目旨在弄清20种肿瘤的分子发病机制(详见*Science*、

15 September 2006, p. 1553)。Mardis等人正在利用基因组测序技术对多种肿瘤开展研究,希望能够找到有价值的、可作为药物作用靶点的基因和信号通路,同时也希望能够借此了解与肿瘤耐药机制相关的信息。Mardis的课题组甚至还用她们自己的专业知识挽救了课题组一位同事的性命。这位同事不幸患上了白血病,而且在经过治疗之后又再次复发

了。Mardis等人得知这一情况之后立即向他 推荐了一种Maserati方案,这是一种基因组系 列分析方法,通过这种方法,马上就帮这位同 事找到了一种特效药,阻止了病情的发展。 Mardis希望将来有一天这种方法可以成为临 床诊疗常规中的一项标准手段。Mardis目前 的计划是,让这种Maserati方案尽可能地给临 床患者提供帮助。

## 1. 学习测序专业知识

Mardis不仅仅是因为她的学问才会在人 群中脱颖而出的。她喜爱粉红色,而且也乐于 让大家知道这一点, 所以艳粉色的唇膏就成了 她的标志。现在,她就正穿着一件粉红色的冬 装,和一双玫瑰色的蛇皮靴。在她的办公室里 也延续了这种大胆的用色风格,每一个来到她 办公室的人都会注意到那好几排各式各样的红 色玩具车。在Mardis的办公桌上整齐地摆放 着几十辆风火轮赛车。在Mardis的面前还有 一个展示柜,里面陈列着1962年和2002年的 法拉利车模。据Mardis介绍,她曾经在加利 福尼亚州开过她朋友的一辆法拉利Scuderia Spider跑车,这是最让她愉快的一次经历。 Mardis真的很喜欢车,尤其是红色的车。 Mardis现在开的是一辆红色的奥迪S5,但是 她最喜欢的还是保时捷911Carrera S。

不过下了公路,DNA测序仪就成了Mardis最钟爱的机器,测序速度越快她越喜欢。受到当化学教授的父亲的影响和鼓励,Mardis也选择了科学作为她的终身职业。这名来自美国内布拉斯加州(Nebraska)的姑娘考上了美国俄克拉荷马大学(University of Oklahoma, OU)的生物学专业。大学毕业之后,俄克拉荷马大学的生化学家Bruce Roe又说服Mardis从研究果蝇遗传学的方向转到了分子生物学方向,"因为在遗传学领域,DNA只是一个抽象的物体,可是在分子生物学领域,我却可以切切实实地对DNA展开研学领域,我却可以切切实实地对DNA展开研

究。"Mardis回忆说。在转入Roe门下之后,她们实验室很快就买了一台世界上第一批商业化的DNA"测序"仪,Roe让Mardis来改进这台只能完成机械点样工作的仪器,使其能够用于科学研究。"我们当时就想,怎么样才能将这种测序方法放大,使其成为一台流水线式的、真正能够产生DNA序列数据的测序仪呢?"Mardis回忆说。

为了打造这条流水线,Mardis在攻博期 间对某种细菌长达5000bp的序列进行了测 序,这在1989年时可算是一项难度不小的工 作, 堪称是一项壮举。从那时起, 技术开发 就成了Mardis的工作目标。博士毕业之后, Mardis进入美国加利福尼亚州的Bio-Rad公司 从事产品开发工作。当时Bio-Rad公司正在开 发一种新型的DNA聚合酶,这种酶对于DNA 测序工作具有非常重要的价值。Mardis在Bio-Rad公司学到了全套的(包括好的和坏的)管 理理念和技能,也亲身经历了一款生物产品开 发的全过程,对产品开发有了深刻的体会。这 段经历进一步提升了Mardis的价值,所以在 1993年, Roe实验室的Wilson又将Mardis重 新召回了华盛顿大学,让她到大学新成立的基 因组研究中心工作。Mardis的新工作是开发 高通量DNA测序技术和高通量DNA测序仪。

Mardis只用了短短几个月的时间就拿出了一个分离待测DNA样品的新方案。当时, 大家都在用一种由M13病毒和细菌组成的系 统制备(扩增)待测DNA样品,但是使用这 种方法还需要将待测DNA和细菌自身的基因 组分开, 纯化出待测DNA样品才能进行后续 的测序操作,比较麻烦。采用Mardis发明的 新方法则会大大提升测序工作的效率,而且 对操作的要求也没那么苛刻,实验中用到的 有刺激性气味的试剂也会少一些, 更重要的 是,不会像传统的酚提取方法那样产生有害的 试验废物。这项成果是Mardis在近二十年来 取得的一系列技术改进成果中的第一项。其 间,她和她的同事用这种方法于1998年首次 对动物的基因组进行了测序,得到了秀丽隐杆 线虫(Caenorhabditis elegans)的基因组序 列,后来又进一步向人类基因组序列发起了冲 击,她们的人类基因组测序工作比私人公司开 展得还要早(Science, 16 February 2001, p. 1177)。很快,Mardis的课题组就扩充到了 20多人, 在她的科研团队中除了生物人才之 外还吸纳了电子工程师和机械工程师等其他方 面的专业人才,依靠这些人才,才能让DNA 测序工作走向标准化、自动化。"我们总是定 下非常'远大的'目标,虽然这些目标在当时 看来都是不可能完成的任务,但是我们可没有 想过失败的结果。所以我们只能卷起袖子,想 尽办法解决我们面临的一切问题。" Mardis 介绍说。

对于这一点,Mardis的同行们有着非常深刻的印象。比如在美国马萨诸塞州坎布里奇白头研究院(Whitehead Institute in Cambridge, Massachusetts)里从事大规模人类基因组测序工作的Chad Nusbaum就评价说: "从某种角度来说,Mardis是我们这些人的榜样和标杆。干我们这一行的面对的可

都是科学大难题,所以我们都得充满了干劲才行,Mardis要比我们中的大多数人还要积极。"

科学家们在2001年首次公布了人类基因组草图。就在这幅草图即将完工的时候,华盛顿大学的科学家们又开始向新的目标进发了,他们要对小鼠、鸡、鸭嘴兽、玉米和斑马鱼进行基因组测序。再后来,华盛顿大学的基因组研究所(就是当年的基因组研究中心,该中心在成立两年之后更名为基因组研究所)又开始对更多的人进行人类基因组测序。从2010年开始,该研究所又逐渐转向了肿瘤基因组测序方向,到今天,他们有60%的研究工作都与人体肿瘤基因组相关。

之所以会出现这种转变,有部分原因是因为从2005年左右开始,DNA测序技术的成本有了大幅度的降低,而且测序速度也有了飞速的提升。据Mardis回忆,有了这些新的测序仪,她们总算是看到了基因组测序的希望,当时他们都有'总算是等到了这一天'的感觉。

虽然测序业内人士依旧把Mardis看作是行内的关键人物,而且Mardis也的确是业内的大腕,她是一年一度的DNA测序大会(DNA sequencing conference,业内人士每年都可以在这个大会上汇报最新的研究进展,与同行们进行交流)的创始人和组织者,不过Mardis可没有松懈下来,肿瘤基因组测序这个新挑战仍然激励着她在科学的道路上继续前进。Mardis这样说道:"我对技术的热爱从来就没有停息过,可我的工作重点的确调整了好几次,我现在认为,开展肿瘤基因组学研究是当务之急。"

