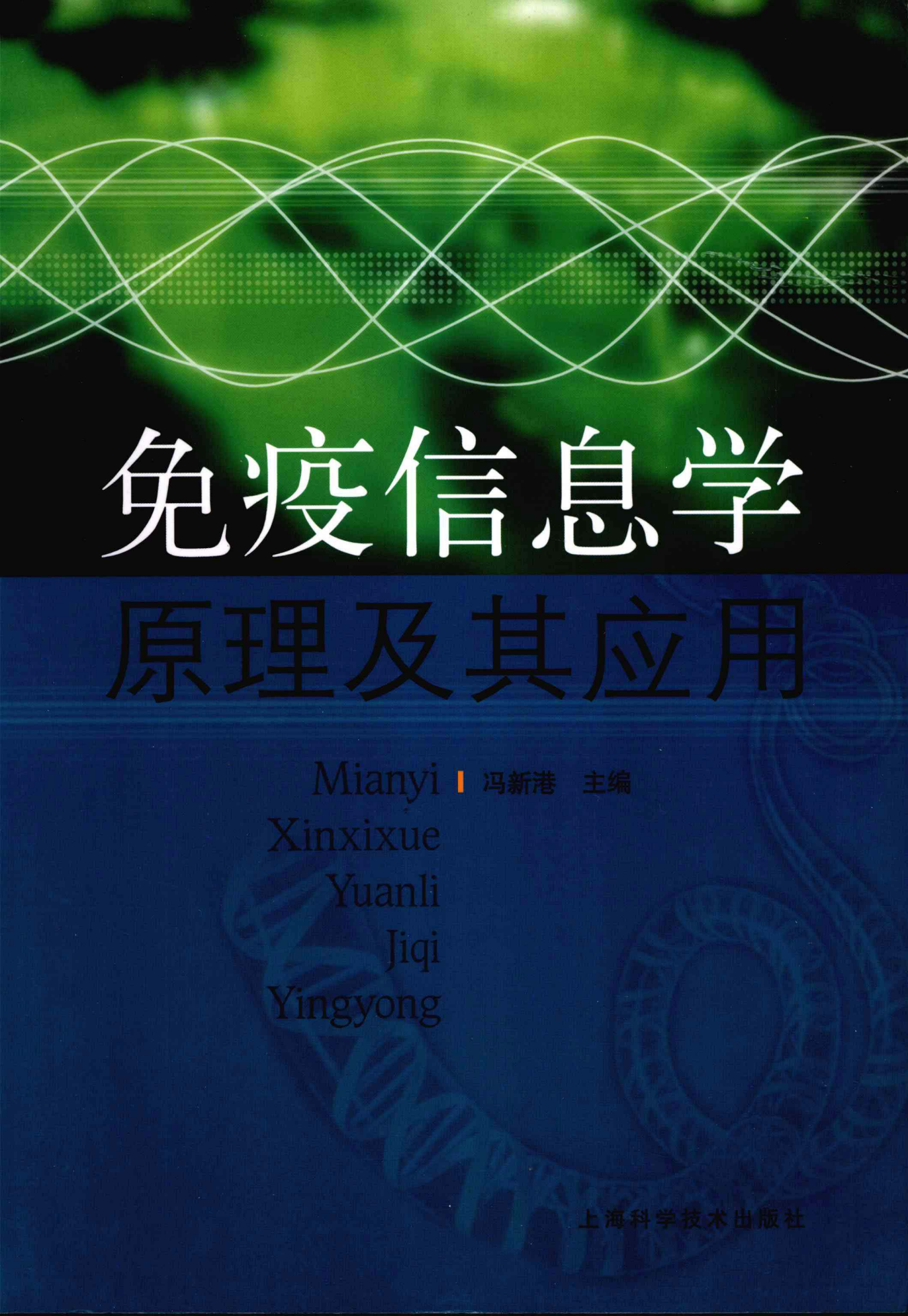




免疫信息学

原理及其应用

Mianyi | 冯新港 主编
Xinxixue
Yuanli
Jiqi
Yingyong

上海科学技术出版社

免疫信息学

原理、工具、应用

陈建峰 王明 王明 王明

科学出版社

科学出版社

免疫信息学原理 及其应用

冯新港 主编

上海科学技术出版社

图书在版编目(CIP)数据

免疫信息学原理及其应用/冯新港主编. —上海:
上海科学技术出版社, 2009. 6
ISBN 978-7-5323-9740-2

I. 免... II. 冯... III. 免疫学—生物信息论
IV. Q939.91

中国版本图书馆 CIP 数据核字(2009)第 006353 号

上海世纪出版股份有限公司 出版、发行
上海科学技术出版社

(上海钦州南路 71 号 邮政编码 200235)

新华书店上海发行所经销

苏州望电印刷有限公司印刷

开本 787×1092 1/16 印张 15.25

字数:306 千字

2009 年 6 月第 1 版 2009 年 6 月第 1 次印刷

印数:1—2 250

ISBN 978-7-5323-9740-2/R·2626

定价:40.00 元

本书如有缺页、错装或损坏等严重质量问题,
请向工厂联系调换

► 内容简介

本书内容包括免疫生物学基础、免疫系统功能建模、免疫信息学数据和数据挖掘、B 细胞表位的预测、T 细胞表位的预测、抗原处理与 TAP 转运预测、感染性疾病的传播动力学及免疫流行病学模型、抗病毒疫苗研发、反向疫苗学在细菌疫苗中的应用、免疫信息学在抗原虫病疫苗研发中的应用、抗蠕虫疫苗研发、免疫信息学和抗肿瘤疫苗的研发、生物信息学在抗自身免疫病疫苗研发中的应用、免疫信息学在实验与临床免疫学中的应用。

