文章编号: 1007 - 4252(2022)04 - 0004 - 356　　DOI: 10.20027/j.gncq.2022.0044

材料基因组技术在集成电路材料研究中的应用进展

孔维悦[2],林紫威[1],李　鑫[1],黄　琪[2],朱　雷[1,2,*]

(1.中国科学院上海微系统与信息技术研究所　信息功能材料国家重点实验室,上海 200050; 2. 上海集成电路材料研究院,上海 200050)

摘要: 传统材料科学的“试错式”研究方法日益成为限制现代材料工业发展与突破的瓶颈,而材料基因组技术作为材料科学研究的革新技术,可以加速新材料的研发和应用,缩减研究成本和周期,成为当下研究的热点。本文围绕材料基因组技术,介绍了高通量实验、高通量计算与材料数据库三大要素的主要内容和典型研究范例。重点梳理了材料基因组技术在集成电路领域的研究进展,指出了集成电路材料研发的复杂性和广阔的可优化空间,提出了要结合材料基因组技术加速集成电路材料的创新,促进未来集成电路材料领域的发展。

关键词: 材料基因组; 高通量实验; 高通量计算; 材料数据库; 集成电路材料

中图分类号: TN40　　**文献标识码:** A

Application Progress of Materials Genome Technology in the Research of Integrated Circuits Materials

KONG Wei-yue[2], LIN Zi-wei[1], LI Xin[1], HUANG Qi[2], ZHU Lei[1,2,*]

(1. State Key Laboratory of Information Functional Materials, Shanghai Institute of Microsystems and Information Technology, Chinese Academy of Sciences, Shanghai 200050, China; 2. Shanghai Institute of IC Materials, Shanghai 200050, China)

Abstract: Traditional materials research relaying on trial and error has increasingly become the bottleneck for the development and breakthroughs of modern materials industry. Innovative materials genome technology has become a hotspot in materials science research due to it can shorten the cycle of new materials development and reduce research costs. This paper mainly introduces three parts of materials genome technology: high-throughput calculation, high-throughput experimentation, material database. Some representative works are also mentioned. Then the materials genome researches in the field of integrated circuits (IC) are reviewed. In addition, the complexity and optimize ability of the IC materials genome research are pointed out, and we recommend accelerating the innovation and development of integrated circuit materials by using materials genome technology.

收稿日期: 2022-06-10;　　**修订日期:** 2022-07-15

基金项目: 国家重点研发计划项目(2020YFB0704502)

作者简介: 孔维悦(1995-),女,硕士,工程师,主要从事集成电路材料技术研究(E-mail: weiyue.kong@sicm.com.cn).

通信作者: 朱　雷(1988-),男,博士,工程师,主要从事集成电路材料基因组技术研究(E-mail: leizhu@mail.sim.ac.cn).

Key words: Materials Genome; High-throughput Experimentation; High-throughput Computing; Material Database; Integrated Circuits Materials

0 引言

传统材料科学研究主要依靠“试错法”，通常依照“提出假设-实验验证”的顺序开展材料研究，利用材料的现有理论与知识经验，通过调整材料配比，进行表征测试和检验，从而不断逼近目标材料，但这种方法研发周期长、研发成本高，无法满足现代工业对新材料发展的需求[1-2]。2011 年 6 月，时任美国总统的奥巴马提出了“材料基因组计划”(Materials Genome Initiative，MGI)，其核心是利用高通量材料计算、高通量材料实验和材料数据库技术三大要素研究突破、找寻和建立材料从原子排列到微观组织形式再到材料性能之间的关系[3-4]，将材料“发现-研发-生产-应用”的速度至少提高一倍，成本至少降低一半为目标[5]，发展以先进材料为基础的高端制造业[6-8]。

1999 年，中国材料科学界召开主题为“发现和优化新材料的集成组合方法”的第 118 次香山科学会议，寻找加速发现新材料的有效途径。2011 年，中国科学院和中国工程院主办主题为“材料科学系统工程”的第 S14 次香山科学会议，研究中国应对 MGI 的策略[9]。同时，国内多所高校、科研院所也在探索如何利用材料基因组技术提高效率，并开始关注材料基因在工程技术等实用领域的应用。经过多年发展，材料基因组技术在催化材料[10-11]、锂电材料[12-13]、热电材料[14-15]等领域均有一定的应用。但是其在集成电路(Integrated Circuit，IC)材料领域的研究和应用缺乏系统性整理。本文主要分为两大部分，第一部分将围绕材料基因组的三种主要工作模式，即基于高通量理论计算的计算驱动模式、基于高通量制备与表征实验的实验驱动模式以及基于大数据、机器学习和数据挖掘的数据驱动模式[16]，介绍目前材料基因组技术的研究进展。第二部分侧重于介绍材料基因组技术在 IC 材料领域的应用与发展情况。

1 材料基因组三大工作模式

1.1 高通量计算技术

材料基因组框架下的材料高通量计算是以高性能计算平台和软件为基础，借助并发式自动流程高通量算法，实现新材料成分、结构、性能等方面的高效筛选[17]，降低了研发成本、缩短了研发时间，目前高通量计算已在新材料研发设计中实现了广泛的应用。

1.1.1 计算平台

材料基因组要求建立更强的计算工具，创建更强的数据共享、管理和分析平台。基于这一理念，集成计算材料工程成为了其重要组成部分。在保证计算精度以及可靠模拟的前提下，利用高通量材料集成计算关键技术和平台最大程度地减少材料研发过程中的实验成本。

由中国科学院计算机网络信息中心杨小渝团队研发出来的 MatCloud(http://matcloud.cnic.cn)是国内首个高通量材料集成计算基础平台和软件框架，主要包括高通量计算驱动引擎、高通量计算环境、材料计算数据库以及材料数据分析与挖掘 4 个部分，将计算和数据集成于一体[18-19]，并可实现通过 Web 浏览器进行在线作业设置、作业提交和监控。基于该平台可以进行多通道、多任务、高并发的材料计算模拟，并通过结构筛选、元素替代、性能与成分优化等尽可能地建立起材料组分、结构和性能的定量关系模型以帮助加速发现新材料[20-22]。

上海大学于 2017 年上线了具有自主知识产权的高通量第一性原理计算平台：材料信息平台(Material Information Platform，MIP)(https://mip.shu.edu.cn/)，该平台实现了基于第一性原理等方法的材料数据计算、基础/计算材料数据检索及在线可视化展示、面向材料结构与功能的数据预测分析方法和应用模型、基础/计算材料数据及相关软件下载等主要功能，同时由于集成了电输运算法模块，MIP 成为目前唯一具有准确热电性质模拟的高通量平台。席丽丽等人[23]基于 MIP 平台，以新型高性能热电材料为研发对象，对硫族类金刚石结构的热电化合物做了高通量的筛选工作，最终提出一类新型类金刚石结构化合物，并通过实验验证所提出的化合物的热电优值 ZT 大于 1.0。