## 2. 肿瘤基因组学

在Mardis办公桌上摆放着的那些跑车模型中,我们还可以看到Mardis对肿瘤基因组学研究的关注,在她的桌上还放着一本由遗

传学家Robert Weinberg编撰的经典教材——《肿瘤生物学》(*The Biology of Cancer*)。 通过阅读这些书籍,参加肿瘤学会议,以及 与临床医生进行沟通,Mardis现在对肿瘤已经有了一个大致的认识和了解,在很多肿瘤学家来向她寻求基因组测序方面的帮助时,Mardis也能够和他们聊几句肿瘤专业的话题了(不过很讽刺的是,Weinberg曾经多次在公开场合质疑过肿瘤基因组学研究对于临床患者的意义和价值,他认为进行肿瘤基因组学研究和传统的肿瘤单基因研究相比没什么太大的区别)。

比如在8个月前,Mardis曾经在意大利做过一次演讲,会后有一位来自美国哈佛大学的临床医生找到Mardis,这位医生多年来一直在收集肺癌患者在对抗癌药物出现耐药性之前和之后的肿瘤组织样本,他表示愿意和Mardis分享这些珍贵的研究材料,共同开展研究,希望通过DNA测序研究了解肿瘤细胞演进的动态过程。由于Mardis在基因组测序方面积累了多年的经验,所以她非常自信地第一个踏入了这个研究领域,对肿瘤组织进行全基因组测序,目前她所涉及的肿瘤类型已经包括乳腺癌、脑癌、肝癌、前列腺癌、肺癌和血癌等七种肿瘤。

早在2007年, Mardis、Wilson和华盛顿 大学的另外一位肿瘤学家Timothy Ley就开始 了肿瘤基因组方面的研究工作, 他们当时选择 的研究目标是白血病,希望利用基因组测序技 术找到急性髓细胞样白血病 (acute myeloid leukemia, AML)的致病突变,他们之所以选 择这种疾病是因为在美国每年大约有1.3万人 患上这种疾病,有8800多名AML患者会因此 病死亡。2007年以前,肿瘤基因组学家都认 为找到染色体易位(translocations)等大规 模染色体畸变(large-scale aberration),或 者研究可疑致病基因才是王道, 因为之前的研 究表明,这些异常都与AML发病有关。不过 她们三人可不这么认为,她们决定另辟蹊径, 利用新一代的基因组测序技术,对肿瘤细胞的 基因组进行一番彻底的检查, 然后与患者本人 正常细胞的基因组进行比对, 看看都有哪些差 异,发生了哪些改变。与此同时,还有一些科 学家对肿瘤细胞进行了外显子组(exomes,即基因组中编码蛋白质的所有序列,这些序列在整个基因组中只占很小的比例)测序,试图发现其中的突变。相比之下,全基因组测序发现的信息更多,能够从全局的角度审视肿瘤细胞,除了能够发现蛋白质编码基因的问题之外,还能够发现基因缺失或基因表达调控区域里的异常等其它情况。

据Mardis回忆,她们当时为这个研究项 目重新修改了申请基金的标书,向NIH申请 了100万美元的研究经费,但是当时负责基金 评审工作的评委"讨厌"她们的计划。不过 Mardis等人并没有气馁,又把标书寄给了当 地的一名慈善家Alvin J. Siteman, 华盛顿大 学的肿瘤研究中心就是以Siteman的名字命名 的。第二天Siteman就给她们转了价值100万 美元的股票作为她们的科研经费, 让她们能 够开展这项类似于人类基因组计划的工作。 Mardis等人首先花费了16个月学习如何使用 新一代测序仪测出的DNA短片段序列,然后 才开始对这些数据进行分析和研究。据在美 国马萨诸塞州坎布里奇博大研究所(Broad Institute in Cambridge) 从事肿瘤基因组研 究工作的Gad Getz介绍,这项工作真的挺难 的,尤其是数据分析工作,那对他们这些人来 说简直就是一场噩梦。

不过辛苦的工作总算是带来了回报,Mardis等人的这项工作成果2008年发表在了《自然》(Nature)杂志上,那可真称得上是一篇让人眼前一亮、眼界大开的文章。Mardis小组一共发现了10个与AML有关的突变,其中有两个突变是之前已经被发现的,另外8个突变都是她们新发现的。Getz说道:"Mardis等人的这项工作让我们这些人第一次意识到,原来对肿瘤细胞进行全基因组测序也是一条不错的肿瘤研究思路。当时很多人看到这篇文章之后都异常兴奋。" Spellman也说道:"似乎就在一夜之间,肿瘤全基因组测序技术就在肿瘤医疗活动中获得了一席之地,是Mardis等人为肿瘤患者开启了个人基因组

测序的新时代。"

被研究的第一例肿瘤标本来自一名50多 岁的女性AML患者,她在初次治疗时取得了 不错的治疗效果,不过在一年之后病情又复发 了,可是这一次她可就没那么幸运了,病情 没能得到控制,她在病情复发一年之后不幸 去世。后来Mardis小组又于2009年在《新英 格兰医学杂志》(The New England Journal of Medicine) 上发表了第二例AML患者的肿 瘤基因组测序结果。这位患者的基因组中共 有12个基因发生了突变,而且在他的调控序 列(包括疑似调控序列)中,还发现了大约 50处碱基突变的情况。通过对其他AML患者 进行肿瘤全基因组测序之后发现, IDH1和 DNMT3A这两个基因都有助于预测AML患者 的预后状况。Mardis解释说,虽然每一位患 者的突变情况各不相同, 但是她们希望通过对 大量AML患者进行肿瘤全基因组测序后,可 以得到一份AML共有突变的遗传图谱,至少 能够发现受到AML影响的信号通路,这将对 开展针对性的治疗大有裨益。

之后Mardis小组又对8名AML患者进行了肿瘤细胞全基因组测序研究,不过她们都是在这些患者被确诊,并且接受过化疗,可是病情又再次复发以后对他们进行测序的。这些复发的肿瘤细胞还是源自患者骨髓组织里最开始恶变的那些肿瘤细胞,不过Mardis小组发现,这些原始肿瘤细胞还是出现了一些改变,她们

在2012年1月11日出版的《自然》杂志上撰文 指出: "我们首次发现, 化疗能够对肿瘤细胞 的DNA造成伤害, 在肿瘤细胞的全基因组范 围内都能发现这种伤害作用留下的印记。所以 说化疗的确会使肿瘤细胞产生新的突变, 这可 能就是患者病情复发的原因。通过我们的研究 证实了这种推测。"



基因组自动测序仪。Mardis发明了这种全自动 基因组测序仪,仪器可以自动挑出待测DNA样 品进行测序。这套设备为人类基因组计划节约 了大量的人力和时间。

#### 3. 让肿瘤基因组测序走进临床

两年前,肿瘤基因组学还只是Mardis开展的一项研究工作而已。可是在2011年的7月,Ley实验室的博士后,曾经身患急性淋巴细胞白血病(acute lymphoblastic leukemia,ALL),但是被成功"治愈"的肿瘤科医生Lukas Wartman的病情再次复发了。当时Ley和Mardis也才刚刚开始ALL肿瘤基因组测序方