主 编 冯新港

编写人员 冯新港 苑纯秀 王孝波 周 杰

邵东华 姚利晓 杨健美 陈兆国

夏艳勋 傅志强

前言

FOREWORD

免疫信息学是一门新兴的交叉学科,有许多理论和技术问题仍有待相关领域的研究工作者深入探讨。对于我们这些从事感染性疾病研究的工作者来说,最感兴趣的课题是预防和治疗产品的研发,这涉及许多学科,其中免疫学占据十分重要的地位,因为它给许多防治产品的成功研发提供了实验和理论基础。但是,对一些老大难的传染病如艾滋病、乙肝、结核等,新发和再现的感染性疾病如 SARS 和禽流感等,以及一些被忽略的热带病如疟疾、血吸虫病等等,仍缺乏十分安全有效的防制产品,因而迫切要求应用包括免疫信息学在内的新理论和新技术,以期针对这类疾病的产品研发取得新的突破。这也是我们决定编写这本书的原因之一。

促使我们编写本书的另一个理由则是课题研究方向拓展的需要。本课题组近年来在血吸虫功能基因组学和蛋白质组学等方面开展了一些工作,获得了大量的基础数据。怎样从这些数据中发现新的疫苗和诊断候选分子,是我们最关心的问题之一,这实际上就涉及反向疫苗学理论和技术的应用。反向疫苗学策略在抗某些病毒、细菌、衣原体,以及寄生原虫的候选疫苗分子的发现中得到了比较成功的应用,它能否应用于抗血吸虫这类多细胞病原体,我们觉得可以尝试一下。我们认为,多细胞病原体感染涉及的生物学和免疫学等过程与机制复杂多样,真正实现基于确定机理的防控产品的研发可能还为时过早,基于试验或经验的筛选仍然是疫苗等产品研发的“硬”途径;但基于数据挖掘和理论模型研究等“软”方法获得的信息和知识,也许可以帮助我们少走弯路,从而加快筛选的进程。这种研究工作的需要促使我们收集、学习相关资料,并以此为基础整理出本书的大部分章节。