北京航空航天大学孙志梅团队基于 Python 和 C++自主开发了集高通量自动计算流程、材料大数

据和人工智能(Artificial Intelligence，AI)算法于一体的可视化平台(Artificial Learning and Knowledge Enhanced Materials Informatics Engineering，ALKEMIE)[24]。近期，该团队发布了分布式高通量自动流程集成计算和数据管理智能平台 ALKEMIE 2.0[25]，平台概况及可视化工作流如图 1 所示，主要包括三类：1)具备建模、数据库画图分析的基础(Basic)类模块；2)集成不同尺度计算软件的计算模拟工作(Works)类模块，包含第一原理计算软件 VASP[26]、分子动力学软件 LAMMPS[27]、热力学计算软件 Gibbs[28-29]，相图相场计算软件 OpenPhase[29-30]等；3)与机器学习和数据挖掘相关的 AI 类模块，可用于开发材料结构能量预测、原子受力预测等模型。该团队基于 ALKEMIE2.0 平台从 81 种二元铝合金中筛选了结构能量稳定的合金化元素，并通过 BoltzTraP 高通量自动工作流[31]，计算了二元铝合金的导电性，实验结果表明 Be、Zn、Ga、Ge、In、Sn 均可相应地提升 Al 合金的导电性能。

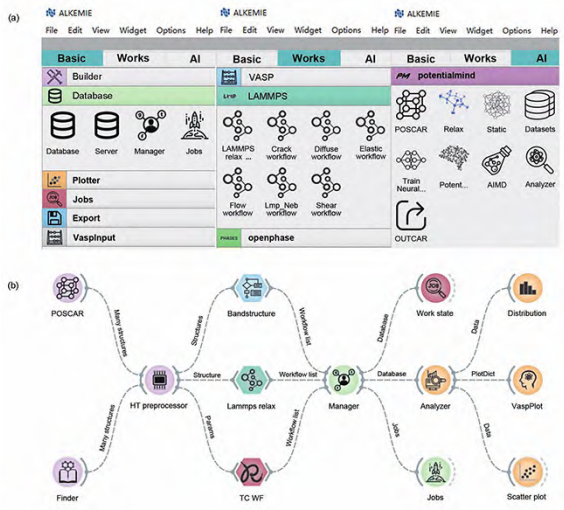

图 1 ALKEMIE2.0 平台概况(AIMD—第一性原理分子动力学，HT—高通量，TC WF—居里温度工作流)(a)模块(b)ALKEMIE2.0 可视化工作流[25]

Fig. 1 The outline of ALKEMIE2.0 platform (AIMD—ab initio molecular dynamics, HT—high-throughput, TC WF—Curie temperature workflow) (a) the modular (b) work area and workflows of ALKEMIE2.0[25]

国际上也存在众多贴合材料基因组理念的关于高通量计算平台的研究。2011 年，麻省理工学院和劳伦斯伯克利国家实验室联合创建的 Materials Project(www.materialsproject.org)，不仅是一个开放的无机材料数据库，使得研究者可以下载材料结构、材料性质、相图等，还是一个高通量计算平台，可进行交互式探索和数据挖掘，甚至用于搜索未知材料[32]。随后美国西北大学 Wolverton 研究团队受到 Materials Project 平台的启发，于 2013 年建立了开放量子材料数据库(Open Quantum Materials Database，OQMD)(http://oqmd.org/)，这是一个基于密度泛函理论(Density functional theory，DFT)计算的材料热力学和结构数据库，目前已计算超过 102 万种材料的热力学和结构特性的数据，其中以钙钛矿数据居多。此外，由德国马普所和柏林洪堡大学于 2014 年联合建立的 Novel Material Discovery(NOMAD)(https://nomad-lab.eu/)是全球最大的 open-access 材料计算数据库之一，供所有人免费使用、上传和下载材料，甚至可为个人数据提供数字对象标识符(DOI)，保证了数据的可引用性。该平台不仅存储了大量材料计算的输入与输出文件，还提供了功能强大的 Archive、Encyclopedia、Artificial Intelligence Toolkit 工具包，用于各类数据下载、检索和分析，加速了研究者们对材料大数据中关键信息的挖掘。

VASP 软件作为材料计算工作的重要工具，每年有众多相关研究发表，但其编译和运行等需要在 Linux 环境下进行，而为了使得 VASP 软件能够在 Windows 环境中运用，Materials Design 公司开发了全功能材料设计与性质预测平台——MedeA(https://www.materialsdesign.com/)。该平台集成了建模、数据库、性质计算与预测等模块，例如数据库模块涵盖了无机晶体数据库 ICSD、晶体结构数据库 COD、金属晶体数据库 NCD、无机材料及相图数据库 MSI Eureka，计算模块则包含了量子力学 VASP、分子动力学 LAMMPS、力场优化工具、蒙特卡洛 GIBBS、介观尺度模拟 Mesoscale、高通量计算工具 MedeA-HT 等工具，使得用户能够创建各种材料类型，并实现材料的多尺度和高通量计算，甚至直接预测材料的理化性质。Chaudhary 等[33]利用 MedeA 软件模拟了二氧化硅纳米颗粒的 x 射线衍射(X-Ray Diffraction，XRD)、能带结构、光谱和态密度。

1.1.2 高通量计算方法

除了基础平台建设，采用合适的简化算法减少传统材料计算的高耗费，也是高通量计算的一大研究热点。针对不同的材料类型，可以开发不同的高

通量简化算法。

Andrew 等人[34]在 AFLOW 工作流框架上，构建了用于简化高通量计算的 AFLOWπ，新算法无需任何额外输入即可产生紧束缚模型的哈密顿算子，且可简化 Hubbard U 校正中的自洽、声子色散、弹性特性、复介电常数和电子传输系数的计算。Restuccia 等人[35]基于 AiiDA 框架发布了一个计算协议，用于通过第一性原理计算高通量获取固体界面的两个内在摩擦特性——粘附能和理想界面剪切强度，该计算协议要求的输入文件简单，这也是第一次在摩擦学中使用高通量方法。Mathew 团队[36]开发了基于 Materials Project 的开源 Python 工具 MPInterfaces，该软件包可自动生成各种界面结构，准备第一性原理计算所需输入文件，并在远程计算资源上部署高通量工作流，该软件已被用于界面体系的高通量计算筛选。Sarikurt 等人[37]利用密度泛函理论与半经典玻尔兹曼输运理论结合用于高通量计算二维层状材料的热电系数和电子适应度函数，从而预测了多个之前未被归类的各向同性和各向异性的二维热电材料。

中国科学院物理研究所陈立泉院士团队[38-41]开展了自动化高通量计算方法的研究，将快速“键价和”法与高精度第一性原理分子动力学方法相结合，从众多材料数据库中收集得到了 20 万余条无机晶体数据，并对其中含锂材料相关性质进行了计算，最终建立起包含电解质与电极材料的数据库。河南大学白莹团队[42]设计了将不同精度计算方法相结合的高通量筛选流程，并综合结构匹配、扩散通道等因素，筛选出了两种可能与锂离子电池富锂正极材料相匹配的包覆化合物 Li_2SiO_3 和 Li_2SnO_3，并通过实验制备证实了这两种材料可有效改善 Li_2MnO_3 电极的循环稳定性。中国科学院物理研究所/北京凝聚态物理国家研究中心的研究组[43]根据晶体材料能带中高对称点能带的对称性，发展了自动计算材料拓扑性质的新方法，在近 4 万种材料中发现了 8 000余种拓扑材料，是过去十几年找到的拓扑材料总和的十倍。清华大学沈洋团队[44]通过结合相场理论与高通量计算，构建了电热耦合击穿模型，定量计算聚合物基电介质介电击穿的热效应，设计出了三明治叠层结构的高能量密度聚合物基纳米复合材料，储能密度提高了 2.44 倍。

