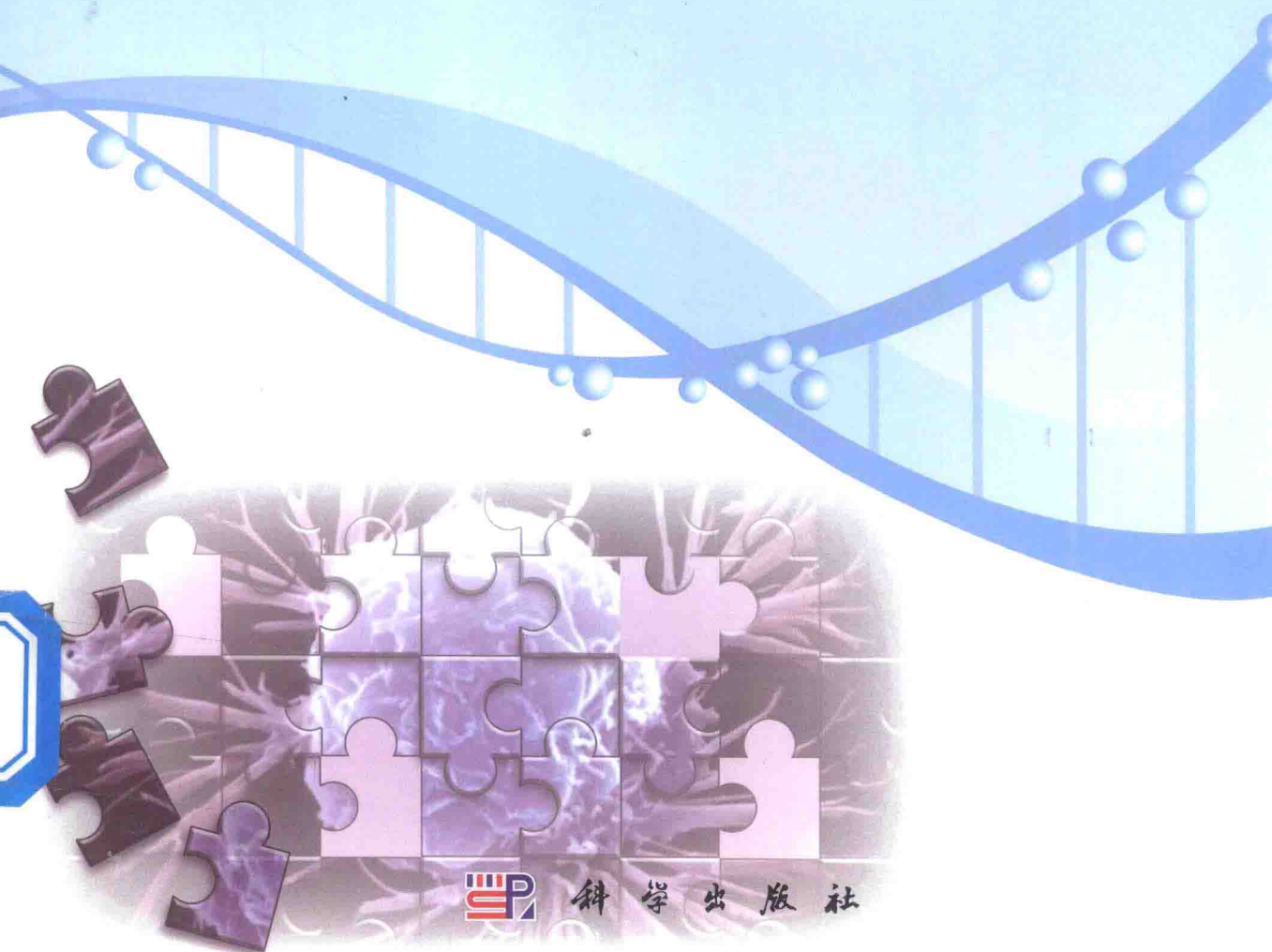


CANCER BIOINFORMATICS

肿瘤信息学

[美] 徐 鹰 崔 娟 D. 普特 著
王 岩 杜 伟 王 浩 孙文靖 董 宁 译



科学出版社

肿瘤信息学

Cancer Bioinformatics

〔美〕徐鹰 崔娟 D.普特 著
王岩 杜伟 王浩 孙文靖 董宁 译

科学出版社

北 京

图字: 01-2015-2057 号

内 容 简 介

深入解读癌症本质,是人类在新世纪面临的重大挑战。本书从生物信息学角度,利用现有多种组学数据对癌症的生物学特征予以描述归纳及探讨,指出癌症是一个进化复合体系统,要以“整体论”来思考研究,并提供了理论框架。

全书分为14章,探讨了癌症的生物学基础、利用组学数据从分子及基因组层面理解癌症发生发展的驱动力、癌细胞在多重压力下存活的路径、表观遗传应答的作用、癌症侵袭和转移的原因、癌症转移后的特征等。此外,还示范了癌症标志物的发掘以及利用计算手段研究癌症的策略等。

希望本书能部分地弥合实验癌症生物学家和计算生物学家之间的沟壑;也希望基础医学、临床医学、生物信息学的研究者们能掌握独特的思维方式,从而加速各自的相关研究。

Translation from English language edition: *Cancer Bioinformatics* by Ying Xu, Juan Cui and David Puett.

Copyright ©Springer Science+Business Media New York 2014.

All Rights Reserved.