至于有些章节涉及的理论免疫学及传播动力学方面的内容,我们认为也是值得向从事感染性疾病防治研究的有关工作者推介的。因为感染性疾病的防控,牵涉群体,如疫苗产品的应用包括免疫程序和效果评估等等,就与病原本身的传播和宿主应答调控的动力学密切相关,因此,对感染过程中宿主和病原群体动力学特征有所了解,对于新防控产品的研究

是具有启发意义的。

由于编写者主要是从事相关领域的应用基础研究工作的,因此在内容的取舍上偏向于免疫信息学技术的应用,在理论原理的系统性论述介绍方面,肯定存在一些不足;加上目前国外同类著作出版很少,国内尚没有类似的参考书,使得我们在内容的选择上难以做到全面完整。因此,本书如能起到抛砖引玉的作用,也就使我们感到非常满意了。

本书由冯新港主编,具体编写分工为:冯新港编写第1章、第3章、第5章、第7章、第8章、附录;王孝波、冯新港编写第2章;周杰编写第4章;邵东华、冯新港编写第6章;姚利晓编写第9章;杨健美编写第10章;陈兆国编写第11章;苑纯秀编写第12章;夏艳勋编写第13章;傅志强编写第14章。

在编写本书的过程中,得到了各位编写者的大力支持,对他们的辛勤工作,深表谢忱!在出版过程中,得到了上海科学技术出版社的指导和帮助,还得到了中央级科研院所基本业务费有关课题基金的资助,在此一并致谢!

冯新港

2009年2月

目 录

CONTENTS

第 1 章	绪 论	1
1.1	免疫信息学是一门新兴交叉学科	1
1.2	免疫信息学研究的基本问题	2
1.3	免疫信息学研究的基本方法	2
1.3.1	数据库构建方法	3
1.3.2	计算与统计学方法	3
1.3.3	数学建模与计算机模拟方法	3
1.4	免疫信息学的应用	4
1.4.1	系统免疫学理论与人工智能研究	4
1.4.2	疫苗研发与传染病研究	5
1.4.3	实验与临床免疫学研究	5
	参考文献	5
第 2 章	免疫生物学基础	9
2.1	免疫应答	9
2.1.1	天然免疫	9
2.1.2	适应性免疫	9
2.2	免疫系统的组成与结构	10
2.2.1	免疫器官	10
2.2.2	免疫细胞	11
2.2.3	免疫分子	12
2.3	免疫系统的机制	12
2.3.1	抗原	12
2.3.2	免疫系统的特异性识别	13
2.3.3	免疫记忆与免疫耐受	15
2.4	免疫监督	16

2.5 免疫效应及其调节机制	16
2.5.1 免疫效应	16
2.5.2 免疫调节	17
2.6 免疫系统的复杂性特征	18
2.6.1 复杂性适应系统	18
2.6.2 免疫系统的自组织	18
2.6.3 免疫系统网络	19
2.7 小结	19
参考文献	20

第3章 免疫系统功能模建

3.1 淋巴细胞的免疫识别与增殖动力学	21
3.1.1 T、B淋巴细胞受体库大小估计	21
3.1.2 淋巴细胞增殖与分化的克隆模型	24
3.2 独特型网络与免疫应答	28
3.2.1 独特型网络与免疫记忆	28
3.2.2 独特型网络与免疫耐受	29
3.3 免疫系统的调节	29
3.3.1 免疫相关细胞的分子调控网络	29
3.3.2 免疫应答的细胞调节	31
3.3.3 宿主免疫应答的系统水平的调节	32
3.4 小结	34
参考文献	34

第4章 免疫信息学数据库和数据挖掘

4.1 数据库概念	37
4.2 免疫信息学数据库	39
4.2.1 简介	39
4.2.2 基于网络的免疫信息学公共数据库利用	39
4.2.3 免疫多态性数据库(IPD)	42
4.2.4 SYFPEITHI	44
4.2.5 IMTECH 免疫信息学数据库	44
4.2.6 抗原决定簇数据库(IEDB)	45
4.3 数据挖掘与知识发现	45
4.3.1 数据挖掘在免疫信息学中的作用	45
4.3.2 数据挖掘在免疫信息学中的应用步骤	46
4.3.3 短肽预测的数据挖掘实例	46

