

电子细胞的研究现状与展望

赵明生¹, 尚 彤², 孙冬泳², 蒋景宏¹, 汤 健², 吴佑寿¹

(1. 清华大学电子工程系, 北京 100084; 2. 北京大学医学部心血管所, 北京 100083)

摘 要: 电子细胞是利用计算机电子信息技术等先进手段, 模拟再现细胞内外部生命活动的现象和过程, 并用于探索细胞生命活动的潜在规律. 它是一种人工生命复杂系统, 代表了生命科学和信息科学的交叉学科前沿, 有十分重要的作用. 本文综述了电子细胞的研究发展, 就其研究内容、实例、与其关系密切的学科领域、以及它的作用和意义进行了扼要阐述, 最后, 对它的发展做出展望和建议.

关键词: 生物信息学; 电子细胞; 虚拟细胞; 人工生命

中图分类号: Q81 **文献标识码:** A **文章编号:** 0372-2112 (2001) 12A-1740-04

The Current State and Prospect of E-Cell

ZHAO Ming-sheng¹, SHANG Tong², SUN Dong-yong², JIANG Jing-hong¹, TANG Jian², WU You-shou¹

(1. Dept. of Electronic Engineering, Tsinghua University, Beijing 100084, China;

2. Cardiovascular Institute, Peiking University Health Science Center, Beijing 100083, China)

Abstract: Electronic cell or E-cell simulates and reveals life activities around cell internally and externally by using advanced tools of computer science, electronic information technology and others. It can be used as a new powerful tool to explore the potential nature laws of cell life. E-cell represents the frontier of life science and information science. It is a type of artificial life and a very complex system. This paper gives survey on the history and cause of E-cell research and development. And briefly presents its research contents, two real examples, other scientific areas closely related to it, and its functions and great significance. Finally, we address its future prospects and some suggestions on its research progress.

Key words: bioinformatics; electronic cell; E-cell; virtual cell; artificial life

1 电子细胞研究的发展

一个多世纪以来, 人们一直十分重视对细胞的研究. 近几十年来, 随着计算机仿真计算和可视化能力的提高, 以及在数学、信号与信息处理科学领域的发展, 于 90 年代中期以后诞生了一种研究细胞的新的技术手段——电子细胞 (E-Cell, Electronic Cell)^[1, 3, 4, 7, 8, 9, 11, 12, 18].

电子细胞亦称虚拟细胞 (Virtual Cell). 它是在计算机上模拟真实细胞的结构、物质组成、生命活动的动力学行为和生命现象, 用虚拟现实的方式实现友好人机交互, 以便研究者构造细胞结构和其内外部环境物质组成, 考察、记录细胞实验现象和功能, 再现细胞生命活动和发现新的生物学现象规律. 因此, 电子细胞亦称人工细胞, 是人工生命的重要基础部分.

电子细胞的实现极为复杂, 它综合了生物学、生理学、生物化学、数学、物理学、化学和信息科学等多学科的理论知识, 达到对细胞的结构和功能进行分析和整合的一门新兴交叉学科方向. 很明显, 它是新兴的生物信息学和生命信息学研究的最重要内容之一^[2, 4].

自从发明显微镜以后, 人类对生物学的研究开始在生命的最基本单位——细胞的层次上展开探索, 产生了细胞学说. 电子显微镜的出现, 使细胞学上升到分子生物学的新微观境界, 从基因的对基因和基因的深入研究, 产生了划时代的生命科学研究成果. 特别是近 20 年来, 基因组计划^[15]的实施, 人

们对生命的本质有了更深刻的认识, 积累了极为丰富的数据和资料. 迄今还完成了 40 多组生物的全基因组测序工作, 包括人类基因组约 30 亿碱基对的测序工作已经完成. 仅登录在美国 GenBank 数据库中的 DNA 序列总量已超过 70 亿碱基对, 可以预计今后的 DNA 序列数据的增长将更为惊人. 生物学的的数据并不仅仅表现在 DNA 测序方面, 与之同步的还有蛋白质的一级结构, 即氨基酸序列的增长. 此外迄今为止, 已有一万多种蛋白质的空间结构以不同的分辨率被测定. 在这些数据的基础上派生整理出来的数据库已达 500 余个. 这一切构成了生物数据的海洋, 这种科学数据的急速和海量积累, 在人类的科学研究历史上是空前的^[29].