本书彩图可扫描书后二维码

图书在版编目(CIP)数据

肿瘤信息学/(美)徐鹰等著;王岩等译. —北京:科学出版社,2016.6

书名原文: *Cancer Bioinformatics*

ISBN 978-7-03-047555-8

I. ①肿… II. ①徐…②王… III. ①肿瘤学—研究 IV. ①R73

中国版本图书馆 CIP 数据核字 (2016) 第 044415 号

责任编辑:李悦 李晶晶 / 责任校对:蒋萍

责任印制:张伟 / 封面设计:北京铭轩堂设计公司

科学出版社出版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

北京京华虎彩印刷有限公司印刷

科学出版社发行 各地新华书店经销

*

2016年6月第一版 开本:787×1092 1/16

2016年6月第一次印刷 印张:19 1/2

字数:465 000

定价:118.00元

(如有印装质量问题,我社负责调换)

致 谢

本书的出版与很多人的勤奋工作及鼎力支持是密不可分的。感谢徐鹰教授给予我们翻译出版这本重要著作的机会，还要感谢他在百忙之中的多次悉心指导，以及持续地鼓励与启发。

感谢梁艳春教授在本书翻译过程中提供的中肯意见。本书从系统生物学的高度，探讨癌症这一复杂命题，涉及了大量的跨学科、跨领域知识，给我们这个翻译小组带来了严峻的考验。所幸之处在于，来自吉林大学、哈尔滨医科大学、美国佐治亚大学等多所国内外院校的生物学、医学、计算机、数学等不同领域的多位专家同道，组成了我们背后强大的支持团队。他们无私的支持与倾力的合作，让本书更加严谨。我们深深地感谢这些令人尊敬的老师及同仁，国外专家包括：崔娟、杨青、王鹏、张驰、曹莎等；国内专家包括：崔久嵬、王林聪、刘琦、王国庆、周丰丰、黄艳新、唐彤宇、张灵、孙连坤，管仁初、俞琼、李瑛、吴春国、苏静、刘丙强、胡良海等。同时，感谢包括孙莹、孙慧妍、欧歌、吕一丹、谭永飞、杨晓春等多位同学，认真阅读译文初稿并提出了宝贵意见和建议。

真诚地感谢科学出版社的李悦等编辑，在本书出版过程中进行了大量庞杂而细致入微的校对工作，他们良好的专业修养以及高效的工作也为本书的出版提供了有力保障。

最后，感谢所有在本书的撰写、翻译、修订过程中付出辛勤劳动的学者、同学以及朋友们，没有你们的帮助和共同协作，本书的翻译出版不会如此顺利的进行。

本书翻译得到了国家自然科学基金面上项目的资助（项目编号 61572227、61472159）。

王岩（吉林大学计算机与技术学院）

杜伟（吉林大学计算机与技术学院）

王浩（哈尔滨医科大学附属第二医院）

孙文靖（哈尔滨医科大学基础医学院）

董宁（吉林大学白求恩第一医院）

译者序

诸如“生物信息学”、“数据分析”、“二代测序”，以及“RNA-seq”等概念显然已经为生物学、医学和计算机科学等诸多领域带来了深远的影响。在“大数据时代”到来的今天，这些“高、大、上”的新理念、新技术、新方法到底会为人类带来哪些福音，又会如何深刻影响人们的生活呢？美国佐治亚大学生物化学及分子生物学系的徐鹰教授（Prof. YingXu）以其独到的视角，结合自己近年来的最新研究成果，为我们了解并学习“生物信息学”和“肿瘤信息学”指明了一条捷径。

本书名为 *Cancer Bioinformatics*，是以肿瘤信息学研究为中心，主要内容围绕肿瘤信息学的题目展开，并以相关的生物信息学技术贯穿其间。书中的一些精彩部分也包含于徐鹰教授在吉林大学主讲的由中国科学院资助的“龙星计划”《肿瘤信息学》课程中。作为吉林大学计算机学院的长江学者讲座教授，“千人计划”讲座教授，国际著名的计算生物学家、生物信息学专家，徐鹰老师长期致力于癌症生物信息学、微生物信息学和结构生物信息学等相关领域的研究。他以数理统计、机器学习理论、信息技术为基础，将建模、算法创新等技术手段应用到生物信息学的多个领域，为推动生物信息学和系统生物学的快速发展作出了重要贡献。多年来，他还始终致力于生物信息学研究在国内的传播，培养了众多的优秀人才。

作为徐鹰教授的学生，我们很荣幸有机会为传播徐老师的理论贡献绵薄之力。我们这个翻译小团队主要包括吉林大学计算机学院长期从事生物信息学研究的王岩、杜伟，哈尔滨医科大学长期从事恶性肿瘤临床治疗与基础研究的王浩、孙文靖，以及吉林大学白求恩第一医院的董宁等人。同时，来自于吉林大学、哈尔滨医科大学、佐治亚大学等多所国内外院校的 20 余位知名教授、专家、学者以及同学都在翻译的不同阶段提供了巨大的支持与帮助。

本书的英文版由 Springer 出版社于 2014 年 8 月正式出版，而中文版的翻译前期工作也随即逐步展开。然而，囿于本书涉猎内容广泛，加之译者们的能力所限，本书的翻译工作持续了一年多的时间。所幸的是，徐老师具有前瞻性的理论正在不断地获得来自学术界的印证，这也恰好为读者们提供了思考比对的空间。

本书最初是希望能够帮助计算生物学、生物信息学的学生快速进入肿瘤信息学领域。出乎意料的是，来自于肿瘤临床与基础研究一线的反响也异乎寻常的积极。所以，我们也真诚地希望能有更多来自于医学领域的读者可以从中获益。

本书所涉及的数据分析流程和技术，对于具有数理基础和计算机基础的同学而言，不会带来太多阅读和理解的障碍；而对于缺乏此类背景知识的医学领域的读者来说，也无需为此担心，因为你们的生物医学背景甚至会有助于更好地理解作者的本意。

与国内外众多优秀的生物信息学著作或译作不同，本书作者更多地把重点放在了数据分析结果的生物学意义上，强调了对数据的生物学解读：癌症的基本生物学特征；原癌基因、癌基因与抑癌基因的作用；影响癌症发生发展过程中的重要因素，如微环境、免疫系统、表观遗传组改变、代谢组改变等。更为简化的概括是：要以持续进化的观点