1.2　高通量实验技术

材料高通量实验作为材料基因组的三大核心要素之一，主要包含高通量制备技术与高通量表征技术，旨在将传统的顺序迭代方法改为并行处理，利用较少的实验次数筛选组分最优的实验体系，提高实验效率。高通量实验在材料基因组中扮演着承接作用，所获得的丰富实验结果可以为材料模拟计算提供基础数据及实验验证；同时，高通量实验还可以为材料数据库提供海量数据支持，优化材料筛选、预测模型。

早在 20 世纪 60 到 70 年代，Kennedy[45]和 Hanak[46]分别在各自的研究领域中不断重复样品合成-表征过程，自发地提出了“多样品”的实验设想；20 世纪 80 年代中期，组合化学的概念兴起，以 Affymax[47]为代表研制的生物芯片表明组合化学可显著提高生物、化学领域的研发效率；到了 20 世纪 90 年代中期，项晓东和 Schultz 等发展和完善了现代高通量组合材料实验方法，并很快在多种新型功能材料的筛选中进行了应用与示范推广[48]；20 世纪 90 年代末，高通量实验相关技术已经较为广泛地应用于材料科技工业领域[49]；经过 40 余年的发展，材料高通量实验技术已形成覆盖薄膜、块体、粉末等多种材料形态的高通量制备和满足热学、光学、电学、电磁学等各类材料性能高通量表征的完整实验技术体系。

1.2.1 高通量制备

非晶合金作为典型的多组元合金材料，通常具备元素多样性和复杂性，结合高通量制备技术可明显加速高性能非晶合金材料的按需设计及开发。2019 年，柳廷辉、汪卫华团队利用高通量组合方法研制出的高温块体非晶合金：铱/镍/钽金属玻璃（以及其他含硼的金属玻璃），其玻璃转变温度高达 1 162 K，同时其在 1 000 K 的高温下表现出较高的强度（3.7 GPa）[50]。近期，该团队基于高通量实验及数据分析实现了非晶合金新材料的流程化研发（图 2（a）），并与耶鲁大学 Jan Schroers 团队合作从 12 个合金体系中制备了大约 5 700 种合金，并对 XRD 图中第一个衍射峰的半峰全宽 Dq 进行了表征，发现不同合金的非晶形成能力与 Dq 存在明显关联。这一结果相比一般的高通量实验方法效率提高 100 倍，证实了材料基因组在非晶合金研发中的有效性和高效率[51]。

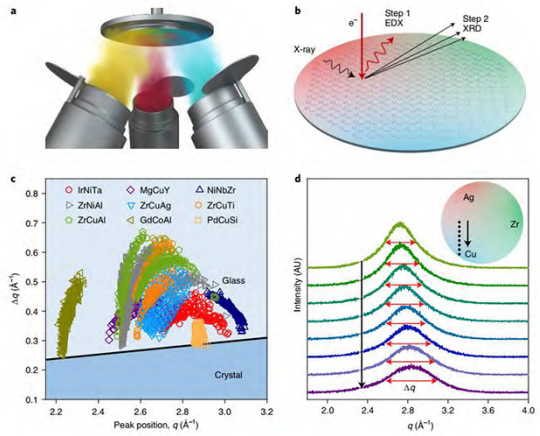

图 2 非晶合金高通量研发流程，包括组合材料制备、自动化成分分析和结构表征、衍射峰自动分析[51]

Fig. 2 Flow chart illustrating the present combinatorial method and the obtained dataset [51]

近期，刘轶团队[52]与美国 MTI 公司合作开发了一套全流程高通量合金制备实验设备，可多工位、批量自动化制备具有离散成分的大块合金样品，比传统单样品制备实验过程至少加速 10 倍。该团队在机器学习方法的预测指导下设计优化了 138 种成分硬质合金 CoxCryTizMouWv，并采用该套全流程高通量设备进行了合成。

在化学反应合成方面，如药物研发环节，传统的化学合成、分离、结构鉴定、活性筛选模式存在诸多不足，例如目标不明确、耗时长、工作量大、成本高、操作繁琐等等。近年来，随着自动化技术的发展以及各种分子细胞水平药物活性评价方法的出现，新药研发中出现了高通量筛选技术，大大加快了候选化合物的活性筛选速度，同时这一筛选技术的出现也导致化学样品的需求量迅猛增长，并促进高通量合成技术的发展。目前，Merck 公司、Pfizer 公司陆续开发并推出自动化高通量反应筛选平台[53-54]，其中，Merck 公司通过结合自动化高通量纳摩尔量级化学反应与亲和选择质谱法不仅实现了用最少的原料在最短的时间内评估数千个纳摩尔量级的反应，筛选反应底物和反应条件，并且无需分离纯化，即可对大量产物混合物进行全面快速的靶标蛋白亲和性筛选和排序，以预测这些产物的生物活性，并且由此指导后续的合成、纯化、生物活性测试等等工作。

1.2.2 高通量表征技术

在样品表征方面，高通量制备需要配套的高通量表征技术以快速分析高通量样品，实现高通量实验的有效性。近二十年来，发展较成熟的高通量表征技术，包括高通量成分与结构表征技术，如同步辐射光源在从红外至硬 X 射线全光谱范围内均能实现高亮度微聚焦，同时还具有高准直性、全光谱、高偏振、宽波段、高亮度、高纯净等优秀特性，从而能够很好地满足高通量组合材料样品所需的亮度和空间分辨率要求。高通量微区热力学表征技术如 Vlassak 等[55-56]开发的一套并行纳米扫描量热系统(PnSC)可以并行测量材料的焓变、热容、相变温度等热力学参数，温度范围从室温到 900 ℃左右，灵敏度可达 10 nJ/K。高通量微区电化学表征技术如 AMETEK 公司下的普林斯顿应用研究(Princeton Applied Research)开发的微区扫描电化学工作站(VersaSCAN)，空间分辨率可达 50 nm，可广泛用于高通量样品包括锂电池正负极、薄膜电解质、半导体等重要材料的电化学性能的精确测试。高通量微区磁学表征技术如采用衰逝微波探针显微镜配以自动化的样品台控制和数据采集，可以实现组合材料芯片的高通量、全自动电磁学特性测量[57-60]。高通量力学表征技术，包括使用原位拉伸数字图像处理技术[61]、纳米压痕仪等技术手段，目前已有的商业化纳米压痕仪器设备可以实现高通量、全自动的微区测试。对材料表面形貌的高通量表征包括使用高通量扫描电子显微镜平台[62]、高频原子力显微镜等。高通量光学表征技术如利用连续光谱椭偏仪[63]、激光椭偏仪[64]、阴极荧光计[60,65]、光致荧光测试仪[57]等实现高通量微区光学性质表征。此外，针对其他不同微区性能的高通量表征技术[66-67]均已发展出相对成熟的技术手段。