数据并不等于信息和知识, 但却是知识和信息的源泉. 关键在于如何从中找到规律和一般现象. 与生物学数据的指数增长相比, 人类关于细胞的生命活动规律和现象的认识却增长十分缓慢^[5, 6, 13, 14, 29], 人们对复杂的细胞行为尚不能做出预测估计. 借助于人们对细胞已有知识的掌握和对积累的大量生物学数据的计算机分析, 利用计算机和信息技术, 研究和构造细胞的计算模型, 并对模型加以模拟和修正, 是当前看来深入研究细胞的唯一最好方法和手段. 在可行性、研究代价、成果的产出速度和产出率上无疑是最优的和不可替代的.

电子细胞的出现是科学发展的必然. 借助于先进的计算工具和计算理论, 对实验数据进行分析并综合信息处理结果

必将达到对生命活动更快、更好、更省和更深入的认识。

电子细胞的发展历史不足十年。日本和美国自 90 年代开始提出和重视计算机模拟细胞的计划,到 90 年代末,对整体细胞模型的研究取得了相当的成果。E-cell 计划发起于 1996 年,主要的发起单位是日本的 Keio 大学,他们率先提出了世界上第一个电子细胞模型,并开发出整个电子细胞仿真的软件环境^[9,11,22,23]。美国也于近年来提出 Virtual Cell 研究计划, Schaff、Goryanin 等人已成功组织开发出类似系统^[10,28]。

电子细胞的发展大体上可以分为三个阶段,首先是构造阶段,这一阶段主要研究电子细胞的物质构成、基本功能模拟,细胞内外环境与组成物质的动力学行为特性知识库系统建立,方便的人机交互可视化界面及开放的网络设计支撑平台(包括计算的数学库,图像图形工具,网络通信工具、设计工具、计算环境、专业数据库等);其次,是细胞功能行为模拟试验和优化完善阶段;第三是应用阶段,这一阶段电子的细胞已具备了强大的模拟真实细胞的能力,在细胞的类型、数量、种类上极为丰富,功能上也基本完善实用,成为医学、生物学、生物工程、能源和环境等领域实际应用的一种不可或缺的重要工具,是相应产业的重要支撑技术。目前,国际上的主要工作仍处于第一阶段,即构造研究阶段。

全面和真实地模仿一个细胞尚有不少的困难,最主要的问题之一是缺少定量的数据,大多数生物学的知识是以定性的方式提供的,而实现模拟需要量化,正在积累的大量数据和先进的信息处理方法,在使模型更加优化并逐步接近真实结果方面,已起到越来越重要的作用。

美国国家普及医药研究院(NIGMS)已提出了对心血管细胞信号传导和 B 细胞的资助计划^[10,11,18,19,24],最终的计划目标是创造这些特别类型细胞的电子细胞。美国国家能源部继基因组计划之后设立了微生物细胞计划(Microbial Cell Project-MCP)^[20],其目标是在计算机上建造虚拟的微生物细胞,这对能源、环保、生物免疫能力的研究有十分重要的意义。

2 电子细胞的研究内容

电子细胞是在电子系统设备上,主要是计算机及其网络,模拟再现细胞生命活动的全过程,基本问题是细胞的生长、发育、繁殖和凋亡过程的建模模拟,涉及细胞内外环境和细胞相关物质对细胞的生命活动的作用表现,这包括离子通道、信号传导、细胞的组成结构及相应功能,细胞内外物质活动和生化反应、基因表达、蛋白质、酶、能量物质等相互作用的动力学行为的真实再现。另外,涉及细胞的医学、病理学、药理学、生物学上的功能,以及对环保、能源等领域的作用,也是电子细胞的重要研究范畴。

下面将列出其几个主要研究方向,并只作简单介绍:

(1) 细胞组成、结构和功能建模。主要问题是细胞的各组成成分物理结构、生化功能、物质流动、信号生成等,是对构成整个细胞的模拟模型有作用的各个有机成份及其相互作用功能作用的研究。