理解并认识癌症，能量代谢改变是核心，癌症的异质性是必然，转移后的癌症反而具有了特殊的共性可资利用；癌症研究要恰当地结合整体论与还原论。

与常见的肿瘤研究相关专著相比，徐鹰教授提出的理论与假说更多地建立在组学数据分析的基础上，这些数据基本上来源于已经公开发表的各种肿瘤及对照组的标本，感兴趣的读者可以通过 GEO 和 TCGA 等相应的数据库自行验证。书中的一些设想与提问同样也可以为有志深入研究的读者提供方向。

能以此方式进行写作的最根本原因在于，作者是从组织学水平而不仅仅是细胞学水平甚至分子水平来看待肿瘤。在当前的肿瘤研究领域，更多的研究结论来源于以细胞系为基础的实验室研究，相应的方法论基础是经典的还原论。这就要求研究者把待研究问题与更多的影响因素——分离，看似条分缕析，却导致相应的结论经常与客观现象不一致，相似研究的结果也常常不尽相同。

如何强调整体论的观点都不为过。然而，否定一切还原论指导下的研究结果显然不是徐鹰老师的目的。正是因为有了当代的组学技术和数据分析，我们才开始有机会站在这样的高度。分析肿瘤组织标本为来源的数据本身就把研究者放在了更高一级的水平上，当然要使用另外的方法论。当然还原论也同时意味着放弃了已有的人类对肿瘤认识的基础。

本书中的研究探讨都有大量的来自各个领域的文献作为支持信息，这也反映了当代肿瘤生物学研究和生物信息研究的基本要求。相信读者会因此而受到启发，组建属于自己的多学科研究团队，提出自己的观点，并加以验证。您刚开始阅读时，可能会被一系列的专有名词所困扰；不要担心，因为即使是作为译者的我们也曾有这样的感受。作为将学习整体论的您，就当是一种小小的练习吧。当您读到之后的章节可能会顿悟，原来如此！当然，您也可能会对其中的观点产生怀疑，我们也愿意和您一起探讨，争取能通过深入的研究解决有意义的问题。书中的观点仅代表作者们在当时的思考，我们相信，恰当地结合转录组数据与蛋白质组数据分析并结合扎实的实验室验证必将成为未来研究的立足点。在此，我们真诚地鼓励有兴趣的读者能大胆地进入肿瘤信息学，为人类应对癌症作出自己的贡献。徐老师本人也非常欢迎学术上的交流（xyn@bmb.uga.edu）。

我们真诚地期待您能以轻松的心态阅读本书，就像是阅读一本高级的科普读物。您会发现，不知不觉中您了解了人类认识肿瘤的上百年历史，并会渐渐地提出各种有趣的问题。请不要为某个基因的具体细节而执迷，也不需为所涉及的广泛知识领域而恐惧。阅读过英文版原著的一位普通外科医生朋友曾这样描述：“我用了三天时间阅读，非常有意思，有点儿上瘾了！”希望您阅读翻译版时也有类似的感受。

在本书的翻译过程中，我们力求忠实、准确地表达原著的内容，也曾力求保留原著的语言风格。由于水平有限，尤其是本书内容涉及与化学、生物、医学、计算机等多门学科的交叉，因此难免会有错误和不准确的地方，敬请广大同行及读者不吝指教。同时，译者根据自己的理解对原著中的部分内容添加了译注，不当之处也请读者多多指正。

王岩 杜伟 王浩 孙文靖 董宁

2015 年 11 月

前言

2010年，普利策奖得主慕克吉（Mukherjee）在其获奖作品《疾病之王：癌症传记》中曾经提及：人类历史上对癌症最早的记载可追溯至约4500年前，即公元前2500年。杰出的埃及医生伊姆霍特普（Imhotep）清晰而简洁地记述了一例乳腺癌病例^①。两千多年以后（约公元前400年），希腊医生希波克拉底（Hippocrates）将此疾病命名为karkinos（希腊语螃蟹之意，与巨蟹座同词），并逐渐演化为现在的癌症（cancer）一词。大约五、六个世纪后（公元130~200年），受希波克拉底学派影响的希腊医生克劳迪阿斯·盖伦（Claudius Galen）将癌症的病因归结为黑色胆汁过多（该学派认为人体是由四种体液组成：血液、黏液、黄胆汁和黑胆汁）。又经过了上千年，维萨里（Vesalius，16世纪）和贝利（Baillie，18世纪）才推翻了黑胆汁假说。这也间接地推动了外科医生们开始切除实体肿瘤（实际上无畏的外科医生们在此之前就已经开始对癌症患者进行外科手术治疗，但罕有历经磨难与感染的患者能存活下来）。直到19~20世纪，随着麻醉学的发展、抗生素的广泛应用，以及洁净的手术环境的出现，才快速推动了外科手术的发展（后来又出现了放射疗法，俗称“放疗”），使之逐渐成为治疗癌症的主要手段，并且沿用至今。直到20世纪中叶，化学疗法（俗称“化疗”）和激素疗法才作为外科手术治疗和“放疗”的补充或替代手段开始应用。

随着癌症治疗手段的进步，生物医学研究者们也开始努力探寻对癌症本质更深层次的认识，探索更有效的方法来阻止癌症的恶化甚至实现彻底根治。在过去的几十年中，人们提出了大量关于驱动和促进癌症发生、发展以及转移的现代生物学理论，并用于指导癌症研究。这些理论中，也包括了Otto Warburg提出的极具洞察力的理论。他认为，导致癌症的根本原因在于能量代谢方式的改变。他在20世纪60年代还曾指出：“癌症……具有无数的继发性因素，但是……只有一个根本性原因，即机体正常细胞的有氧呼吸被糖酵解所取代。”^②40年前，两组杰出的科学家分别发现了第一个原癌基因和抑癌基因，自此之后，癌症研究界的主流观点逐渐开始将癌症当成一类基因组疾病。