1.3 数据库和机器学习技术

近年来，得益于越来越多开源数据集的建立、算法的不断改进以及计算机算力的飞速增长，基于数据库与机器学习的材料信息学成为了材料科学领域的重要研究手段。不同于利用计算机进行硬编码模拟实验结果的传统计算方法，机器学习方法依据已有的材料数据建立模型来学习数据集的构成规则，发现数据间的潜在联系，然后利用该模型对材料属性进行定性或定量预测，因此该方法又称为数据驱动方法。数据驱动方法可明显加快研发效率、缩短开发周期、降低计算成本。不论是微观还是宏观层

面，该方法都可以被应用于材料科学领域中的新材料发现、材料特性预测等。

1.3.1 数据库

现代计算数据库的快速发展为快速筛选、设计或预测材料等奠定了坚实基础。除目前已开发的许多通用数据库[68-69]，还有许多专用数据库，例如，针对二维材料构建的 2DMatPedia 数据库（http://www.2Dmatpedia.org/），包括 6 000 多种 2D 材料的结构、热力学、机械、电子和磁性能，通过对数据库进行智能挖掘，加速对二维材料、光伏材料、电子材料等材料的研发。此外，国内刘淼、孟胜团队开发建立了无机晶体材料数据库 Atomly（https://www.atomly.net/），目前包含了超过 20 万个无机晶体材料的第一性原理计算结果，包括化合物的晶体结构、电子结构和能量等。该团队提出了与电负性差和配位数有关的描述符，并基于 Atomly 数据库的高通量量子力学计算数据，训练机器学习模型，显著提高了材料的形成能的预测精度，验证发现其最高可达 R^2 = 0.982、MAE = 0.07 eV · atom^{-1}[70]。

实现大量现有实验和计算数据的高效积累以及有效应用，才可以更好地满足材料科学研究者们的研发需求，因此材料数据库的建设显得尤为重要。由于数据库中数据的质量、数量和存储形式等将会直接影响后续机器学习等数据挖掘技术的精度和准确度，建设数据库应当朝着以下方向进行：1）具备多样化的存储形式，可存储结构化、非结构化和半结构化的数据；2）实现专业化存储，针对特定材料建立专业数据库，便于行业人员使用；3）进行数据挖掘，利用材料数据库存储的大量数据，进行机器学习，发现材料成分、工艺和性能之间的构效关系[71]。数据库的建设是一项长期工作，虽然目前我国在数据资源储备量方面还存在很多不足之处，但这也意味着国内的材料数据库建设具有很大的发展空间。

1.3.2 机器学习

大数据时代，机器学习为加速材料科学常规的实验设计带来了新机遇。以往通过组合不同合成配方以获得所需材料性能的方法昂贵又费时，而机器学习技术极大地帮助了新材料的设计与预测，同时能够迅速实现高通量实验最佳条件的优化。目前机器学习技术已广泛应用于材料科学领域，例如在材料结构预测方面，Nusrat 等[72]利用神经网络架构预测多主元合金（Multi-Principal Element Alloys，MPEAs）相结构的准确性达到 99%以上，并比较了不同描述符的相对重要性。四川大学蒲雪梅团队[73]开发了一种基于图神经网络（Graph Neural Networks，GNN）的新型深度学习框架，将先验知识有效地嵌入到分子图的端到端学习中，来预测共晶体的形成以加速筛选共晶体（图 3），利用该框架所建立的模型在独立测试集上的准确率达到了 98.86%，对含能共晶体的分布外预测准确率高达 97.11%，具有较强的稳健性和泛化能力。

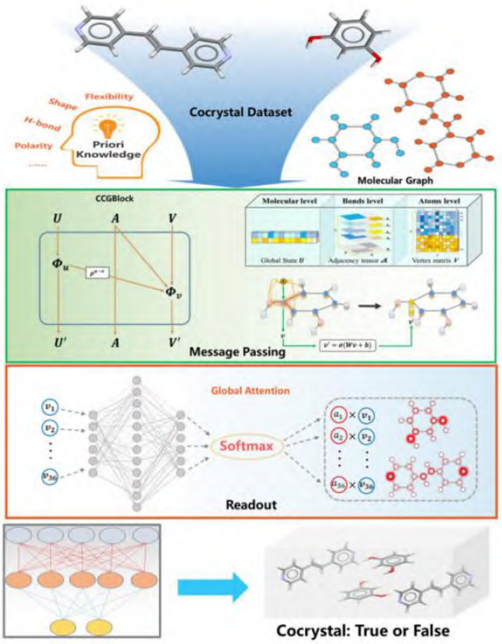

图 3　计算流程简图（结合分子图和先验知识来表示共晶，基于 GNN 框架（该框架由消息传递阶段和读出阶段组成，其中消息传递阶段中使用 CCGBlock 处理 3 种特征矩阵、读出阶段引入全局注意机制聚合节点信息）利用数据特征进行共晶预测）[73]

Fig. 3 Graphical abstract of the calculation process[73]

在材料性质预测方面，郑伟达等[74]从钙钛矿型材料自身的特征性能参数出发，分别采用随机森林（Random Forest，RF）、岭回归（Ridge Regression，RR）以及基于径向基核函数和线性核函数的支持向量回归（Support Vector Regression，SVR）等 4 种机器学习算法对钙钛矿材料自身特征性能参数进行预测，发现不同的机器学习算法对数据样本分布的敏感程度不同，因此预测不同的性能参数需要选择不同方法，如：RF 方法可有效预测钙钛矿材料的密度、带隙性能；RR 方法可实现对密度的预测；线性核函

数的 SVR 方法可以帮助预测形成能。上海大学施思齐团队[75]基于 Ge_xSe_{1-x} 玻璃结构的连续变化在某一组分处发生转折导致转变点温度的两阶段差异性和变化规律不一致特性，构建的两阶段支持向量回归模型，预测精度高达 97.62%，证明机器学习有助于实现玻璃转变点温度的快速精准预测。

在新材料设计与发现方面，Segler 的团队[76]将蒙特卡罗树搜索与神经网络相结合，搭建了一个用于设计化学合成路线的模型，相较于传统的计算机辅助搜索方法速度快 30 倍，由该模型设计出的合成路线可行性达到了化学专家的同等水平。孙宽等[77]根据文献数据建立了一个包含 1 719 个已知材料的数据库用于有机光伏(Organic Photovoltaic，OPV)电池的研究。针对电池的能量转换效率(Power Conversion Efficiency，PCE)值，通过多种分子不同类型的表达方式，例如结构图像、美国信息交换标准代码(American Standard Code for Information Interchange，ASCII)、描述符和分子指纹，成功训练出了预测准确率高达 81.76%的分类模型。在此基础上，10 种新型 OPV 合成材料被成功合成以验证模型的预测。Ren 等人[78]根据先前的金属玻璃研究报道的观察结果和物理化学理论的参数来训练机器学习模型，用于指导高通量实验，该模型在 Co-V-Zr 众多体系中找到了新的三元金属玻璃材料。同时利用实验观察和模型预测的三元成分定量差异来迭代训练模型，以提升模型预测准确性。