(2) 细胞生命活动相关数据库及相应管理软件。搜索、发现、整理和存贮细胞生命活动的相关试验测试数据,建立数据库,在此基础上创造出知识库,如基因相关、蛋白质相关、生化

反应及其它反应链、代谢过程、细胞各物质组成及其相互作用的知识库。这是逼真地再现和预测细胞生命活动的基础。

(3) 细胞间通信和信号传导。主要问题是细胞生命活动信号的产生机理、传导机理和细胞间相互作用的建模、模拟和有关现象规律的研究。

(4) 染色体和基因表达。主要问题是细胞染色体活动的模拟、基因表达的产生条件、产生过程、动力学模型、基因的功能、遗传、变异规律等。

(5) 蛋白质合成、结构分析预测。从细胞基因出发,应用分子力学和分子动力学的原理建模,并观察发现蛋白质的结构规律、合成规则,分析已探明蛋白质的生成机理,预测和模拟产生新的蛋白质。

(6) 细胞代谢过程模拟和分析。主要是细胞的物质代谢和能量代谢过程的机理、生化反应、动力学建模,模拟分析等。

(7) 分子进化和比较基因组学。利用不同物种中基因序列的异同研究细胞中氨基酸序列甚至相关蛋白质的功能结构,模拟这一过程,通过对比来研究分子进化和细胞进化。近年来,由于较多模式生物基因组测序任务的完成,为从整个基因组角度来研究分子和细胞进化提供了条件。

(8) 疾病诊疗。生物的病变在很大程度上反映在细胞的病变,研究细胞的物质成份组成和表型变化等来诊断病变,并采用多种手段在电子细胞上再现治疗控制过程,将对发现新的治疗和诊断方法有十分重要的作用。这方面对人类健康意义重大,可望得到丰硕的成果。

(9) 药物设计。电子细胞的目的之一是可视化再现和阐明细胞生命活动的规律,刻划和预测细胞中各种成份的结构功能、相互作用以及与各种人类疾病之间的关系,以寻求各种防治药物和方法。在电子细胞上进行药物设计和药理实验,既经济又快速安全。如为了抑制某些酶和蛋白质的活性,可以在已知其结构的基础上,利用分子对接算法,在电子细胞上设计抑制剂分子,作为候选药物,这种发现新药的方法有强大的生命力,也有巨大的经济效益。

(10) 其它。如多细胞相互作用的动力学现象,细胞之间的相互作用影响,细胞营养和毒性研究,细胞的整个生命过程的建模和模拟,细胞相关的组织器官研究,高通量技术,在环保、能源、化工工程中的应用等,将逐渐成为并丰富电子细胞的重要研究内容,这里不再赘述。

3 电子细胞模型实例介绍

目前已有两种电子细胞的模型,一个是 1997 年日本学者建立的原核细胞能量代谢的模型^[9,11,23];另一个是 1999 年美国学者建立的真核细胞钙转运的模型^[10,24]。此外,美国国立卫生研究院(NIH)和能源部(DOE)投资数千万美元,正在筹建关于细胞信息传递和生物利用能量的虚拟细胞^[16]。

3.1 虚拟原核细胞模型——日本的电子细胞

日本 Keio 大学的 Masaru Tomita 教授是电子细胞的先驱者,他领导的研究组以原核细胞生物支原体为对象,在 1997 年实现了世界上第一个虚拟电子细胞^[9,11,22,23],应用这一细胞模型可以模拟细胞内糖、脂肪、甘油和能量代谢过程及机理。

该电子细胞以支原体生殖糖代谢过程为主,选取了 127 个与代谢过程相关的基因,对细胞内与代谢过程密切相关的物质与能量活动过程建立了数学模型,采用面向对象方法实现计算机模拟。

其最主要的模型化方法是将细胞内的物质和单元分为三类:物质、反应和系统。其中,“物质”代表了细胞参与反应的组成成份,包括基因、蛋白、酶等,基因表达和反应物浓度的变化为物质主要的动态属性;“反应”代表了细胞中各种反应的类型,以及这些反应的动力学方程和相关参数。“反应”的功能主要表达参与反应的物质之间的关系和数量的变化;“系统”代表了细胞中生物反应的特定场所和相互关联的各环节。