面的研究工作,于是她们决定对Wartman进行一番彻底的基因组测序,而且还对肿瘤细胞的转录组(即细胞内所有的RNA)进行了一番分析,看看都有哪些基因被激活了。据Wartman回忆,他们当时的工作更像是一项单纯的学术研究,与改变他的治疗方案几乎没有什么关系。然而,当Wartman的化疗失败

之后,病情变得越来越重,于是这场测序工作似乎就变成了一次与死神的赛跑。

四周以后,基因组测序结果出来了,可是结果却让人很失望。在Wartman的肿瘤细胞里发现的所有突变全都不是现有抗癌药物作用的靶点。可是几天之后拿到的转录组测序结果却给他们带来了惊喜,在Wartman的肿瘤细胞里*FLT3* RNA的含量非常高,这些转录产物能够促使肿瘤细胞生长和分裂。

Obi Griffith和Malachi Griffith这对双胞胎都是Mardis课题组的成员,他们俩的工作是建立一个数据库,将基因突变与可以抑制这些基因编码蛋白质的药物一一对应起来,这样就可以选择更有针对性的药物对患者进行抗癌治疗,减少副作用的发生,而且与传统的化疗药相比,还可以降低药物诱发肿瘤细胞产生突变的机率。多亏有了这个数据库,Mardis等人很快就找到了特异性抑制FLT3蛋白活性的药物,这款药物已经获得了美国食品与药品监督管理局(FDA)的批准,被用于肾脏肿瘤的临床治疗工作。



积极的生活态度。自律而又非常努力的 Mardis对待跆拳道就和她对待工作一样 认真、有追求。

事实证明,这种药物非常管用,Wartman的病情再一次得到了控制。据Mardis介绍,那一次的实践结果让她非常满意,她希望能够在更多的肿瘤患者身上体会到那种成功的快感。

所以Mardis等人开始关注如何将肿瘤全基因组测序结果应用于临床诊疗工作。Mardis认为,需要将肿瘤全基因组测序结果与肿瘤转录组,甚至还包括外显子组测序结果结合起来,只有这样才能对肿瘤有一个全面的认识和了解,这就是她提出的Mardis式肿瘤诊疗方案(Maserati approach)。不过曾经与Mardis一起对乳腺癌进行深入研究的华盛顿大学的肿瘤学家Matthew Ellis认为,还应该加上蛋白质组学(即肿瘤细胞里所有的蛋白质)的研究结果。Getz说道:"他们都是我们这个领域里的先锋,现在整个学科都在朝着他们开拓的方向前进。"

不过要让肿瘤基因组测序技术走进临床还需要在测序技术上取得更大的突破。 Mardis的课题组就正在想办法解决肿瘤样品的处理技术,因为目前的样品处理方法还不能满足测序的需要。要么是肿瘤组织样品的量太少,不能进行测序;要么是因为肿瘤组织被福尔马林泡过,或者是在石蜡里包埋过,不能进行测序。Mardis等人还在想办法缩短测序工作的时间。现在完成一次测序大约需要1个月,她们希望能够缩短到数天。现在Illumina公司推出了一款最新的测序仪,可以将以往10天的工作量缩短到1天,Mardis等人正在使用这台测序仪进行试验。

Mardis还在尝试将肿瘤基因组测序结果应用到其它领域。比如她们最近新开展的一项乳腺癌项目就是研究如何利用肿瘤细胞全基因组测序结果,找出某一位患者肿瘤细胞表达的特有蛋白,并以此蛋白为靶点,开发个性化的免疫抗癌疗法,或者开发抗癌疫苗,防止肿瘤将来复发。还有一项刚刚起步的脑癌研究项目则计划对大脑中的小胶质细胞(microgliacell)进行转录组测序研究,因为小胶质细

胞是大脑中的一种支持细胞(supportive cell),可以为脑肿瘤细胞提供肿瘤刺激信号(cancer-stimulating signal)。

为了能够完成这么多的研究工作, Mardis每天早上四点半就起床了,然后离开 她们夫妇在2011年建造的法式田园风格的别 墅,驱车48公里,7点就坐到了办公桌前。 Mardis夫妇是一对空巢"老人",她们的女 儿正在上大学,家里只有两只约克郡犬和一只 德国牧羊犬。近几年,Mardis每年由于忙于工作,都有大约150天的时间不在家。所以现在Wilson不得不在Mardis的电脑上贴了一张便签,上面写着大大的"NO",希望她不要做那么多工作。可是急脾气的Mardis可没听那一套,这对于肿瘤患者可是一个大大的福音。Roe表示,Mardis总能冒出新点子,想出新办法,也总能提出新的问题。

#### 原文检索:

0. Kilpivaara and L. A. Aaltonen. (2013) Diagnostic Cancer Genome Sequencing and the Contribution of Germline Variants. *Science*, 339:1559-1562.

Howard L. McLeod. (2013) Cancer Pharmacogenomics: Early Promise, But Concerted Effort Needed. *Science*, 339:1563-1566.

Mario L. Suvà,, Nicolo Riggi, Bradley E. Bernstein. (2013) Epigenetic Reprogramming in Cancer. *Science*, 339:1567-1570.

JOCELYN KAISER. (2013) The Downside of Diversity. Science, 339:1543-1545.

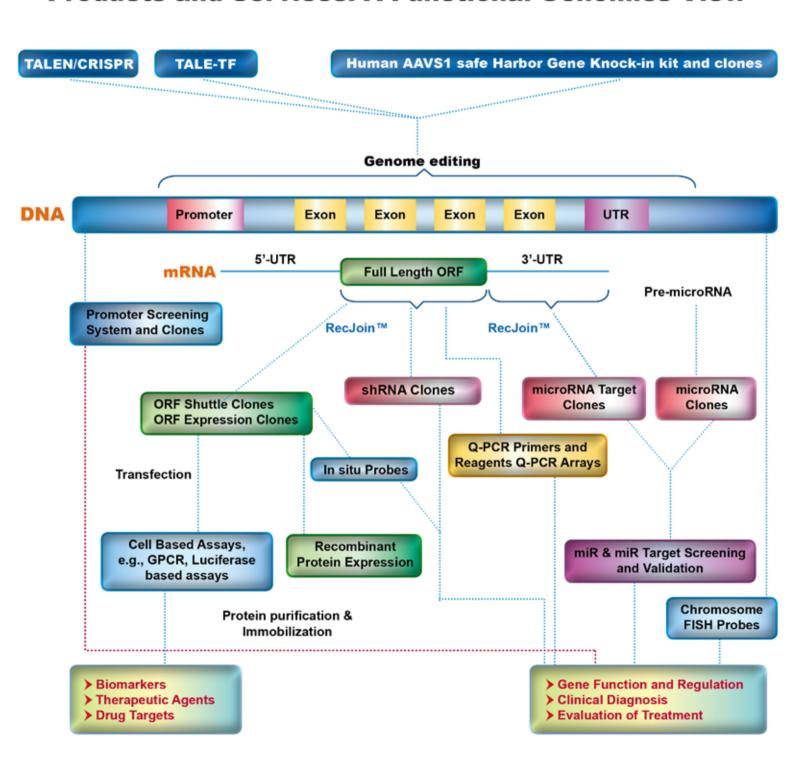
ELIZABETH PENNISI. (2013) Steering Cancer Genomics Into the Fast Lane. *Science*, 339:1540-1542.