2 材料基因组技术在 IC 材料领域的应用

随着摩尔定律发展，IC 对于器件集成度和性能的要求不断提高，要想突破先进技术节点的限制，对于 IC 新材料的研究是不可避免的。但是 IC 材料种类繁多，除了常用的几类衬底材料，还包含了阻挡层材料、氧化物势垒层材料等制备成为各种 IC 功能器件的功能材料，光刻胶材料、抛光材料、湿化学品等工艺材料以及封装材料等。IC 材料制备工艺复杂，传统方法费时费力，材料基因组技术在研究 IC 材料领域具有巨大的潜力，例如高通量计算可用于 IC 新材料的设计预筛选，高通量实验和表征技术可用于 IC 薄膜材料的制备与表征。本节就此介绍一些材料基因组技术在 IC 材料领域或相关可借鉴领域初步的研究进展。

2.1 材料基因组技术在 IC 功能性材料的研究情况

带隙是评价 IC 材料电热输运性质、载流子迁移特性、发光能谱等性质的重要判据，在热电堆温度传感芯片、场效应晶体管、硅光器件等 IC 领域中应用广泛，因此需要研究者能够快速准确地预测特定材料的带隙值。徐永林等人[79]设计了一种带隙预测模型，基于已计算过的材料体系数据，利用单组元组分替换策略结合查重技术获得 356 个相似材料体系。同时借助机器学习技术构建了具有 77.73%预测精度的集成学习带隙预测模型 Ensemble，验证了 50 个相似材料体系的带隙值。

为了保持 IC 器件的持续微缩，高 k 介质材料因其具有高的相对介电常数、强击穿电场等优点，可以防止介质击穿，有效抑制栅极漏电流，逐渐地取代了传统的 SiO_2 栅介质。针对高 k 介质材料，Koinuma 等人[80]利用高通量脉冲激光沉积(Pulsed Laser Deposition，PLD)技术制备了 $HfO_2-Y_2O_3-Al_2O_3$ 三元化合物，并发现当 HfO_2: Y_2O_3: Al_2O_3 = 6: 3: 1 时，化合物具有稳定非结晶相结构和较高的介电常数(图 4)。

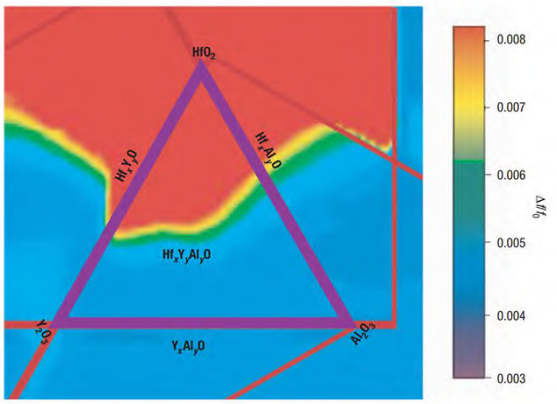

图 4 $HfO_2-Y_2O_3-Al_2O_3$ 三元组分介电图(较高的频移 $\Delta f/f_0$ 对应较大的介电常数)[80]

Fig. 4 The dielectric map of the $HfO_2-Y_2O_3-Al_2O_3$ ternary composition determined by microwave microscopy, with higher frequency shift ($\Delta f/f_0$) corresponding to larger dielectric constant[80]

Akira 等人[81]为基态晶体结构的氧化物设计了两种随机森林回归模型，基于对氧化物进行的密度泛函理论(Density Functional Theory，DFT)和密度泛函微扰理论(Density Functional Perturbation Theo-

ry，DFPT) 计算建立了包含各类化学成分和结构类型的氧化物介电常数数据库(图 5)，并使用数据库中的数据构建了电子和离子对静态介电常数贡献的预测模型(所用的递归特征消除过程见图 6)。

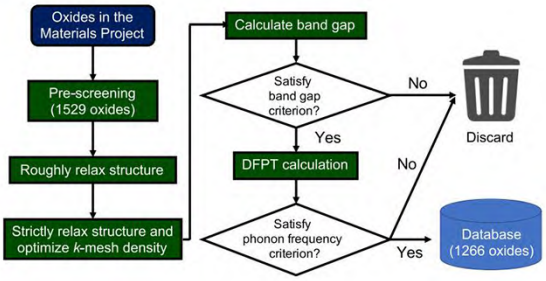

图 5　构建氧化物介电常数计算数据库的工作流[81]

Fig. 5 The workflow for constructing a computational database of the dielectric constants of oxides [81]

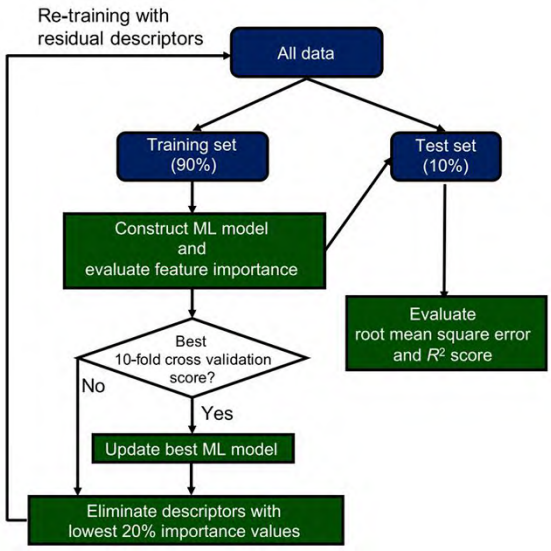

图 6　递归特征消除过程的示意图[81]

Fig. 6 The schematic diagram of the recursive feature elimination process [81]

合金材料在 IC 中扮演着重要角色，其中高熵合金(High-Entropy Alloys，HEAs) 作为多基元无序合金，偏向于形成简单固溶体相，因而具有高强度、耐腐蚀、耐氧化、迟滞扩散等优异性能，适用于 IC 中铜互连的扩散阻挡层。北京科技大学宿彦京团队[82]采用了一种基于遗传算法的通用算法来选择最佳的机器学习模型与材料描述符组合，并用于 HEAs 相结构预测，优化后的分类模型对固溶(Solid-Solution，SS) 和非固溶(Non-Solid-Solution，NSS) HEAs 的识别精度达到 88.7%，对体心立方(Body Centered Cubic，BCC)、面心立方(Face Centered Cubic，FCC) 和双相(Dual-Phase，DP) HEAs 的识别精度达到 91.3%(图 7)。该团队[83]还为 HEAs 设计了一种将机器学习模型与实验设计(Design of Experiments，DOE) 结合的材料设计策略(图 8)，便于在 Al-Co-Cr-Cu-Fe-Ni 系统中搜索具有大硬度的合金，仅通过 7 个实验就制造出了几种硬度比原始训练数据集中最佳值高 10%的合金。