3.2 虚拟真核细胞模型——美国的虚拟细胞

美国的虚拟细胞研究主要侧重于细胞内部的生化反应动力学过程、信号传导等各功能模块的研究。美国的虚拟细胞计划实施以后,国家能源部、NRCAM (National Resource for Cell Analysis and Modeling) 等部门又分别支持微生物、心血管肌细胞及 B 细胞等研究^[18,19]。在因特网上公开提供了一种分析细胞部分功能表现的虚拟细胞应用演示软件^[24],该软件侧重于计算机可视化技术、网络化技术、后台数据管理和计算技术,其贡献是为细胞研究者提供了一个灵活、开放的环境,使用者可以灵活地定义细胞组成结构和内部动力学模型。

利用虚拟细胞可以完成离子通道的钙离子流动、RNA 的转运,线粒体的作用和细胞核膜的作用等一系列关于真核细胞的生物学活动和功能。在使用虚拟细胞作实验时,首先由用户通过客户端向服务器提交实验“说明书”——试验的要求和涉及的物质与反应等,服务器根据用户对实验的定义以及实验细胞的模型,从后台数据库、知识库中提取数据,并用服务器上的软件分析系统进行处理,完成虚拟试验。最终的试验结果以类似生物学实际试验结果的形式,包括图像、表格返回给用户。在整个实验中,用户可以控制实验过程和各种参数。

虚拟细胞的构建过程一般分为四大步骤:首先,应由生物领域专家根据生命科学对细胞及其生理过程的认识程度,选择建模对象,包括选择细胞的种类、可以模拟的功能等。其次,要在对相关数据资料整理加工的基础上,建立相关数据库和知识库。第三,设计与实现。这是构建虚拟细胞最关键的一步。即依据虚拟对象——细胞与其功能,利用相关数据库和知识库,通过系统的数学计算分析,将生物反应过程实现数字化模拟,建立虚拟细胞的模型。这需要用高性能计算实现其数学计算。如美国采用了“Beowulf”结构的并行计算机系统^[25],这种计算机系统采用多机并行网络计算技术完成计算任务,具有非常强大的计算能力。第四,组装、测试和维护。主要包括电子细胞的控制界面、分析控制系统、数学计算系统和反应界面等几部分的综合集成。限于篇幅,这里就不赘述。

4 与电子细胞关系密切的学科领域

电子细胞是多学科领域的一项交叉学科技术,从广义上说,它应该属于新兴的生物信息学科范畴。它包含了细胞生命活动信息的获取、处理、储存、分发、分析、综合和解释等在内的所有方面,它综合运用数学、信息科学技术和生物学的各种

工具,来阐明和理解大量数据所包含的生物学意义^[1,2,3,4,7,8,]。

对细胞来说,一方面,人们长期以来积累了大量相对零散的细胞实验数据和知识,另一方面,我们在医学、药物、农业、环保、能源等方面对新知识的渴求,这些新知识将帮助人们改善生存环境,预防和诊疗疾病、提高生活质量。因而,从积累的大量数据和部分知识上升到系统化的整合知识成为当今十分迫切的研究课题,电子细胞的研究目标也正基于此。

为达到这一目标,单靠生物学的方法已显得难以奏效,还必须借于数学,信息科学的理论、方法和技术。涉及的相关学科领域和技术主要有:统计学;概率论与随机过程理论,特别是隐马尔科夫链模型(HMM);信息论和信号处理学;最优化理论;神经计算技术^[27];传统的人工智能,知识推理和新兴的模糊技术;计算机和网络技术;数据库和知识库技术;数据挖掘技术等。

生物学,包括遗传生物学、神经生物学、细胞学、微生物学和分子生物学、医学、免疫学、药理学、生物物理学、生物化学等,无疑是与电子细胞最为密切的学科。其它如农业科学、营养学、食品、环境、能源等科学,与生物学的物种生长发育、微生物的生化作用十分相关,其实质是与细胞极为相关。