参考文献	47
------------	----

第 5 章	B 细胞表位的预测	49
5.1	线性 B 细胞表位的预测方法	49
5.1.1	预测算法的基本性质	50
5.1.2	倾向性标度简介	50
5.1.3	常用预测方法与服务器	52
5.2	构象表位预测方法	54
5.2.1	基于序列和 3D 结构的方法	54
5.2.2	基于结构的预测方法	55
5.2.3	基于序列-结构的预测方法	55
5.2.4	模拟位预测法	56
5.3	小结	56
	参考文献	57

第 6 章	T 细胞表位的预测	60
6.1	CTL 结合肽预测	60
6.1.1	基于序列的方法、基于结构的方法	61
6.1.2	基于结构的方法	63
6.1.3	常用程序及服务器	64
6.1.4	发展方向	65
6.2	辅助 T 细胞(Th)结合肽预测	65
6.2.1	基于序列的方法、基于结构的方法	66
6.2.2	常用程序及服务器	69
6.2.3	发展方向	69
6.3	小结	70
	参考文献	70

第 7 章	抗原处理与 TAP 转运预测	74
7.1	HIV 蛋白酶裂解预测	74
7.1.1	HIV 蛋白酶裂解预测的结构基础	74
7.1.2	预测原理和方法	75
7.2	蛋白酶体裂解位点的预测	77
7.2.1	结构基础	77
7.2.2	预测原理和方法	78
7.3	抗原处理相关转运蛋白(TAP)的预测	80
7.3.1	TAP 结构和特征	80

- 7.3.2 TAP 结合预测原理与方法 80
- 7.3.3 机器学习方法 81
- 7.4 整合的方法 82
 - 7.4.1 SMM 法 82
 - 7.4.2 NETCTL 法 83
 - 7.4.3 WAPP 法 83
 - 7.4.4 EpiJen 法 84
- 7.5 小结 85
- 参考文献 85

第 8 章 感染性疾病的传播动力学及免疫流行病学模型 88

- 8.1 感染性疾病的传播动力学模型 88
 - 8.1.1 三个基本模型 89
 - 8.1.2 HIV 病毒动力学模型分析 90
 - 8.1.3 同性恋人群中 HIV 感染的基本流行模型分析 91
 - 8.1.4 SARS 传播动力学分析 92
- 8.2 重要寄生虫病的免疫流行病学模型研究 93
 - 8.2.1 免疫流行病学 94
 - 8.2.2 疟疾的免疫流行病学模型分析 94
 - 8.2.3 血吸虫免疫流行病学模型分析 95
 - 8.2.4 放牧动物线虫免疫流行病学模型分析 97
- 8.3 小结 99
- 参考文献 99

第 9 章 抗病毒疫苗研发 103

- 9.1 抗 HIV 病毒疫苗研发 103
 - 9.1.1 减毒活疫苗 105
 - 9.1.2 灭活疫苗 105
 - 9.1.3 病毒颗粒样疫苗 106
 - 9.1.4 亚单位疫苗 106
 - 9.1.5 裸 DNA 和活重组载体疫苗 106
 - 9.1.6 免疫信息学在 HIV 疫苗研究中的应用——抗原表位疫苗 107
- 9.2 抗 HBV 病毒疫苗研发 109
 - 9.2.1 有效抗原表位的筛选 110
 - 9.2.2 免疫机理的研究 110
- 9.3 小结 111