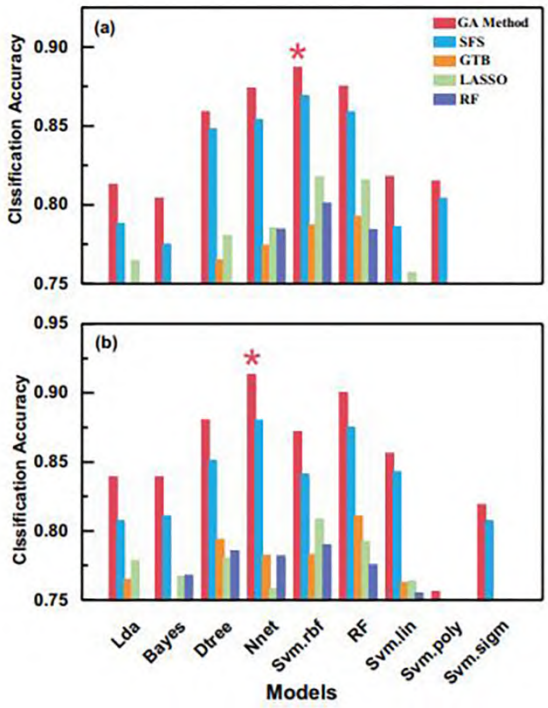

图 7　(a) 不同分类模型对 SS 和 NSS HEAs 的分类精度；(b) 不同分类模型对 FCC、BCC 和 DP HEAs 的分类精度[83]

Fig. 7 (a) Classification accuracy of different classification models for SS and NSS HEAs; (b) Classification accuracy of different classification models for FCC, BCC and DP HEAs [83]

2.2　材料基因组技术在 IC 工艺材料的研究情况

集成电路技术的不断发展，对先进工艺材料的研发生产提出了更高的要求。例如: 光刻工艺作为 IC 领域重要工艺之一，对应的光刻材料成为了光刻技术革新的重点。高端光刻材料需要能够满足光刻技术的高分辨率、器件结构平坦化、多次台阶刻蚀和深层次结构刻蚀等要求，但目前存在工艺配方设计复杂、成膜树脂和光敏剂设计合成困难和规模化原材料纯化技术提高困难等系列技术挑战。材料基因

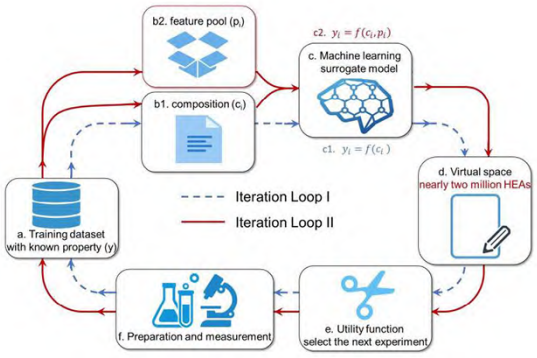

图 8 基于机器学习和 DOE 的迭代设计循环示意图，用于加速高熵合金设计[83]

Fig. 8 Schematic diagram of an iterative design loop based on machine learning and DOE for accelerating HEAs design[83]

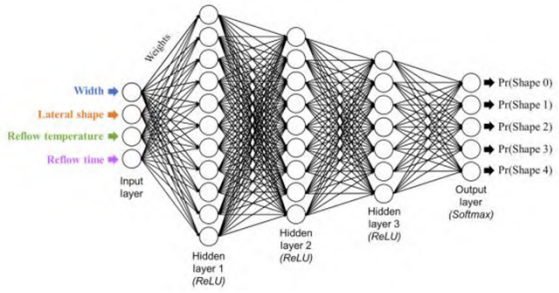

图 9 用于训练、预测回流后光刻胶形状类型的基于神经网络架构的分类模型[85]

Fig. 9 The neural network architecture-based classification model for training and predicting photoresist shape types after reflow[85]

组技术在光刻胶材料方面已有相关研究应用，例如 Lee 等人[84]将光引发剂合成与自动化机器学习框架相结合加速新型光引发剂设计，他们利用自动化机器学习算法从数百个以自动化方式生成、排序和验证的模型中快速识别最佳定量结构-特性关系模型。Chia 等人[85]基于常见光刻胶 SPR 220 回流的大量经验数据，以光刻胶初始形状、宽度、回流温度和时间为输入变量，开发了基于全连接神经网络的回流行为模型(图 9)，该模型可用于预测光刻胶回流后的形状类型及其完整的横截面高度轮廓。Tate 等人[86]利用 Hansen 模型光刻胶聚合物在给定溶剂混合物中的溶解度，借助高通量工作流收集和构建预测溶解度模型所需的实验数据，从初始混合到应用程序优化光刻胶溶解度。Tanaka 等人[87]使用石英晶体微量天平方法在四甲基氢氧化铵水性显影液中观察聚(4-羟基苯乙烯)(poly(4-hydroxystyrene)，PHS)的溶解动力学，结合使用多项式回归分析观察得到的数据，阐明了重均分子量和分子量分布对 PHS 薄膜溶解动力学的影响。

化学机械抛光技术(Chemical Mechanical Polishing，CMP)作为 IC 领域同样重要的关键工艺，要求抛光材料趋于多样化，且能适应更精确的 CMP 工艺控制，这对抛光液和抛光垫等核心材料的性能要求更高。但是抛光液配方复杂，抛光垫基材配方及沟槽设计加工难度高，亦使得抛光材料研发生产存在难点。Terrell[88]研究开发了一种集成计算建模方法，在不使用经验拟合参数的情况下考虑了流体力学、接触力学和粒子动力学等重要的机械相互作用，此建模方法可用于预测实际样品形貌的表面磨损。张九思等[89]首次提出借助深度学习方法中的残差卷积神经网络(Residual Convolutional Neural Networks，ResCNN)预测 CMP 工艺材料去除率(Material Removal Rate，MRR)的数据驱动方法(图 10)，其预测性能优于现有文献的其他方法。何平等人[90]将粒子群优化算法和支持向量机相结合，构建了 CMP 抛光液组分优化模型，以 CMP 抛光液组分及 PH 值作为输入，Cu 去除速率为输出，对 Cu-CMP 实验数据进行训练和预测，可优化抛光液配比。