随着多种组学数据源的快速扩张，如基因组、表观遗传组、转录组、代谢组和药理学组学数据等，过去的20多年里，我们在对癌症的理解方面已经取得了卓越成就，尤其是在理解微环境和免疫系统如何促进肿瘤发生发展等方面。然而，即使是有了这些长足的进步，仍有大量复杂的问题有待解答。本书的作者们深信，针对解答这些问题所需要的信息可能早已存在于目前的组学数据中，并且仍有大量类似的数据实际上并没有得到充分地挖掘和利用。我们认为，出现这种情况的原因之一在于，计算癌症生物学家作为一个群体，尚未学会以独立的方式来思考癌症信息学，而只有这种独立的思考方式才能使他们充分利用隐含在组学数据中的全部信息。目前，计算癌症生物学研究几乎都在遵

① Mukherjee, S. (2010). The emperor of all maladies: a biography of cancer, Scribner.

② Warburg, O. H. (1969). The prime cause and prevention of cancer, K. Triltsch.

循着类似于实验性研究者的思维方式，即一种高度还原论的思想。遗憾的是，有关癌症的很多根本问题之间，在本质上是彼此紧密关联且不可分割的。例如，很多癌症研究没有考虑到相关的细胞外基质和免疫应答，而这二者都与癌症密不可分。如今，面对着现有的组学数据，以一种崭新而全面的思维方式来解析势在必行，这样才能使人们对癌症生物学的理解提高到一个新的水平；换言之，我们可以在更复杂的背景下研究癌症。

本书目的是从信息学角度对癌症生物学进行宏观描述，同时示范如何挖掘组学数据来获得新的视角和新的理解，从而能更广泛并更本质地对癌症的生物学问题加以探讨和解答。纵观全书，作者力图逐步确立以下述及的主要观点：①癌症，就其本质而言，是病变细胞在微环境压力不断增加且越加难以应对的条件下得以存活的持续进化过程，而微环境本身是与病变细胞共同进化的；②细胞增殖是癌症克服生存压力得以存活的一种方式；③进化中的细胞必须克服的主要挑战来自于周围的组织，所以癌症更主要是组织水平的问题，而不仅仅是细胞水平的问题；④对具体癌症而言，其生存通路的选择，往往不取决于功能异常的基因或基因突变，而在很大程度上取决于人类基因组中那些规模较大的编码序列，且其功能原本另有他用；⑤癌细胞的不同亚群为启动存活通路创造了条件；⑥在散发癌症中，基因突变的主要功能意义可能在于以更高的效率、更好的延续性替代了现有的功能；与之相对的是，在遗传性癌症中，基因突变主要起到了驱动性作用，但从某种角度看又不同于现有文献中所定义的驱动性突变；⑦原发癌和转移癌的细胞增殖在本质上是不同的，前者主要用来克服生存压力，而后者则是细胞应对压力过程中衍生的副产品，这就提示，在治疗转移癌和原发癌的策略上应该有根本性的差异。

在本书的每一章，作者都坚持在进化的背景下探讨癌症的不同主题，并以如下的方式提出问题，例如：进化中的病变细胞所处的微环境为其带来了何种压力？不断进化的细胞又是如何适应相关压力的呢？此外，我们努力利用实例来演示从现有组学数据中提取所需信息的过程，既包括如何提出问题，也包括了如何利用“假说驱动”的数据挖掘方法来解决这些问题。例如：原发癌和转移癌的主要驱动力之间的区别是什么？通过寻找所有转移癌的共同上调基因，并与相应原发癌的共同上调基因进行比较，进而利用通路的富集分析结果进行探讨。

我们希望本书能在某种程度上弥合实验癌症生物学家和计算生物学家之间的沟壑，从而有助于揭示隐藏于癌症组学数据中的大量有益信息；也希望能帮助计算癌症生物学家掌握独特的思维方式，从而有别于那些专长于实验研究的同行。作者确信，更多定性或定量地利用组学数据一定会有利于提高我们对癌症生物学的整体认知，并进一步提高早期检出癌症的能力，开发更多有效的治疗手段，从而帮助提高癌症患者的生存质量。

我们期待通过读者的反馈，指出本书中需要修订的内容，以帮助我们纠正谬误、改进表述或调整内容。这些信息必将成为改进和完善后续版本的重要参考。

Athens, GA, USA	Ying XU
Lincoln, NE, USA	Juan Cui
Chapel Hill, NC, USA	Daivid Puett