参考文献	111
------------	-----

第 10 章 反向疫苗学在抗细菌疫苗中的应用 113

10.1 疫苗学的发展历程	113
10.2 传统抗细菌疫苗的现状 & 反向疫苗学的创建	113
10.2.1 活疫苗	114
10.2.2 死疫苗	114
10.2.3 细菌的代谢产物	114
10.2.4 亚单位疫苗	114
10.2.5 基因重组疫苗	114
10.2.6 DNA 疫苗	114
10.3 反向疫苗学在抗细菌疫苗中的应用	115
10.3.1 脑膜炎奈瑟球菌 B	115
10.3.2 结核分枝杆菌	118
10.3.3 炭疽杆菌	119
10.3.4 肺炎链球菌	121
10.4 小结	122
参考文献	123

第 11 章 免疫信息学在抗原虫病疫苗研发中的应用 127

11.1 免疫信息学在抗疟疾疫苗研发中的应用	128
11.1.1 疟疾的危害与重要性	128
11.1.2 疟疾疫苗研究的必要性、现状及存在的问题	128
11.1.3 免疫信息学在抗疟疾疫苗研发中的应用现状与展望	129
11.2 免疫信息学在抗利什曼病疫苗研发中的应用	136
11.2.1 免疫信息学在抗利什曼病减毒活疫苗研发中的应用	136
11.2.2 免疫信息学在抗利什曼病基因工程亚单位疫苗研发中的应用	138
11.2.3 免疫信息学在抗利什曼病 DNA 疫苗研发中的应用	138
11.3 免疫信息学在抗弓形虫病疫苗研发中的应用	139
11.3.1 免疫信息学在抗弓形虫病遗传改良疫苗研发中的应用	139
11.3.2 免疫信息学在抗弓形虫病重组亚单位疫苗研发中的应用	140
11.4 免疫信息学在抗隐孢子虫病疫苗研发中的应用	140
11.5 免疫信息学在抗球虫病疫苗研发中的应用	141
11.6 小结	143
参考文献	143

第 12 章	抗蠕虫疫苗研发	151
12.1	吸虫疫苗的研发	151
12.1.1	应用反向疫苗学理论与技术研发抗血吸虫疫苗	151
12.1.2	基于血吸虫基因组序列信息的疫苗候选分子发现技术	152
12.1.3	血吸虫保护性抗原或表位的实验筛选与鉴定	153
12.1.4	保护性 B 细胞抗原或表位的筛选与鉴定	154
12.1.5	T 细胞抗原或表位的筛选与绘制	155
12.2	绦虫疫苗研发	156
12.2.1	绦虫感染免疫生物学——绦虫疫苗研发的实验基础	156
12.2.2	基于基因组序列信息的候选疫苗分子的发现	157
12.2.3	抗原表位预测及鉴定	157
12.3	线虫疫苗的研发	158
12.3.1	大规模 EST 序列的测定与分析	159
12.3.2	高通量筛选差异表达分子	159
12.4	小结	160
	参考文献	160
第 13 章	免疫信息学和抗肿瘤疫苗的研发	163
13.1	免疫信息学在肿瘤疫苗候选抗原分子鉴定上的应用	164
13.1.1	CTL 筛选法	164
13.1.2	重组 cDNA 表达文库的血清学方法	164
13.1.3	多肽洗脱法	165
13.1.4	cDNA 的差示分析技术	165
13.1.5	蛋白质组学技术	165
13.1.6	组合肽库技术	165
13.1.7	肿瘤抗原 B 细胞表位、T 细胞表位的免疫信息学预测	166
13.2	肿瘤疫苗研究进展	167
13.2.1	肿瘤细胞疫苗	167
13.2.2	肿瘤抗原疫苗	168
13.2.3	多肽疫苗	170
13.2.4	核酸疫苗	171
13.2.5	树突状细胞(DC)疫苗	171
13.3	小结	174
	参考文献	174
第 14 章	生物信息学在抗自身免疫病疫苗研发中的应用	180
14.1	免疫信息学在自身免疫病发病机制研究中的应用	181