YORK、筱玥/编译



## 功能基因组研究线路概览(GeneCopoeia产品与服务)

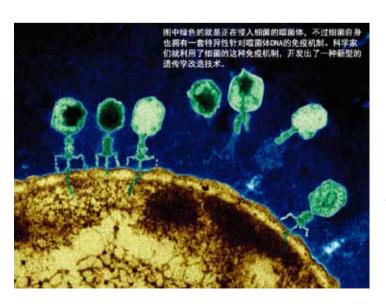
Products and Services: A Functional Genomics View





# CRISPR热潮

科学家们有望利用一种细菌免疫机制开发出革命性的新型基因组编辑技术(genome-editing technique)。



细菌虽然不会对我们真核生物产生太多的同情之心,但是它们也的确是会生病的。而且细菌得病也不是和我们人类毫无关系的,至少就会对乳制品业带来很大的影响,因为在生产酸奶和奶酪的过程中都会用到嗜热链球菌(Streptococcus thermophilus)这样一类细菌。嗜热链球菌能够将奶中的乳糖(lactic acid)。可是噬菌体(bacteriophages,简称

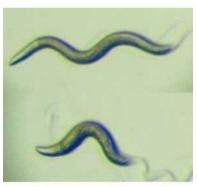
phages)却会让这些细菌"生病",极大地影响酸奶等乳制品的生产质量。

2007年,Danisco这家来自丹麦哥本哈根的食品原料公司(现在该公司已经被DuPont公司收购)的科学家们发现了一种方法,能够极大地增强工业细菌对噬菌体的抵抗力。他们的方法就是用噬菌体来免疫细菌,结果取得了很不错的效果(Science, 23 March 2007, p. 1650)。DuPont公司收购了Danisco公司之后,利用这种技术培育出了抵抗力更强的工业生产用菌。而且这项工作还证明,细菌也是拥有适应性免疫系统(adaptive immune system)的,而且依靠这种免疫机制抵抗同一种噬菌体的多次攻击。

很快, 其他并非从事食品科学和微生物 研究的科研人员也意识到了这种细菌免疫机制 的重要价值,因为他们发现这种细菌免疫机制 有一个特点, 那就是能够针对特定的DNA序 列发挥作用。今年的1月,有四个课题小组都 报道称他们对这种细菌免疫机制进行了成功的 应用尝试——对人类细胞里的特定基因进行了 定向破坏。科学家们将这种技术称作CRISPR 技术。在随后的8个月里,更多的实验室也开 始使用这种技术,对人类细胞、小鼠、大鼠、 斑马鱼、细菌、果蝇、酵母细胞、线虫和多种 农作物进行了大量的试验, 成功地对基因进行 了定向缺失、插入、活化或抑制等遗传改造操 作,从各个方面证明了CRISPR技术的潜在应 用价值。虽然生物学家们最近也开发出了几种 新方法,可以对基因进行更加精确的操作, 但是美国哈佛大学(Harvard University)的 George Church认为,CRISPR技术的高效率 和易用性还是其他技术无法匹敌的。Church 的实验室是世界上首批使用CRISPR技术成功 对人类细胞进行遗传改造的实验室之一。

科学家们可以使用CRISPR技术,用比以往快得多的速度构建人类疾病的小鼠动物模型,也能以更快的速度对基因进行研究,而且还可以一次对细胞内的多个基因进行遗传学改造,研究这些基因之间的相互作用。虽然伴

随着近年来逐渐流行起来的的CRISPR技术狂潮,也慢慢暴露出了这项技术的一些不足和限制,使这股热潮进一步蔓延的速度有所减退,但是Church和其它的CRISPR技术先驱们还是坚信,这项技术有着美好的前景,他们还组建了公司,希望能够利用CRISPR技术治疗各种遗传性疾病。美国蒙大拿州立大学波兹曼分校(Montana State University in Bozeman)的生化学家Blake Wiedenheft就评价道:"我还没有发现在哪个领域内,有哪种技术能够像CRISPR技术这样发展得如此迅猛。"







在短短的8个月之内,科学家们已经利用CRISPR技术成功地对线虫(上图)、斑马鱼胚胎(中图)和果蝇进行了遗传学改造,获得了更粗短的线虫,腹部组织体积更大的斑马鱼胚胎,和眼睛颜色更深的果蝇,这些成果都表明CRISPR技术在动物基因改造方面有着巨大的应用潜力。

#### **【艰难的开始**

这种新技术最早可以追溯到1987年,当 时有一个科研小组观察到,在细菌编码的一个 基因末端有一段非常奇怪的重复序列。不过 当时这个发现并没有引起太多人的关注。直到 十年之后, 越来越多的生物学家们开始从事微 生物基因组的解析工作,这时他们也都发现了 这种奇怪的现象,即在细菌的基因末端,一段 DNA序列会紧接着一段它自己的反向序列, 然后再接一段大约30bp左右的、貌似是由碱 基随机排列而成的DNA序列,科学家们称之 为"空格DNA(spacer DNA)",然后再紧 接着一段与前面DNA序列相同的DNA片段, 之后再接另外一个空格DNA。在一个微生物 的基因组中,这种情况可以出现好多次,不过 每一个奇怪的序列都可以由不同的重复片段 和空格片段组成。在大约40%的细菌和90% 的古细菌(archaea)中都能够观察到这种现 象,于是科学家们给这种序列取了一个名字, 叫作CRISPR,即成簇的、规律间隔的短回文 重复序列(clustered regularly interspaced short palindromic repeats) .

起初,很多科研人员都认为这种CRISPR 序列是毫无价值的垃圾DNA,可是在2005 年,有3个生物信息学课题组报道称,这些 CRISPR序列里的空格DNA总是能够与噬菌 体的DNA序列互补、匹配,这提示我们, CRISPR序列很有可能与细菌的免疫保护机 制相关。"这是一条非常重要的线索。" 美国加州大学伯克利分校(University of California, Berkeley)的生化学家Jennifer Doudna这样评价道。于是,美国马里兰州 美国国家癌症生物技术信息中心(National Center for Biotechnology Information in Bethesda, Maryland) 的Eugene Koonin等人 提出了一个新想法,即细菌和古细菌能够吸收 噬菌体的DNA,然后将其留作己用,并转录 出相应的RNA,与入侵的外源DNA结合,这 有点类似于真核生物采用的RNA干扰(RNA interference, RNAi) 机制。