缩 写 词

以下缩写词贯穿本书。

A2AR	Adenosine A2a receptor	腺苷A2a受体
ABC	ATP-binding cassette	三磷酸腺苷结合盒
ABL1	ABL proto-oncogene 1, non-receptor tyrosine kinase	ABL1原癌基因, 非受体酪氨酸激酶
ACTA	Actin, alpha 1, skeletal muscle	骨骼肌肌动蛋白
ACVR	Activin A receptor	激活素A受体
AIDA	Axin interactor, dorsalization associated	Axin相互作用蛋白
AIF	Apoptosis-induced factor	凋亡诱导因子
AKT	V-AKT murine thymoma viral oncogene homolog	v-akt鼠科胸腺瘤病毒癌基因同源物
ALK	Anaplastic lymphoma receptor tyrosine kinase	间变性淋巴瘤激酶
ALL	Acute lymphoblastic leukemia	急性淋巴细胞性白血病
AML	Acute myelogenous leukemia	急性髓细胞性白血病
APAF	Adaptor protein apoptotic protease-activating factor	凋亡蛋白酶激活因子
APO2L	Also called TNFSF10, tumor necrosis factor superfamily member 10	肿瘤坏死因子超家族成员10
APC	Adenomatous polyposis coli	腺瘤性结肠息肉病
ARID	AT-rich interactive domain	AT丰富结合域
ARSE	Arylsulfatase E	芳基硫酸酯酶E
ASXL	Additional sex combs like	附加性梳样
ATM	Ataxia telangiectasia mutated	共济失调毛细血管扩张突变基因
ATP	Adenosine triphosphate	三磷酸腺苷
ATP	ATPase, H ⁺ /K ⁺ transporting	氢/钾ATP酶通道蛋白
ATR	Ataxia telangiectasia and Rad3 related	共济失调毛细血管扩张Rad3相关蛋白
BAD	BCL2-associated agonist of cell death	BCL2 拮抗因子
BAK1	BCL2-antagonist/killer 1	BCL2同源拮抗剂1
BARD1	BRCA1 associated RING domain 1	乳腺癌易感基因环状结构域蛋白1
BAX	BCL2-associated X protein	BCL2相关X蛋白
BCC	Basal cell carcinoma	基底细胞癌
BCL2	B-cell lymphoma 2	B淋巴细胞瘤/白血病2基因
BCL3	B-cell lymphoma 3	B淋巴细胞瘤3基因
BCL2L	BCL2-like protein	BCL2样蛋白
BCR	Breakpoint cluster region	断裂点簇集区

BID	BH3 interacting domain death agonist	BH3结构域凋亡诱导蛋白
BIM	Also called BCL2L11, BCL2-like 11 (apoptosis facilitator)	BCL2样11 (细胞凋亡辅助蛋白)
BOK	BCL2-related ovarian killer	BCL2相关抗卵巢癌
BRAF	B-RAF proto-oncogene serine/threonine-protein kinase	B-RAF原癌基因丝氨酸/苏氨酸蛋白激酶
BRCA	Breast cancer	乳腺癌易感基因
CAS	CRK-associated substrate	CRK-关联底物
CAD	Carbamoyl-phosphate synthetase	氨甲酰磷酸合成酶
CCL	Chemokine ligand	趋化因子配体
CDC	Cell division control protein	细胞分裂周期蛋白
CDK	Cyclin-dependent kinase	细胞周期蛋白依赖性激酶
CDKN	Cyclin-dependent kinase inhibitor	细胞周期蛋白依赖性激酶抑制剂
CEACAM	Carcino-embryonic antigen-related cell adhesion molecule	癌胚抗原相关细胞黏附分子
CFD	Complement factor D	补体因子D
CHRM3	Cholinergic receptor, muscarinic 3	毒蕈碱型胆碱受体M3
CLDN1	Claudin 1, also called SEMP1, senescence associated epithelial membrane protein 1	细胞间紧密连接蛋白1
CLL	Chronic lymphoblastic leukemia	慢性淋巴细胞性白血病
CML	Chronic myelogenous leukemia	慢性髓细胞性白血病
CNN1	Calponin	钙调节蛋白-1
CNV	Copy number variation	拷贝数变异
COL	Collagen	胶原蛋白
COX	Cyclooxygenase	环氧化酶
CPS	Carbamoyl-phosphate synthase	氨甲酰磷酸合成酶
CRC	Colorectal cancer	结直肠癌
CREB	CAMP-response element-binding protein	环磷酸腺苷酸反应元件结合蛋白
CSPG	Chondroitin sulfate proteoglycan	硫酸软骨蛋白多糖
CTC	Circulating tumor cells	循环肿瘤细胞
CTNN	Catenin (cadherin-associated protein)	环连蛋白
CTSB	Cathepsin B	组织蛋白酶B
CYP	Cytochrome P450	细胞色素 P450
DCC	Deleted in colorectal cancer	结直肠癌缺失基因
DEFA	Defensin, alpha	α -防御素
DES	Desmin	结蛋白
DHPS	Deoxyhypusine synthase	脱氧羟腐胺赖氨酸合酶
DMN	Synemin, intermediate filament protein	中间丝蛋白 β -synemin
DNAPK	Protein kinase, catalytic polypeptide	DNA 依赖性蛋白激酶催化亚基
DNM2	Dynamin 2	发动蛋白 2
DNMT	DNA methyltransferase	DNA 甲基转移酶

DPT	Dermatopontin	皮肤桥蛋白
DR	Death receptor	死亡受体
E2F	E2 promoter binding factor	E2 启动子结合因子
EC	Enzyme classification	酶分类
ECM	Extracellular matrix	细胞外基质
ECT2L	Epithelial cell transforming 2 like	上皮细胞转化序列 2 样基因
EED	Embryonic ectoderm development	胚胎外胚层发育蛋白
EFNA	Ephrin-A	肝配蛋白-A
EGR	Early growth response protein	早期生长反应蛋白
EGF	Epidermal growth factor	表皮生长因子
EGFR	Epidermal growth factor receptor	表皮生长因子受体
EIF1A	Eukaryotic translation initiation factor 1A	真核起始因子 1A
EMMPRIN	Extracellular matrix metalloproteinase inducer	细胞外基质金属蛋白酶诱导因子
EMT	Epithelial-mesenchymal transition	上皮细胞-间充质转化
EP300	E1A-binding protein P300	E1A 结合蛋白 P300
EP	Prostaglandin E receptor	前列腺素 E 受体
ER	Endoplasmic reticulum (subcellular compartment)	内质网
ER	Estrogen receptor	雌激素受体
ERBB2	ERB-B2 receptor tyrosine kinase 2	ERB-B2 受体酪氨酸激酶 2
ERK	Extracellular signal-regulated kinase	胞外信号调节激酶
EZH2	Enhancer of zeste homolog	zeste 基因增强子同源物 2
FADD	FAS-associated protein with death domain	Fas 死亡结构域相关蛋白
FAK	Focal adhesion kinase	黏着斑激酶
FAM26D	Family with sequence similarity 26	序列相似家族 26
FAS	Fas cell surface death receptor	Fas 细胞表面死亡受体
FGF	Fibroblast growth factor	成纤维细胞生长因子
FGFR	Fibroblast growth factor receptor	成纤维细胞生长因子受体
FH	Fumarate hydratase	富马酸水合酶
FHL	Four and a half LIM domains	4 个半 LIM 结构域
FLT3	Fms-related tyrosine kinase 3	FMS 样酪氨酸激酶 3
FN	Fibronectins	纤连蛋白
FOS	FBJ murine osteosarcoma viral oncogene homolog	FBJ 鼠科骨肉瘤病毒癌基因同源物
FXR	Farnesoid X receptor	法尼醇 X 受体
GAD	Glutamate decarboxylase	谷氨酸脱羧酶
GADD	Growth arrest and DNA damage	生长抑制 DNA 损伤基因
GATA3	GATA-binding protein 3	GATA 结合蛋白 3
GCLC	Glutamate-cysteine ligase, catalytic subunit	谷氨酸-半胱氨酸连接酶催化亚基
GCLM	Glutamate-cysteine ligase, modifier subunit	谷氨酸-半胱氨酸连接酶调节亚基