14.1.1	免疫信息学在 I 型糖尿病疫苗相关分子研究中的应用	181
14.1.2	免疫信息学在类风湿关节炎疫苗相关分子研究中的应用	182
14.1.3	免疫信息学在系统性红斑狼疮疫苗相关分子研发中的应用	183
14.1.4	免疫信息学在其他自身免疫病发病机制研究中的应用	184
14.2	生物信息学在自身免疫病蛋白疫苗研发中的应用	185
14.2.1	生物信息学在 I 型糖尿病蛋白疫苗研发中的应用	185
14.2.2	免疫信息学在类风湿关节炎肽疫苗研发中的应用	187
14.2.3	免疫信息学在系统性红斑狼疮肽疫苗研发中的应用	189
14.2.4	免疫信息学在其他自身免疫病肽疫苗研发中的应用	189
14.3	免疫信息学在自身免疫病 DNA 疫苗研发中的应用	191
14.3.1	生物信息学在 I 型糖尿病 DNA 疫苗研发中的应用	191
14.3.2	生物信息学在类风湿关节炎 DNA 疫苗中的应用	194
14.3.3	生物信息学在系统性红斑狼疮 DNA 疫苗研发中的应用	194
14.3.4	免疫信息学在其他自身免疫病 DNA 疫苗研发中的应用	195
14.4	免疫信息学在自身免疫病细胞疫苗中的应用	196
14.5	小结	197
	参考文献	197

第 15 章	免疫信息学在实验与临床免疫学中的应用	205
15.1	免疫信息学在实验免疫学中的应用	205
15.1.1	T 细胞识别与活化功能分析	206
15.1.2	免疫系统相关分子的鉴定及其作用机制	207
15.2	免疫信息学在临床免疫学中的应用	211
15.2.1	变应原预测	212
15.2.2	免疫疾病相关基因的鉴定	213
15.3	免疫信息学改进器官移植、肿瘤预后结果分析	214
15.4	小结	216
	参考文献	216

附录 A	220
附录 B	227

第 1 章

绪 论

1.1 免疫信息学是一门新兴交叉学科

免疫信息学(immunoinformatics)是建立在现代免疫学和信息学基础上的一门新的交叉学科,是运用信息学的理论和方法解决免疫学及疫苗学问题的学科,也是研究免疫系统调控和免疫应答过程中信息传递规律的科学^[1]。immunoinformatics 一词由以色列魏茨曼研究所的已故著名的应用数学家、数学生物学家和理论免疫学家 Lee A. Segel 于 2001 年提出^[2],最初用来描述免疫系统中由细胞信号分子如细胞因子、趋化因子等介导的调控作用机制,其中的信息学含义与信号发送与传递等性质相似。此后,一些生物信息学者、疫苗学者以及免疫学者对免疫信息学概念的内涵和外延进行了拓展,并被人们所接受^[3]。严格说来,要给免疫信息学一个非常具体的学科界定和教科书式的阐释是比较困难的,这主要是由于免疫信息学与相关的一些新兴学科有着密切的联系。

我们知道,20 世纪 90 年代以来,由于多学科理论与技术(如高通量测序及生物芯片等技术)的进展与突破,各种组学(如基因组学、蛋白质组学及代谢组学等)蓬勃发展,给科学界提供了海量的生物学基础数据。人们相信这些通过还原论方法和技术(以分子生物学技术为主流)获得的生物学数据深刻而又复杂,必须采用综合论方法与技术对其进行阐释^[4];而综合论方法则是以信息论方法为基础的,它利用信息技术分析大量的生物数据,从中发掘出生物学的一般原理,并试图通过各个生物组成部分的研究推导出系统行为^[5]。因此,一门以计算机为工具对生物信息进行储存、检索和分析的新兴学科——生物信息学(bioinformatics)应运而生,并在现代生命科学、医学、农学等多学科领域中得到了广泛的应用,也在一定程度上推动了上述实验学科研究范式的转变进程,即从单纯的“假说驱动”(hypothesis-driven)到假说驱动辅以数据驱动(data-driven)。免疫学在经历了一个世纪的实验为主的研究后,也同样积累了大量的数据资料,迫切需要应用生物信息学理论与技术对其进行综合分析,以发现其中蕴涵的免疫学规律,故而有的学者认为免疫信息学实质上是生物信息学的一个分支学科^[6],从方法论上说,这个界定是合理的。