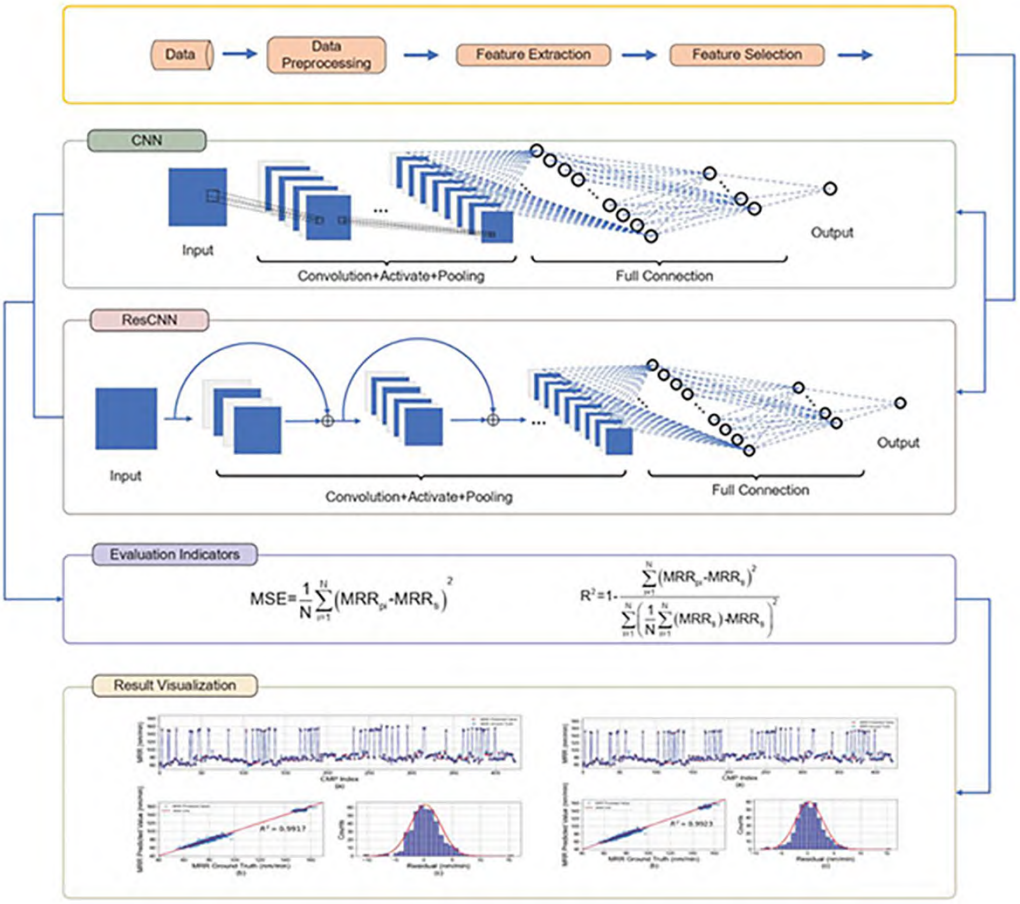

图 10　基于 ResCNN 方法的 CMP 材料去除率预测网络[89]

Fig. 10 The prediction network for CMP-MRR based on ResCNN method [89]

湿化学品是 IC 湿法工艺必不可少的基础化工材料，可用于清洗、刻蚀、合成、表面处理等，这对材料的纯度和洁净度要求非常高，但是湿化学品的纯化技术、检测技术和配方设计技术都亟待提升。材料基因组技术已成为湿化学品研究的有效手段，例如 Zabula 等人[91]报道了一种高通量实验方法，用于快速测量分子抑制剂的腐蚀保护活性，并筛选出了由羧酸盐和镧系元素组成的新型金属缓蚀剂。Kohji 等人[92]将“组合化学”和基于自动化设备进行快速检测的“高通量筛选”结合，设计制造了具有 96 条平行产线的高通量筛选反应器，以优化用于甲醇合成的铜锌催化剂，并使用反应器获得的数据集构建神经网络模型以筛选催化剂组成。Putt 等人[93]研发了基于滴定化学反应的高通量微量滴定板检测方法，用于连续测定过氧化氢和溶液中过氧化物的总量，适用于过氧化氢的快速定量测定。电子封装是 IC 制造的一个重要环节，对封装的基片和密封材料力学性能、电绝缘性、化学稳定性和热传导性等性能有着较高要求。目前主要的封装材料为陶瓷与塑料，材料基因组技术在此方面也有相关研究应用。Shin 等人[94]开发了新的微力学方法，使得变分渐进法能够均匀化异质微结构，并通过多尺度材料基因组模型预测碳化硅陶瓷的弹性特性与断裂强度。哈尔滨工业大学谭惠丰团队[95]利用全原子分子动力学研究了热固性聚合物的纳米级交联过程，并通过拉伸实验进行了验证，展示了对热固性聚合物纳米交联反应的可靠理解，有利于热固性塑料的设计与特性调控。华东理工大学林嘉平团队[96]为了实现以乙炔封端聚酰亚胺(Acetylene-Terminated Polyimides，ATPI)作为添加剂增强含硅芳炔(Silicon-Containing Arylacetylene，PSA)树脂韧性的目的，建立了材料基因组方法来设计和筛选出 ATPI(图 11)，并

考察了 ATPI 含量与预聚合温度对 PSA 树脂性能的影响，该工作为加速设计和筛选高性能聚合物材料提供了新方法。

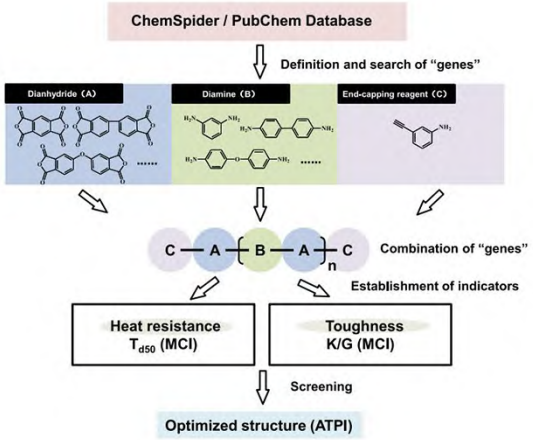

图 11 ATPI 的设计与快速筛选而开发的程序(A、B 和 C 基因分别为二酐单体、二胺单体和封端剂，Td50 和 K/G 由分子连接指数计算得出)[94]

Fig. 11 Procedures developed for the design and rapid screening of ATPI [94]

面对 IC 领域工艺材料的研发困境，材料基因组或可成为工艺材料研发的新范式，成为技术突破的重点。工艺材料基因组技术的主要研究思路应如下：借助高通量计算与高通量实验获取海量数据，建立起光刻胶、抛光液等工艺材料的专业材料数据库；然后引入机器学习算法等技术对专业数据库中的数据进行挖掘，获得材料微观结构、化学组成、材料合成方法以及加工工艺等与材料性能之间的潜在联系；依据机器学习得到的关系，进行新工艺材料的设计与合成，并利用高通量实验表征进行材料性质验证；验证所得数据录入数据库，进行多次迭代设计获得更为符合 IC 工艺要求的新材料。

中国科学院上海微系统与信息技术研究所的相关团队围绕材料基因组技术这一新的研究范式，将高通量计算-高通量工艺-数据库技术融合，打造集成电路材料基因组全技术平台。以具体材料的应用为导向，将计算、实验和数据库技术贯穿整个研究。在面向 5G 高性能滤波器的掺杂氮化铝研究中，优化掺杂算法，开发了第一性原理高通量计算流，目前已计算超过 150 种掺入氮化铝的元素组合，筛选出 9 种符合谐振器件性能的掺杂组合，为制备高压电性能的掺杂氮化铝薄膜提供了理论依据。实验方面，该团队利用其在 2020 年成功研制的 8 英寸多靶共溅射高通量物理气相沉积(Physical Vapor Deposition，PVD)设备[97]，系统研究了不同比例钪掺杂氮化铝薄膜的工艺，制备得到掺钪浓度为 20%的 AlN 薄膜，结合同步辐射 XRD 开发了高通量压电性能表征系统对制备的薄膜压电性能进行了表征，进一步探究了钪掺杂浓度对 AlN 压电薄膜材料特性及器件性能的影响[98]。该团队将其研发过程中产生的大量数据包括高通量计算数据、高通量 PVD 制备及表征实验数据以及采用 MES 自动采集的工艺参数、设备状态数据等，存储至集成电路芯片关键材料数据库，为未来的机器学习、数据分析等提供了数据基础，从而有利于加速芯片关键材料的研发。