GDT	Growth-to-differentiation transition	生长分化转换
GFRA	GDNF family receptor alpha	GDNF 家族受体 α
GKN1	Gastrokine-1	胃动蛋白 1
GNAS	Guanine nucleotide-binding protein	鸟嘌呤核苷酸结合蛋白
GNPNAT	Glucosamine phosphate <i>N</i> -acetyltransferase	葡萄糖胺磷酸盐 <i>N</i> -乙酰转移酶
GPI	Phosphoglucose isomerase	磷酸葡萄糖异构酶
GPR	G-protein coupled receptor	G 蛋白偶联受体
GPX	Glutathione peroxidase	谷胱甘肽过氧化物酶
GSR	Glutathione reductase	谷胱甘肽还原酶
GSTP1	Glutathione S-transferase P1	谷胱甘肽 S-转移酶 P1
HARE	Hyaluronic acid receptor for endocytosis	透明质酸内吞受体
HAS	Hyaluronic acid synthase	透明质酸合成酶
HDAC	Histone deacetylase	组蛋白去乙酰酶
HDGF	Hepatoma-derived growth factor	肝癌衍生长因子
HER2	Human epidermal growth factor receptor 2	人类表皮生长因子受体 2
HGF	Hepatocyte growth factor	肝细胞生长因子
HIF	Hypoxia inducible factor	缺氧诱导因子
HLAA	Major histocompatibility complex, class I, A	主要组织相容性复合体, I 型, A
HMOX1	Heme oxygenase 1	血红素加氧酶 1
HNRPAB	Heterogeneous nuclear ribonucleoprotein A/B	不均一核核糖核蛋白 A/B
HO1	Heme oxygenase 1	血红素加氧酶 1
HPV16L1	Human papillomavirus 16 L1	人乳头瘤病毒 16 型 L1
HSF	Heat shock factor	热休克因子
HSPA2	Heat shock protein family A (Hsp70) member 2	热休克蛋白家族 A (Hsp70) 成员 2
HSPB1	Heat shock protein family B (small) member 1	热休克蛋白家族 B (小) 成员 1
HYAL	Hyaluronidase	透明质酸酶
IAP	Inhibitor of apoptosis protein	凋亡抑制蛋白
ICAD	Inhibitor of caspase-activated DNase	Caspase 激活的脱氧核糖核酸酶抑制剂
ICGC	International Cancer Genome Consortium	国际癌症基因组联盟
IDH	Isocitrate dehydrogenase	异柠檬酸脱氢酶
IFITM2	Interferon induced transmembrane protein 2	干扰素诱导跨膜蛋白 2
IGF	Insulin-like growth factor	胰岛素样生长因子
IGFBP	Insulin-like growth factor-binding protein	胰岛素样生长因子结合蛋白
IKK	Inducible I kappa-B kinase	入 κ B 抑制蛋白激酶
IKZF	IKAROS family zinc finger	IKAROS 家族锌指
IL1B	Interleukin 1 beta	白介素 1 β
IL7R	Interleukin 7 receptor	白介素 7 受体
IFI	Interferon-induced protein	干扰素诱导蛋白
IFNGR	Interferon gamma receptor	干扰素 γ 受体
ING	Inhibitor of growth family	生长抑制因子
INMT	Indolethylamine <i>N</i> -methyltransferase	吲哚乙胺- <i>N</i> -甲基转移酶