Danisco公司的研究人员Rodolphe Barrangou、Philippe Horvath等人在2007年发现,他们只要对嗜热链球菌进行一番遗传学改造,插入或者去掉几个与噬菌体序列互补的空格DNA片段,就可以改变细菌对噬菌体的免疫力。不过现在已经来到美国北卡罗莱纳州立大学(North Carolina State University in Raleigh)工作的Barrangou当时并没有充分意识到CRISPR机制巨大的应用潜力,他说道:"我们当时完全没有想到,这些DNA序列可以被用于基因组改造方面的工作。"

目前分别就职于德国HeImholtz感染性疾病研究中心(HeImholtz Centre for Infection Research)和德国HeImholtz医学院(Hannover Medical School in Germany)的Doudna和Emmanuelle Charpentier则分别把工作往前推进了一步。他们俩各自对不同的CRISPR相关蛋白进行了研究,了解了这些蛋白的作用,阐明了DNA空格序列在细菌的免疫防御机制中的具体作用机制。不过他们俩很快就都把目光锁定在了Cas9这种蛋白上,因为依赖Cas9蛋白的CRISPR系统要比其它CRISPR系统更简单。

当有外源的噬菌体入侵时,细菌的CRISPR系统就会被激活,将空格DNA和重复DNA序列转录成一个长链RNA分子,然后细菌会将该长链RNA分子切割成短链的、包含空格DNA序列的小RNA分子(short spacerderived RNA),我们称之为crRNA。Charpentier团队发现,有一种名为tracrRNA的RNA分子与Cas9蛋白共同作用,负责生成crRNA,该成果发表在了2011年的《自然》(Nature)杂志上。Charpentier等人推测,Cas9、tracrRNA和crRNA一起能够降解与crRNA互补的外源DNA分子。

Doudna和Emmanuelle Charpentier都发现, Cas9蛋白是一种能够降解DNA分子的核

酸酶(nuclease),其中含有两个酶切活性位点,每一个位点负责切割DNA双螺旋中的一条链。而且这两个课题组还发现,他们可以在不破坏Cas9蛋白与特异性DNA靶点结合能力的前提下,使Cas9蛋白中的一个、或者两个酶切位点失活,该研究成果对于将CRISPR系统应用于基因组改造工作意义重大。据Doudna回忆,仅仅需要一种酶,再搭配各种RNA就够了,这种技术要比之前的方法简单多了。

不过在CRISPR技术投入使用之前,Doudna和Charpentier的课题组都必须证明,他们能够非常精确地控制Cas9蛋白的切割位点,确保不会发生脱靶的情况。Doudna的博士后Martin Jinek提出了将tracrRNA和空格RNA组合起来,形成一个所谓的"向导RNA分子(single-guide RNA)"的想法,并且他们课题组已经在去年成功地构建出了好几个这种向导RNA,而且将其与Cas9蛋白混合在一起,成功地对特定的DNA位点进行了切割(Science, 17 August 2012, p. 816)。"这篇文章可谓是一篇里程碑式的论文。"Barrangou评价道。

该试验表现出的精准的靶向特性一下子 让众多科研人员对CRISPR技术产生了浓厚的 兴趣。遗传工程师们一直以来都希望能够对

各种生物进行基因插入,或者删除的操作。 但是他们一直无法在基因组中进行精确的定 位, 因此不能在特定的位点进行这种基因改 造。十年前,科学家们开发出了锌指核酸酶 (zinc finger nucleases),该蛋白拥有两个 人工组合在一起的结构域,它可以依靠其中的 DNA结合结构域(DNA-binding domains)结 合到基因组中特定的位点上, 然后再依靠其中 的酶切结构域将该位点的DNA链切断。该技 术诞生之后也掀起了一股热潮,甚至有人成 立了公司, 专门以该技术为平台开发艾滋病 治疗手段(Science, 23 December 2005, p. 1894)。最近又出现了另外一种基因组改造 工具,那就是TALEN人工核酸酶,这是一种 比锌指核酸酶更方便的基因组定向改造工具, 并且有可能会彻底取代锌指核酸酶(Science, 14 December 2012, p. 1408) .

现在,CRISPR系统又大有彻底压倒TALEN核酸酶的趋势。CRISPR系统是以RNA作为基因组定位工具,而锌指核酸酶和TALEN核酸酶则都是以人工开发的特异性DNA位点结合蛋白作为基因组定位工具。我们都知道,合成RNA要比合成蛋白质容易得多,所以Barrangou评价道: "用CRISPR技术只需要几个星期就可以拿到确定的实验结果,可人工核酸酶技术至少需要好几个月。"

## ■ CRISPR技术的优势

方便、快捷并不是CRISPR技术的唯一优势。Church的课题组一直都在致力于将TALEN核酸酶应用于人体细胞的改造工作,但是当他了解到Doudna和Charpentier等人的工作之后,他们立刻尝试了CRISPR技术,构建了一段特异性针对某个位点的RNA分子,尽管他们早已拿到了针对同一位点的TALEN核酸酶。Church等人用三种人类细胞系进行了试验,结果发现CRISPR技术的效率要比TALEN技术高得多,而且采用CRISPR技术

可以对更多的基因进行遗传学改造(Science, 15 February, p. 823)。为了更进一步证明 CRISPR技术的优势,Church课题组甚至构建了一个向导RNA文库,合成了数万条向导 RNA,这些RNA分子能够特异性地与90%的人类基因结合。"有了CRISPR技术,你可以对人类基因组进行任意操作。" Church这样说道。

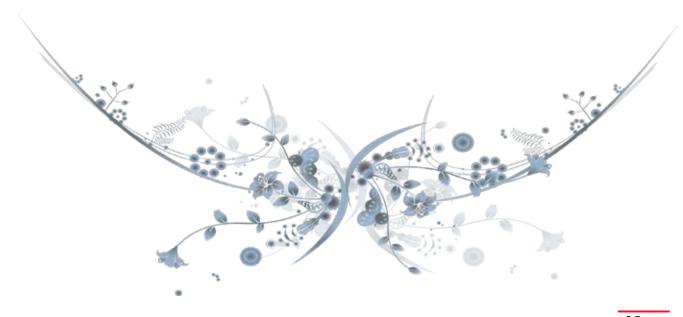
这就意味着,我们几乎可以利用Cas9 蛋白对任何一个基因进行遗传学改造,既可 以利用Cas9蛋白的核酸酶功能将DNA链切断,使基因失活;也可以将DNA链切断之后在该位点插入一个新的基因。美国马萨诸塞州坎布里奇博大研究所(Broad Institute in Cambridge, Massachusetts)的合成生物学家Feng Zhang的文章几乎与Church的论文在同一时间发表。Zhang等人发现,采用CRISPR技术可以同时对人类细胞里的两个基因进行特异性的切除(Science, 15 February, p. 819)。接下来,Zhang又与美国马萨诸塞州坎布里白头研究所(Whitehead Institute for Biomedical Research in Cambridge)的发育生物学家Rudolf Jaenisch合作,一次性破坏了小鼠胚胎干细胞(embryonic stem cell)里的5个基因。