5 作用意义及发展展望与建议

5.1 电子细胞的作用和意义

研究电子细胞是为了能使人们更好、更快地掌握和发现细胞生命活动的规律及知识,服务于医疗、教育、科学研究和社会生活。它可以有以下明显的作用。

医疗:用于病变的早期预防、病变的诊断、疾病的治疗模拟、保健、新药物的实验和发明等。

建立正常和病理的电子细胞模型,不仅可以发现细胞内部活动和调节的生理机制,而且可以了解和揭示疾病发病过程,寻找到有效致病分子和标记分子,进行疾病的预警诊断,提出防治和干预措施,设计和试验新药物,建立新的医疗保健模式——e-Doctor,发展新的生物高技术产业。

教育:电子细胞以其生动和可视化的表现形式可以改变传统医学生物学教学模式,部分代替和辅助传统医学生物学实验和教学,可以生动、直观、透彻、逼真地重现细胞生命活动的各环节,还可以用于实现远程教育 and 电子教育,使医学教育社会化、个性化,提高医学教学的效率和效果。在经济性和易用性上有显著的优势,甚至会成为将来教学实验的主要方式。

科学研究:采用电子细胞可代替或辅助真实细胞进行各种科学研究。可实现实际实验中很难实现的条件,预测试验的结果,发现一些在真实细胞实验很难观察到的现象和规律。可以设定各种复杂的实验条件,进行全方面的干预、简化实验操作、节约实验材料和时间。因此,电子细胞亦是芯片上的生物实验室或研究所,在科学研究中的价值不可低估。

社会生活:电子细胞是一个多能的仿生环境。应用电子细胞,可以观察环境因素对人体的影响及其作用途径,提出防治措施;应用电子细胞可以虚拟生物对能量的摄取和利用,指导能源的有效利用和开发。通过电子细胞可以模拟营养、药物、毒物对细胞的作用,避免对人和生物的直接作用所产生的不

良后果。可以虚拟和模拟基因和蛋白质的结构,预测基因和蛋白质的功能,加速新的物种、产品的开发等等,这一切对国民经济和社会生活都有重大的意义。

电子细胞是生命科学技术的重大突破,将使生命科学和信息学科的研究内容及方法手段产生革命性的创新。虽然这一点在国内尚无明确的认识,但在国外已初现端倪。人类研究历史也表明,科学数据的大量积累将导致重大科学规律的发现。例如对数百颗天体运行数据的分析导致了开普勒的三大定律和万有引力定律的发现;数十种元素和上万种化合物数据的积累导致了元素周期表的发现;氢原子光谱学数据的积累促成了量子理论的提出,为量子力学的建立奠定了基础。历史的经验值得注意,有理由认为,过去的科学发现和重大进展是靠单一的人脑实现的,现在学会借助于强大信息处理能力的计算机,可使工作效率成亿倍地提高,必将会从今日巨大积累的细胞学数据中,极大地加速导致细胞和生命科学重大规律的发现,而电子细胞技术将会是实现这一飞跃的最重要手段。其巨大作用如计算机的发明对当今社会的影响相类似,六、七十年代的人们大多不能料想到计算机机会像今天这样成为不可或缺的工具。可以预计,电子细胞在十年左右的时间会成为医学、生物学、药学、营养学、生态、环境、农业等学科研究产业领域不可或缺的重要工具。

5.2 展望和建议

电子细胞的出现仅有不足五年的研究历史,采用信息技术手段对生命复杂体系的研究在国外亦属刚刚起步阶段。目前我们的差距主要在相关的计算机、可视化计算、网络技术等多项技术与生命科学技术的综合应用水平上。实现电子细胞的分项技术国内外水平没有太明显差距。近年来我国在数学计算、大型计算机研制、人工智能、人工神经网络、细胞和分子生物学基础、临床医学等方面已经取得了十分可喜的进展,积累了大量的数据和方法,奠定了良好的技术和物质基础。此外,我国科学家长期受辩证唯物主义的思想教育,擅长于综合分析,辩证论治的哲学思维,富于顽强拼搏的精神。我们的基础工作已经具备,所拥有的条件也大大改善,相信在国家的支持下,通过多学科合作,在 3~5 年内建立我国自己的电子细胞技术平台,赶上国外先进水平并成为电子细胞技术强国是有可能的。