IRF	Interferon regulatory factor	干扰素调节因子
JAK	Janus kinase	贾纳斯激酶
JNK1	c-Jun N-terminal kinase 1	c-Jun 氨基末端激酶 1
JUN	Jun proto-oncogene	Jun 原癌基因
KAT	K (lysine) acetyltransferase	赖氨酸乙酰转移酶
KC	Keratinocyte chemo-attractant	角质细胞趋化因子
KCNN	Potassium intermediate/small conductance calcium-activated channel	中电导钙离子激活钾通道
KEAP	Kelch-like ECH- associated protein	胞质接头蛋白
KGF	Keratinocyte growth factor	角质细胞生长因子
KIT	V-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog	v-kit Hardy-Zuckerman 4 猫科肉瘤病毒癌基因同源物
KRAS	Kirsten rat sarcoma viral oncogene homolog	Kirsten 大鼠肉瘤病毒癌基因同源物
LANCL	Lanc lantibiotic synthetase component c-like protein	LanC 羊毛硫抗生素合成酶组分 C 样蛋白
LDLR	Low density lipoprotein receptor	低密度脂蛋白受体
LIPF	Lipase, gastric	脂肪酶, 胃
LMAN1	lectin, mannose-binding, 1	凝集素, 甘露糖结合, 1
LMNB1	Laminin B1	核纤层蛋白 B1
LMOD	Leiomodin	平滑肌蛋白
LYVE1	Lymphatic vessel endothelial hyaluronan receptor 1	淋巴管内皮透明质酸受体 1
MAO	Monoamine oxidase	单胺氧化酶
MAPK	Mitogen-activated protein kinase	丝裂原活化蛋白激酶
MCL	Myeloid cell leukemia	髓样细胞白血病
MCP	Macrophage chemo-attractant protein	巨噬细胞趋化蛋白
MDK	Midkine (neurite growth-promoting factor 2)	中期因子 (神经轴突生长促进因子)
MDM2	Human homolog of mouse double minute 2	鼠双微粒体 2 的人同源基因
MEK1	Mitogen-activated protein kinase kinase 1	丝裂原活化蛋白激酶激酶 1
MET	Hepatocyte growth factor receptor	肝细胞生长因子受体
MFAP	Microfibrillar-associated protein	微原纤维相关蛋白
MIP	Macrophage inflammatory protein	巨噬细胞炎性蛋白
MLH	mutL homolog	mutL 同源基因
MLL	Myeloid/lymphoid or mixed-lineage leukemia	髓系/淋巴混合谱系白血病
MMEJ	Microhomology-mediated end joining	微同源介导末端连接
MMP7	Matrix metalloproteinase-7	基质金属蛋白酶 7
MS	Mass spectrometry	质谱技术
MSH	mutS homolog	mutS 同源基因
mTORC	Mammalian target of rapamycin	哺乳动物雷帕霉素靶蛋白复合体
MT3	Metallothionein 3	金属硫蛋白
MUC	Mucin	黏蛋白

MXI	MAX-interacting protein	MAX 相互作用因子
MYC	V-myc avian myelocytomatosis viral oncogene homolog	V-myc 髓细胞瘤病毒癌基因同源物
MYD88	Myeloid differentiation primary response gene 88	髓样分化因子初次应答基因 88
MYH11	Myosin, heavy chain 11	肌球蛋白重链 11
MYLK	Myosin light chain kinase	肌球蛋白轻链激酶
NADH	Nicotinamide adenine dinucleotide	还原型烟酰胺腺嘌呤二核苷酸
NANOG	Nanog homeobox	Nanog 同源蛋白
NAV2	Neuron navigator 2	神经元导向因子 2
NF	Neurofibromatosis	神经纤维瘤病
NFE2L2	Nuclear factor, erythroid 2-like 2	红细胞衍生核因子 2 样蛋白 2
NFκB	Nuclear factor of κ light polypeptide gene enhancer in B-cells	B 细胞κ轻肽基因增强子核因子
NGF	Nerve growth factor	神经生长因子
NHE	Na ⁺ -H ⁺ exchanger	钠氢交换体
NMR	Nuclear magnetic resonance	核磁共振
NOXA	Also called PMAIP1, Phorbol-12-myristate-13-acetate-induced protein	佛波醇 12 肉豆蔻酸 13 醋酸酯诱导蛋白
NQO1	NAD(P)H dehydrogenase, quinone 1	NAD(P)H 脱氢酶 (醌) 1
NOTCH	Notch protein	Notch 蛋白
NOX	Phagocyte-like NADPH oxidase	吞噬细胞 NADPH 氧化酶催化亚基 gp91phox 的一系列同源物
NRAS	Neuroblastoma RAS viral (v-ras) oncogene homolog	神经母细胞瘤 RAS 病毒癌基因同源物
NUP160	Nucleoporin 160 kDa	核孔蛋白 NUP160
OGT	O-linked <i>N</i> -acetylglucosamine (GlcNAc) transferase	氧连 <i>N</i> -乙酰氨基葡萄糖转移酶
OR	Olfactory receptor	嗅觉受体
OCT4	Octamer-binding transcription factor 4	八聚体结合转录因子 4
OX40	Also called TNFRSF4, tumor necrosis factor receptor superfamily member 4	肿瘤坏死因子受体超家族成员 4
P53	Tumor protein 53	P53 肿瘤蛋白
PAI	Plasminogen activator inhibitor	纤溶酶原激活物抑制剂
PAK1	P21 protein (CDC42/RAC)-activated kinase	P21 蛋白 (Cdc42/Rac) -活化激酶 1
PAPPA	Pregnancy-associated plasma protein A	妊娠相关血浆蛋白 A
PAR	Protease activated receptor	蛋白酶激活受体
PAX5	Paired box 5	配对盒基因 5
PCDH7	Protocadherin 7	原钙黏蛋白 7
PDGFR	Platelet-derived growth factor receptor	血小板源性生长因子受体