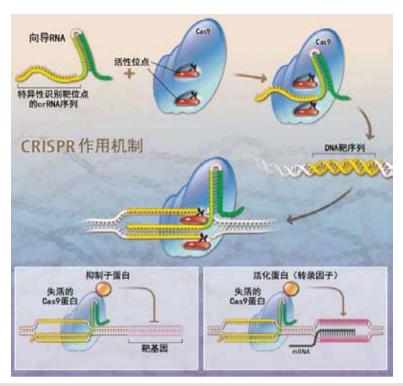
这些工作为将来构建突变小鼠(这是生物学研究工作中急需的科研材料)打下了很好的基础。因为运用此技术可以将经过人工遗传改造过的小鼠胚胎干细胞培育成小鼠胚胎,再继续发育成突变小鼠。但是Zhang等人找到了一种更便捷的方法,他们将编码Cas9蛋白的mRNA和两条向导RNA分子直接注入了小鼠的受精卵细胞当中,同样成功地敲除了两个基因,成功率达到了80%。他们还能够进行更加精细的基因组学改造操作,比如控制Cas9蛋白的酶切活性,使其只能够在DNA链上切开一个小口,而不能将其彻底切断,这样就可

以借助同源重组修复机制(homologydirected repair)在被切开的位点插入一个新的基因,该研究成果已经于今年的5月2日发表在了《细胞》(*Cell* )杂志上。

目前,构建一种小鼠动物病理模型需要经过多轮精心的育种试验,耗时至少1年,可是使用Zhang等人的CRISPR方法,只需要数周的时间就足够了。Zhang认为,这种方法应该不仅限于小鼠这一种动物,对其它动物应该也能取得同样好的结果,他说道:"只要我们能够对动物的胚胎进行遗传学操作,再将改造过的胚胎植入母体内,那就应该就能够得到转基因动物,甚至可能得到转基因的灵长类动物。"

在Zhang和Church的文章上线3周之后,Doudna课题组和一个来自韩国的科研团队也合作发表了一篇文章,Doudna等人成功地用CRISPR技术对人类细胞进行了DNA切割操作。与此同时,另外一个科研小组则成功地用CRISPR技术得到了一种突变的斑马鱼。这一系列的文章吸引了整个生物研究界的注意力。美国杜克大学(Duke University in Durham, North Carolina)的生物医药工程师Charles Gersbach评价道:"如果只是一篇文章,那么可能只会引起一部分人的关注,可是一下子出来了6篇文章,那每个人都会想:'看来我也得做这个了。'"





为DNA"动手术"的外科医生。科学家们发现,仅仅由向导RNA分子和Cas9蛋白组成的CRISPR系统就可以对基因组中特定的DNA序列进行外科手术式的精细改造,在特定的位置敲除、或者插入基因。最近还有人发现,对Cas9蛋白进行修饰之后还可以对特定的基因进行抑制(如左下图所示)或者激活(如右下图所示)操作。

中国科学院北京遗传及发育生物学研究所(Chinese Academy of Sciences' Institute of Genetics and Developmental Biology in Beijing)的Gao Caixia也是在一年前看过Doudna和Charpentier等人的文章之后坚定地投靠了CRISPR系统。之前,Gao的团队一直是在使用锌指核酸酶和TALEN核酸酶对水稻和小麦等农作物进行遗传学改造工作。现在由于使用了CRISPR技术,她们已经成功地让四种水稻基因失活了,这意味着CRISPR这

种技术也可以用于对水稻这种至关重要的粮食作物进行遗传学改造的工作。Gao等人还用CRISPR技术敲掉了一个小麦基因,结果得到了耐白粉病(powdery mildew)的小麦新品种。Gao等人于今年的8月在《自然 生物技术》(Nature Biotechnology)杂志上发表了文章,介绍了她们用CRISPR技术对植物进行遗传学改造的成果,与此同时,也有其他4篇文章得以发表,分别介绍了用CRISPR技术对植物和大鼠进行遗传学改造的工作。



上图展示的是经CRISPR技术改造过的水稻。就在本月(2013年8月)早些时候,科学家们成功地用CRISPR技术对水稻等植物进行了遗传学改造,图右展示的就是基因被敲除之后的水稻,可见某个基因被敲除之后,水稻变得矮小、苍白。

此外,使用CRISPR技术还非常便宜。 比如针对特定基因设计向导RNA的软件就是 免费的;在美国马萨诸塞州的坎布里奇,还 有一个叫做Addgene的机构,他们向学术科 研机构提供用于构建CRISPR系统的DNA分 子,只收费65美元,因为有11个科研团体向 Addgene机构提供可用于构建CRISPR系统 的DNA序列。自今年初至今,Addgene已经向全世界提供了5000多个CRISPR序列,而且仅仅在7月的一个星期之内,他们就收到了100多份订单。Addgene的执行负责人Joanne Kamens说道: "CRISPR系统现在是最火爆的科研技术。"

#### 基因活性调控功能

最初发表的有关CRISPR系统的文章,介绍的全都是如何利用CRISPR系统切割DNA的工作。但是很快,科研人员们就发现了CRISPR系统的新功能。美国加州大学旧金山分校(UC San Francisco)的Lei S. Qi等人与Doudna合作,发明了类似于RNAi技术的CRISPRi技术,能够以可逆的方式使基因失活,这种技术在研究基因功能方面潜力巨大。Qi等人主要对Cas9蛋白进行了改造,这些Cas9蛋白同样能够在向导RNA的带领下与特定的DNA序列结合,但是不会切断DNA链。在细菌细胞里,Cas9蛋白就足以抑制基因的转录,但是在哺乳动物细胞内,Qi等人又在

Cas9蛋白上额外添加了一个抑制子,用来抑制基因的活性。此时的向导RNA主要与靶基因上游的调控序列——启动子区域结合。

就在上个月,Qi团队和另外三个科研小组使用携带了一个人工合成的转录因子(能够启动基因表达的蛋白)的Cas9蛋白,成功地激活了人体细胞基因组中下游靶基因的表达。他们发现,使用1种CRISPR的效率比较低,不过他们已经想到了解决这个问题的办法。据这四个课题组负责人之一的Gersbach介绍,他们针对启动子区域里的不同位点,设计了多个向导RNA分子,结果就看到了非常明显的协同效应。

"我还没有发现在哪个领域内,有哪种技术能够像CRISPR技术这样发展得如此迅猛。"

——美国蒙大拿州立大学Blake Wiedenheft

Gersbach在今年7月25日出版的《自然方法》(Nature Methods)杂志上发表了一篇文章,文章介绍了他们激活的多个与多种人类疾病相关的基因,包括与肌肉细胞分化、肿瘤与炎症,以及胎儿血红蛋白生成作用相关的基因。另外两个课题组也对生物医学意义重大的多个基因进行了遗传学操作。据Gersbach介绍,利用CRISPR技术对这些基因进行改造,可以治疗镰状细胞贫血、关节炎等多种疾病。

当然, CRISPR技术也不是尽善尽美的,

也存在一定的局限和缺陷。比如我们现在就不清楚向导RNA分子的特异性作用机制。据美国波士顿市麻省总医院(Massachusetts General Hospital in Boston)的J. Keith Joung介绍,他们的原始试验数据显示,CRISPR技术也存在明显的脱靶效应(offtarget effect)。Joung在今年的1月也发表论文介绍了利用CRISPR技术对斑马鱼胚胎进行基因改造,并用CRISPR技术活化基因表达的工作。Joung的工作显示,与向导RNA序列类似的非靶点DNA片段也能够被切断、激活或