参考文献:

- [1] Dennis Normile. Building working cells in silico [J]. Science, 1999, 284 (5411) :80 - 81.
- [2] Joel B Hagen. The origins of bioinformatics [J]. Nature Reviews, Genetics, 2000, 1:231 - 236.
- [3] Szostak JW, Bartel DP, Luisi PL. Synthesizing life [J]. Nature, 2001, 409 Suppl :387 - 390.
- [4] Declan Butler. Computing 2010: from black holes to biology [J]. Nature, 1999, 402: c67 - c70.
- [5] Ewan Birney, et al. Mining the draft human genome [J]. Nature, 2001, 409 (15) :827 - 828.
- [6] Roos DS. Bioinformatics-trying to swim in a sea of data [J]. Science, 2001, 291 (5507) :1260 - 1261.
- [7] Palsson B. The challenges of in silico biology [J]. Nature Biotechnology, 2000, 18 (11) :1147 - 1150.
- [8] Nadia S. Halim. Biological informatics: The virtual cell: modeling cellular processes [J]. Scientist, 1999, 13 (10) :6.
- [9] Tomita M, Hashimoto K, Takahashi K, et al. E-cell: Software environment for whole cell simulation [J]. Bioinformatics, 1999, 15 (1) :72 - 84.
- [10] J Schaff, L M Loew. The virtual cell [A], Pacific symposium on bio-computing [C], 1999, 4:228 - 239.
- [11] Masaru Tomita. Whole-cell simulation: a grand challenge of the 21st century [J]. TRENDS in Biotechnology, 2001, 19 (6) :205 - 210.
- [12] John H Carson, Ann Cowan, Leslie M Loew. Computational cell biologists snowed in at cranwell [J]. Trends in Cell Biology, 2001, 11 (6) :236 - 238.
- [13] Michael Y Galperin, Eugene V Koonin. Who's your neighbor? New computational approaches for functional genomics [J]. Nature Biotechnology, 2000, 18:609 - 613.
- [14] Smyth P. Data mining: data analysis on a grand scale [J]. Statistic Methods Med. Res. 2000, 9(4) :309 - 327.
- [15] genomes to life program [R]. Proposed by the office of biological and environmental research and office of advanced scientific computing research of the U. S. Department of Energy, 2001. 2.
- [16] Microbial Cell Project [R]. U. S. Department of Energy Office of Science Biological and Environmental Research Program, 2000.
- [17] Asthagiri AR, lauffenburger DA. Bioengineering models of cell signaling [J]. Biomed. Eng. Annu. Rev. , 2000, 2:31 - 53.
- [18] <http://www.cellularsignaling.org/>
- [19] http://www.nigms.nih.gov/funding/gluegrant_release.html
- [20] <http://microbialcellproject.org/>
- [21] <http://www.nrcam.uchc.edu/conference/index.html>
- [22] <http://www.e-cell.org/poster/guzi/icsb2000html/icsb2000.html>
- [23] <http://www.e-cell.org/poster/system/index.html>
- [24] http://www.nrcam.uchc.edu/vcell_development/vcell_dev.html
- [25] <http://www.dnaco.net/~kragen/beowulf-faq.txt>
- [26] <http://cmbi.bjmu.edu.cn>
- [27] G Schneider, P Wrede. Artificial neural networks for computer-based molecular design [J]. Progress in Biophysics & Molecular Biology, 1998, 70 (3) :175 - 222.
- [28] Goryanin I, et al. Mathematical simulation and analysis of cellular metabolism and regulation [J]. Bioinformatics, 1999, 15:749 - 758.
- [29] 张春霆. 生物信息学的现状与展望 [J]. 世界科技研究与发展, 2000, 22 (6) :17 - 20.

作者简介:



赵明生 男, 1968 年 2 月出生于安徽安庆市。1995 年毕业于清华大学电子工程系获博士学位, 专业为信号和信息处理。现为清华大学电子工程系副教授, 研究兴趣和方向为神经计算与模式识别、网络信息处理、生物信息学。已发表国际学术论文 20 余篇。