PDK1	Pyruvate dehydrogenase kinase, isozyme 1	丙酮酸脱氢酶激酶同工酶 1
PGE2	Prostaglandin E2	前列腺素 E2
PGM3	Phosphoglucomutase 3	葡萄糖磷酸变位酶 3
PI3K	Phosphoinositide 3-kinase	磷脂酰肌醇 3 激酶
PIK3CA	Phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha	磷脂酰肌醇激酶催化亚单位 A
PKB	Protein kinase B	蛋白激酶 B
PKC	Protein kinase C	蛋白激酶 C
PKM2	Pyruvate kinase isozymes M2	丙酮酸激酶 M2
POF1B	Premature ovarian failure, 1B	卵巢功能早期衰退 1B 基因
PPA	Inorganic pyrophosphatase	无机焦磷酸酶基因
PRDX	Peroxiredoxin	过氧化物酶
PRKCE	Protein kinase C epsilon type	蛋白激酶 C epsilon 型
PRK	Phosphoribulokinase	磷酸核酮糖激酶
PRLTS	Also called PDGFRL, Platelet-derived growth factor receptor-like	血小板衍生长因子受体样蛋白
PRRG1	proline rich Gla (G-carboxyglutamic acid) 1	脯氨酸丰富 Gla (G 羧谷氨酸) 1
PSCA	Prostate stem cell antigen	前列腺干细胞抗原
PTEN	Phosphatase and tensin homolog	磷酸酯酶与张力蛋白同源物
PTTG1	Pituitary tumor-transforming gene-1	垂体瘤转化基因-1
PUMA	P53 up-regulated modulator of apoptosis	P53 上调细胞凋亡调控因子
RAC	RAS-related C3 botulinum toxin	Ras 相关 C3 肉毒杆菌毒素
RAF1	Raf-1 proto-oncogene, serine/threonine kinase	RAF 原癌基因丝氨酸/苏氨酸蛋白激酶
RAS	The same as KRAS	Kirsten 大鼠肉瘤病毒癌基因同源物
RB	Retinoblastoma	视网膜母细胞瘤
RBM5	RNA-binding motif protein 5	RNA 结合模序蛋白 5
REEP	Receptor accessory protein	受体辅助蛋白
RET	Ret proto-oncogene	RET 原癌基因
RFE	Recursive feature elimination	递归特征消除
RHAMM	Hyaluronic acid-mediated motility receptor	透明质酸介导的细胞游走受体
RHOA	RAS homolog gene A	RAS 同源基因家族成员 A
RIP	Receptor-interacting serine/threonine protein	丝氨酸/苏氨酸激酶 1 受体相互作用蛋白
RNS	Reactive nitrogen species	活性氮簇
ROS	Reactive oxygen species	活性氧簇
RPL22	Ribosomal protein L22	核糖体蛋白 L22
RTKN	Rhotekin	Rhotekin 蛋白
RUNX1	Runt-related transcription factor 1	矮小相关转录因子 1
S100	S100 calcium-binding protein	S100 钙结合蛋白
SATB	Special AT-rich sequence-binding protein	特殊富含 AT 序列结合蛋白
SETD	SET domain containing	SET 结构域

SH2B3	SH2B adaptor protein 3	SH2B 接头蛋白 3
SLC5A5	solute carrier family 5 (sodium/iodide cotransporter), member 5	钠/碘共转运体蛋白家族第 5 成员
SMAC	Second mitochondrial activator of caspases	第二线粒体衍生的 caspase 激活剂
SMAD4	SMAD family member 4	SMAD 家族成员 4
SNP	Single-nucleotide polymorphism	单核苷酸多态性
SNRP	Small nuclear ribonucleoprotein polypeptide	小核核糖核蛋白肽
SOD	Superoxide dismutase	超氧化物歧化酶
SOX9	SRY (sex determining region Y)-box 9	Y 染色体性别决定区盒 9
SP	Specificity protein	特异性蛋白
SPI1	Spi-1 proto-oncogene	SPI-1 原癌基因
SRB1	Also called SCARB1, scavenger receptor class B, member 1	B 类 1 型清道夫受体
STAT3	Signal transducer and activator of transcription 3	信号转导子和转录激活子 3
SUZ12	SUZ12 polycomb repressive complex 2 subunit	SUZ12 多梳抑制复合物 2
SVM	Support vector machines	支持向量机
TAL1	T-cell acute lymphocytic leukemia protein 1	T 细胞急性淋巴细胞性白血病 1 蛋白
TAM	Tumor-associated macrophages	肿瘤相关巨噬细胞
TCA	Tricarboxylic acid cycle	三羧酸循环
TCF7L1	Transcription factor 7-like 1	转录因子 7 类似物 1
TCGA	The Cancer Genome Atlas	癌症基因组图谱
TERT	Telomerase reverse transcriptase	端粒酶逆转录酶
TGF	Transforming growth factor	转化生长因子
TGFβR	Transforming growth factor β receptor	转化生长因子 β 受体
TIMP	Tissue inhibitor of metalloproteinase	金属蛋白酶组织抑制因子
TLR	Toll-like receptor	Toll 样受体
TMED6	Transmembrane p24 trafficking protein 6	跨膜 p24 运输蛋白 6
TRADD	TNFRSF1A-associated via death domain	肿瘤坏死因子受体 1 相关死亡域蛋白
TRAF	Tumor necrosis factor receptor-associated factor	肿瘤坏死因子受体相关因子
TRAIL	Transforming necrosis factor-related apoptosis-inducing ligand	肿瘤坏死因子相关凋亡诱导配体
TRIP13	Thyroid hormone receptor interactor 13	甲状腺激素受体相互作用因子 13
TRKB	Tyrosine receptor kinase B	酪氨酸受体激酶 B
TSP	Thrombospondin	血小板凝血酶敏感蛋白
TTN	Titin	肌联蛋白
TXNL	Thioredoxin-like	硫氧还蛋白样
UAP1	UDP- <i>N</i> acetylglucosamine pyrophosphorylase 1	UDP- <i>N</i> -乙酰葡萄糖胺焦磷酸化酶 1
UBFD	Ubiquitin family domain	泛素家族域

UGDH	UDP-glucose dehydrogenase	UDP-葡萄糖脱氢酶
UGP	UDP-glucose pyrophosphorylase	UDP-葡萄糖焦磷酸化酶
ULK	Unc-51 like autophagy activating kinase	Unc-51 样自噬激活酶
UPA	Urokinase	尿激酶
VEGF	Vascular endothelial growth factor	血管内皮生长因子
VEGFR	Vascular endothelial growth factor receptor	血管内皮生长因子受体
VHL	Von Hippel-Lindau tumor suppressor	VHL 肿瘤抑制基因
VLDLR	Very low density lipoprotein receptor	极低密度脂蛋白受体
WISP	WNT1 inducible signaling pathway protein	WNT1 inducible signaling pathway protein
WNT1	Wingless-type MMTV integration site family member 1	无翅型 MMTV 整合位点家族成员 1
WT	Wilms tumor	Wilms 瘤
ZNF367	Zinc finger protein 367	锌指蛋白 367