者失活,即存在明显的脱靶效应。

Joung的科研团队在实验中发现,向导RNA分子还能够与靶点序列不同的非靶点序列结合,非靶点序列与靶点序列的碱基差异可以达到5个之多。虽然Zhang等人的结果稍好一点,但是他也认为,特异性问题是一个必须解决的问题,尤其是在如今大家都迫切地希望用CRISPR技术治疗人类疾病的大背景之下。Zhang表示,如果要确保CRISPR技术的安全性,就必须保证CRISPR分子只能够与目的序列结合,绝不能与非目标分子结合。

同样地,我们也必须保证CRISPR系统的组份能准确地到达目标位置。Joung强调说: "CRISPR系统的输送问题也是一大难题,这应该与被作用的细胞,或者作用物种有关。"比如对于他们使用的实验对象斑马鱼,就采用了直接在斑马鱼的胚胎中注入向导RNA和编码Cas9蛋白的mRNA的方法。可是对于哺乳动物细胞,他们使用的则是DNA分子。如何让CRISPR系统进入成体动物,或者用于治疗身患疾病的人类患者,这还都是需要仔细思考

的问题。

最终,根据实际遗传学操作的需要,CRISPR技术和锌指核酸酶,以及TALEN核酸酶都会找到属于它们自己的一席之地。但是目前,大家都还是被CRISPR技术的简便、易用特性,以及无限的应用潜力所吸引。Charpentier等人正在其它细菌中寻找Cas9蛋白的变异体,希望找到一种更好用的Cas9蛋白。微生物学家们也在利用CRISPR系统对细菌进行改造,以抑制耐药基因的进一步扩散。Church、Doudna、Charpentier,以及很多科研人员也都成立了与CRISPR技术相关的公司,探索CRISPR技术在包括基因疗法在内的治疗人类疾病方面的价值。

除此之外,CRISPR技术还有很多其它非常有价值的应用方向。Barrangou就认为,目前CRISPR技术面临的瓶颈就是人类的想象力太有限了。

细菌受感染之后启动的这套机制还真不赖。

原文检索:

ELIZABETH PENNISI. (2013) The CRISPR Craze. Science, 341:833-836.

YORK/编译



# Gene and miRNA qPCR Arrays

GeneCopoeia提供的qPCR检测阵列是基于SYBR® Green I染料法的定量检测技术,可用于基因表达差异研究,具有易操作和可靠等优点。该系列产品有96和384孔板两种规格,专为检测不同组织或细胞中特定信号通路或疾病相关的基因(miRNA)的表达量而设计。检测结果所显示出来的表达差异,可帮助研究者获取关于这些基因的进一步研究线索。



#### 特点与优势:

验证引物: 每对引物均使用专利算法设计并经过实验确证:

性能优越:严谨的质量监控保证产品的高品质、特异性好及高灵敏性(最低可检测4个mRNA分子); **覆盖范围广**:目录产品包含通路分析、癌症和其他研究热点、定制产品可根据客户要求选择设置。

类别	产品名称	描述	配套产品
基因表达量检测阵列	ExProfile™ Cancer Gene qPCR Arrays	癌症相关基因表达 量检测阵列	RNAzol® RT RNA Isolation Kit All-in-One™ First-Strand cDNA Synthesis Kit All-in-One™ qPCR Mix All-in-One™ gene qPCR Validated Primers
	ExProfile™ Pathway- Focused Gene qPCR Arrays	信号通路相关表达 量检测阵列	
miRNA表达量 检测阵列	miProfile™ miRNome qPCR Arrays	覆盖miRBase V19.0	RNAzol® RT RNA Isolation Kit All-in-One™ qPCR Mix All-in-One™ miRNA First-Strand cDNA Synthesis Kit All-in-One™ miRNA qPCR Validated Primers
	miProfile™ Cancer miRNA qPCR Arrays	人与鼠癌症相关 miRNA表达量检测 阵列	
	miProfile™ Disease and Focus-Group miRNA qPCR Arrays	疾病或特定群组相 关的miRNA表达量 检测阵列	
表达量检测阵列 的定制服务	Custom-made Gene or miRNA qPCR Arrays	有96和384孔板供 选择 阵列定制和qPCR检 测服务	
数据分析工具	Online data analysis tool	完全免费,且操作 简便	



## 雄性哺乳动物为何选择一夫一妻制?

迄今为止,社会性单配偶制(即一夫一妻制)的进化现象已经在多种哺乳动物中数次产生,这是雄性动物的交配策略,原因在于这些物种的雌性动物因为生态方面的因素而呈广泛分布。

长久以来,单配偶制(即一夫一妻制) 一直吸引着科学家和普通大众的注意。问题 的关键在于,这种配偶制竟然发生在多种雄 性哺乳动物身上。对此,进化生物学家感到 十分不解。显然,雄性哺乳动物在单位时间 内繁殖后代的潜力要比雌性高得多。鉴于 此,人们必然要弄清楚这种做法的选择性优 势所在,这种优势应该足以弥补雄性因约 束自己,只与一位雌性进行繁殖活动所导 致的潜在生殖力丧失的后果。在本期杂志 (Science Vol 341)的第526页,Lukas和 Clutton-Brock两名研究者为这个问题作出了 证明:雄性哺乳动物并非通过父方养育后代产生的合理受益来弥补其接受单配偶制带来的代价,而是通过配成一对配偶的方式来解决由于生态因素而导致的雌性个体间分布距离远的不足。

在鸟类、某些新鳍类鱼、无尾类动物以及节肢动物之中,父方对后代的直接养育是可以实现的,并且这种做法增加了其后代的生存率。因此,自然选择青睐于这种促进单配偶制产生的雄性繁殖策略。但是,哺乳动物的情况却截然不同,它们的体内妊娠和随之而来的哺乳等必需的亲代养育方式都只局

限于雌性;相对于雄性而言,这种生理性约束 就减少了雌性可能繁殖后代的几率。因此,若 雄性哺乳动物只与一位雌性结为伴侣,那么它 就放弃了额外的配对机会,从理论上说这不利 于它们的繁殖。由此可见,单配偶制能够成为 雄性哺乳动物繁殖的策略,这的确需要解释。

当然,我们可以从雌性方面进行说明。雌性遵从它们自己的繁殖策略,这一点并不受其配偶数量的限制,而是受到资源获取以及父方养育的可能性的限制。所以,动物的饮食特点以及其它生态学因素决定了雌性哺乳动物的空间分布(图1)。根据动物群居的获益和成本之间要达成的平衡这一自然规则,雌性要么聚

居,要么独居。而无论是哪一种方式,个体家 庭的生活范围都可能会与周边的雌性动物产生 重叠。

因此,雌性哺乳动物为它们的种群设立了 基本的规则,而雄性则按照雌性的分布进行匹配。由于配偶数量限制了雄性繁殖的成功率, 因此它们对雌性的空间分布和雌性的性接受期 的同步程度较为敏感。这样,当雄性哺乳动物 面对独居的雌性时,它有两种选择:一是广泛 漫游,尽力覆盖数只雌性动物所处的范围,并 护卫它们的安全;二是只结交一位雌性,并与 其生活,成为一对配偶。