另一个与免疫信息学密切相关的学科是理论免疫学(theoretical immunology)^[7]。在研究免疫系统内部存在的各层次调控网络规律时,主要采用构建理论模型的方法。有两类常用的免疫学理论模型构造形式:即语言描述性的或非数学的和数学的。非数学的理论常由实验免疫学家提出,如先由 Neils K. Jerne 提出、后经 MacFarlane Burnet 作完整阐述的

克隆选择说(clonal selection theory)^[8, 9]和 Neils K. Jerne 的免疫网络学说(immune network theory)^[10],这两个理论假说后来也被发展成定量的学说,合理地解释了免疫学中的一些核心问题,如免疫反应的特异性、免疫识别、免疫记忆和免疫耐受等,不但奠定了现代免疫生物学研究的理论基础,而且对人工免疫系统(artificial immune system)的发展起到了基础性与决定性的作用。另一类理论常用数学模型来描述,如预测表位大小的模型及亲和成熟模型等;数学模型理论在研究免疫系统全局行为时具有明显优势,是揭示实验免疫学获得的数据中潜在规律的有力工具。

理论免疫学研究在 20 世纪 90 年代中期以前一直比较活跃。此后,人们更多关注实验细胞与分子免疫学的工作,因而积累了大量的免疫大分子方面的信息。怎样在多层次水平探索这些分子间相互作用的过程和机制的问题,愈来愈受到重视。免疫组学(immunomics)就是研究免疫相关的全套分子、它们的作用靶分子及其功能的新兴学科,这一界定非常宽泛。因此,也可以将免疫信息学纳入到免疫组学的范畴内^[11]。

1.2 免疫信息学研究的基本问题

免疫信息学研究的基本问题主要有两大类,一是免疫生物学中的基本问题;二是疫苗学与临床免疫学中的问题,其侧重点是免疫相关分子的结构与功能研究,特别强调分子相互作用及其宏观性质特征的研究。正是通过研究上述问题,才奠定了免疫信息学的实验和理论基础。归纳起来主要有如下几方面的问题:

免疫相关分子的系统发生学研究:主要是对抗原受体、MHC 分子以及细胞因子进行不同种间的比较基因组学方面的研究,揭示它们的进化生物学意义^[12, 13, 14]。

免疫细胞和分子的多样性及其机制研究:研究淋巴细胞受体和 MHC 多样性的分子基础及其产生机制,为在个体和群体水平上阐明脊椎动物适应性免疫应答机制提供基础^[15, 16]。

免疫系统网络功能研究:包括免疫识别、免疫记忆和免疫耐受等现象的产生、维持及调节机制研究,以揭示免疫网络调节的复杂性规律^[17, 18, 19]。

免疫学数据的管理、挖掘和利用:疫苗学与实验临床免疫学研究积累了大量的数据资料,通过信息学方法挖掘利用这些数据来探索一些重要问题,如揭示病原与宿主互作机制、建立预测模型、设计与评估免疫干预或治疗效果,以及发现新的知识(如 MHC 多态性与某些免疫疾病敏感性的关系)等多方面的问题,都是值得深入研究的^[20, 21, 22, 23, 24, 25]。

总之,免疫信息学研究的内容与理论免疫学、免疫基因组学、免疫组学以及疫苗学等所关注的问题有一定程度的交叉,但其突出的特点是数据驱动为主的,没有大量完整和正确的实验数据为基础,预测和模拟就不可能正确,这一点要特别引起免疫信息学研究工作者的注意。

1.3 免疫信息学研究的基本方法

免疫信息学对免疫生物学和疫苗学中基本问题的阐释是以数据和模型为基础的,最基