图1 哺乳动物的社会进化。食物来源的分布和数量、天敌以及病原体(左图)决定某种哺乳动物的雌性是独居还是群居。而雄性的社交策略则由随后的下一步情况决定,比如杀婴行为的风险、双亲养育的选择、性内抉择以及接受同类雌性的时空分布。Lukas和Clutton-Brock证明,在以上因素中,雌性的时空分布决定了哺乳动物形成一对一的生活方式。当独居的雌性聚集或群居时,则进化出其它的社会组织方式。

从雄性的角度来看,这儿产生了一个重要的问题:到底是由于雌性的空间分布迫使雄性接受单个配偶的生活呢,还是自然选择无视雄性丧失交配机会的代价而独青睐于这种方式?在此之前,有许多设法解释哺乳动物单配偶制进化现象的假说,可惜都被一一驳倒,其余的假说则被证明更符合鸟类或其它门类的生物。目前,针对雌性空间分布和父方养育的假说也出现了一个反证。而且,此前的大多数研究都

假定单配偶生活的动物始祖都是独居的,但这 并不属实。最近一项关于动物原始社会进化的 研究显示,它们的始祖过的是群居生活。

为了解决这些问题,Lukas和Clutton-Brock从所有的生物种类中搜集了2500多种哺乳动物的社会生态学和生命史数据。他们利用种系发生学的重构法,确认了61种自主转化成为社会性单配偶制的哺乳动物。除了其中一种之外,其余动物都是从雌性独居的生物原种

进化而来。并且,虽然现今大多数属于社会性 单配偶制的动物都以父方养育为特点,但他们 的分析有力表明,父方养育是在单配偶制形成 之后才进化出来的,这就排除了父方养育是哺 乳动物选择进化为单配偶制的主要自然选择动 力的说法。这两位研究者的分析还排除了雄性 为降低后代的风险而产生杀婴行为是促进单配 偶制进化的主要驱动力的可能,这样,对于雄 性以保护妻子和幼儿的形式来完成其"父方投 资"的假说也就不攻自破了。

研究者还证明,与雌性独居的种群相比,现今仍然选择成对生活的动物物种均以较低的分布密度出现,它们与其近邻产生重叠的生活领域也相对较小。因此,成对生活的的雌性动物很可能由于生态方面的原因而分布广泛,当然也有可能是因为它们经历了强烈的雌性哺育竞争而导致了现在的结果。那么可以推测:在哺乳动物形成配对生活的初期,雄性很可能由于能力所限,无法独占空间分布广泛的多个雌性,因此被进入了这种社交组织方式之中;而父方养育和杀婴行为式的保护所带来的益处只

不过是顺水推舟地稳固其习性罢了。

这样, 据此惊人的样本规模和清晰的阐 述,Lukas和Clutton-Brock得以解释此前研究 所产生的各种相互矛盾的结果。他们的研究还 为社会生态学模型(一种为进化性社会系统研 究服务的理论框架,一直遭受批评)提供了强 大的支持。但是,这项研究也提出了一些新的 问题。比如说,为何单配偶制在哺乳动物之中 分布得如此不均?在大多数种群中未出现单配 偶制,这能否仅仅通过(大多未知的)饮食适 应性来获得解释?为何某些种群的成对伴侣一 整年都亲密地呆在一起,而有些种群却一年才 见一次面,就为了进行交配?而在不同种群之 中,与配偶之外的异性交配或者同种动物内社 交组织的变异有多常见? 而这些社交发生的结 果会是什么?上述问题之中,有些也适用于对 人类的研究,尽管人类的单配偶制并非起源于 独居的祖先, 而是还有其他的因素。但是, 也 正是因为如此, 我们必须寻找出单配偶制产生 的原因。

原文检索:

Peter M. Kappeler. (2013) Why Male Mammals Are Monogamous. Science, 341, 469-470.



## 深潜还是浅潜? 控制氧气水平即可

先问你一个小问题:你能在水下憋气多久?不管你的肺功能有多棒,肯定都比不上加利福尼亚海狮。它们在捕食的时候,能够潜水憋气长达10分钟!这些家伙是怎么做到的呢?为了弄清这个问题,美国斯克里普斯海洋研究所(Scripps Institution of Oceanography)的Birgitte McDonald决定测

量它们静脉血的血氧饱和度。对此,她解释说,在静息时,静脉血在为身体供氧之后,只能达到大约78%的血氧饱和度。由此你可以想象一下,在潜水这种缺氧的状况下,血氧饱和度会下降得多么厉害!或许通过这种监测,就可以发现海狮潜水的秘密了。



当然,McDonald无法单独开展这项研究,在其博士后导师Paul Ponganis的帮助下,她在加利福尼亚海岸捕捉了几只哺乳期的海狮,并把它们麻醉,然后将电极插入其静脉血管腔中,从而得以测量海狮体内的氧气局部压力(这是测量氧气水平的一种方法)。在为它们安装了测量这种氧气局部压力的数据记录器以及一台实时深度记录仪之后,McDonald把这些海狮放回了自然。

实验记录完成之后,McDonald取回了那些设备,并将所有的测量结果转换为血氧饱和度百分比。她发现,海狮在浅水和深水的潜游中分别有着完全不同的耗氧方式。在持续不超过3分钟的浅潜之中,海狮的血氧饱和度可下降到10%-80%区间内的任一水平。而在深潜中,海狮并不总会使血氧储存补充回78%的水平,某些情况下,其潜水时的血氧饱和度甚

至低至**15%**,这或许解释了某些海狮似乎只有非常低的血氧饱和度水平的现象。

McDonald还发现,在开始深潜的时候,海狮的静脉血血氧饱和度会比浅潜时增加更多。有时候其水平能达到95%。对此,她解释说,这很可能是由于出现了动脉-静脉血分流,使得来自肺部的新鲜饱和动脉血流经其体内各组织并氧化静脉血的缘故。而通过分流这种动脉血,它们就能够在潜水之前为自身提供最大程度的血氧储存了。在潜水的时候,会产生一种几乎完全性的氧耗。而当再次上浮之后,海狮会在水面停留较长的时间,以补充氧气储存,这一点与它们浅潜之后的做法完全不同。总之,McDonald的数据表明,海狮知道自己何时深潜,何时浅潜,并随之改变其控制氧气的策略。

#### 原文检索:

McDonald, B. I. and Ponganis, P. J. (2013). Insights from venous oxygen profiles; oxygen utilization and management in diving California sea lions. J. Exp. Biol. 216, 3332-3341.







# 慢病毒完整解决方案



40,000个现货克隆,即选即包,完全免费!



ORF表达克隆的病毒低至8000元;miRNA、shRNA克隆的病毒低至3000元!



滴度高达1010copies/mL!



最快20个工作日内送货!



已发表几十篇高分文章,详情见官网。

活动时间: 6月13日-12月31日

注: 以上荧光图为真实实验结果图



请致电(020)32051255

www.LifeOmics.com