3 总结与展望

材料基因组技术作为突破传统材料实验和设计方法的新型技术手段，得到了高度重视和发展，体现了材料科学研究在新时期的重要突破与创新，是解决现代材料工业众多关键瓶颈的重要技术，已经在一定程度上缩减了现代材料工业的研发周期与研发成本。材料基因组的三大要素在技术发展过程中进行了深层次的融合，以高通量计算替代部分重复性实验，以高通量实验反过来验证计算结果。同时不论是高通量计算还是高通量实验，均为材料大数据挖掘与分析提供了海量基础数据，推动材料信息学发展。目前，材料基因组技术的发展还是面临着一些问题，如高通量实验设备结构复杂、研发成本高；高通量计算对计算机算力的高要求和对计算方法精度和耗时的要求；材料大数据对材料数据高质量、高数量的需求等等。这些问题涉及到材料、物理、机械、计算机、信息技术等众多不同领域，说明材料基因组技术是一种典型的“多学科交叉，多领域融合”的技术。

集成电路材料作为集成电路产业链中细分领域最多的一环，贯穿集成电路制造的晶圆制造、前道工艺(芯片制造)和后道工艺(封装)整个过程。后摩尔时代，集成电路技术每次提升都对材料的性能提出新的要求，材料技术的每次发展也都为集成电路新结构、新器件的开发创造新的空间。据估算，材料对先进制程逻辑芯片性能提升的贡献已超过六成，掌握集成电路材料核心技术与知识产权就占领了集

成电路技术高地[99]。面对目前急需解决的问题及前沿颠覆性材料的新特点，包括①涉及的元素多、组分复杂、机理复杂、工艺繁杂、验证周期长等传统材料研究面临的研发问题以及②与 CMOS 工艺不匹配、材料合成的均一性、材料本身的稳定性、易制备性等面向应用的众多问题，集成电路材料的研究不仅需要探索材料本身的机理，更要综合考虑材料的制备包括工艺环境和设备、材料的应用比如芯片性能等，及时找到并反馈 IC 材料对工艺和器件的影响，以应用为导向促进材料创新。

材料基因组技术变革传统的低效率材料研发方式，寻求研发新范式以推动集成电路材料技术高质量发展，将集成多学科进行材料高通量筛选的材料基因组技术运用在集成电路的材料研发中，同时推动建设高通量计算模拟、高通量样品制备与表征以及材料数据库的集成平台，基于材料本身的组成成分、结构、性质，制备工艺的参数、芯片性能等海量实验数据，构建有效的材料成分-工艺-性能间的关系，发展更高效的研究方法，有助于加快 IC 材料的研发周期。集成电路材料基因组需要培养具有传统材料科学领域经验与知识以及多学科交叉背景的研究人才，还要求其掌握集成电路材料的研发-制备-应用机制。探求材料对于集成电路工艺乃至器件性能、终端应用的影响因素，可以加快材料基因组技术在集成电路材料领域的研究，提升国家在集成电路领域的竞争力。

参考文献:

[1] 向　勇，闫宗楷，朱焱麟，等. 材料基因组技术前沿进展[J]. 电子科技大学学报，2016，45(4)：634-649.

[2] White A. The materials genome initiative: One year on [J]. Mrs Bulletin，2012，37(8)：715-716.

[3] 关永军，陈　柳，王金三. 材料基因组技术内涵与发展趋势[J]. 航空材料学报，2016，36(3)：71-78.

[4] 刘梓葵. 关于材料基因组的基本观点及展望[J]. 科学通报，2013，58(35)：3618-3622.

[5] 王兴艳，宋显珠. 我国新材料产业自主创新能力建设反思[J]. 新材料产业，2014，(4)：60-62.

[6] Materials Genome Initiative for Global Competitiveness [EB/OL]. 2001. http://www.whitehouse.gov/sites/default/files/microsites/ostp/materials_genome_initiative_final.pdf.

[7] Holdren J P. Report on the president on ensuring American leadership in advanced manufacturing [R]. Washington D. C.: PCAST，2011.

[8] Adamick J. A national strategic plan for advanced manufacturing [J]. Journal of Applied Clinical Pediatrics，2012，35(5)：267-268.

[9] 汪　洪，向　勇，项晓东，等. 材料基因组——材料研发新模式[J]. 科技导报，2015，33(10)：13-19.

[10] Montoya J H，Persson K A. A high-throughput framework for determining adsorption energies on solid surfaces [J]. npj Computational Materials，2017，3(1)：1-4.

[11] Mamun O，Winther K T，Boes J R，et al. High-throughput calculations of catalytic properties of bimetallic alloy surfaces [J]. Scientific data，2019，6(1)：1-9.

[12] Kirklin S，Meredig B，Wolverton C. High-throughput computational screening of new Li-ion battery anode materials [J]. Advanced Energy Materials，2013，3(2)：252-262.

[13] Aykol M，Kim S，Hegde V I，et al. High-throughput computational design of cathode coatings for Li-ion batteries [J]. Nature communications，2016，7(1)：1-12.

[14] 骆　军，何世洋，李志立，等. 热电材料高通量实验制备与表征方法[J]. 无机材料学报，2019，34(3)：247-259.

[15] Carrete J，Li W，Mingo N，et al. Finding unprecedentedly low-thermal-conductivity half-Heusler semiconductors via high-throughput materials modeling [J]. Physical Review X，2014，4(1)：011019.

[16] 汪　洪，项晓东，张澜庭. 数据 + 人工智能是材料基因工程的核心[J]. 科技导报，2018，36(14)：15-21.

[17] 宿彦京，付华栋，白　洋，等. 中国材料基因工程研究进展[J]. 金属学报，2020，56(10)：1313-1323.

[18] 杨小渝，王　娟，任　杰，等. 支撑材料基因工程的高通量材料集成计算平台[J]. 计算物理，2017，34(6)：697-704.

[19] Yang X，Wang Z，Zhao X，et al. MatCloud，a high-throughput computational materials infrastructure: Present，future visions，and challenges [J]. Chinese Physics B，2018，27(11)：110301.

[20] 陈晓婕，杨小渝，张　平，等. 基于 web 的二元化合物结构推演自动流程实现[J]. 科研信息化技术与应用，2016，7(1)：67-75.

[21] 杨小渝，任　杰，王　娟，等. 基于材料基因组计划的计算和数据方法[J]. 科技导报，2016，34(24)：62-67.

[22] 王　军，杨小渝，王宗国，等. 在晶体表面寻找最佳吸附位点的高通量算法[J]. 科研信息化技术与应用，2016，7(5)：3-12.

[23] Xi L, Pan S, Li X, et al. Discovery of high-performance thermoelectric chalcogenides through reliable high-throughput material screening [J]. Journal of the American Chemical Society, 2018, 140(34): 10785-10793.

[24] Wang G, Peng L, Li K, et al. ALKEMIE: An intelligent computational platform for accelerating materials discovery and design [J]. Computational Materials Science, 2021, 186: 110064.

[25] 王冠杰,李开旗,彭力宇,等. 高通量自动流程集成计算与数据管理智能平台及其在合金设计中的应用 [J]. 金属学报, 2021, 58(1): 75-88.

[26] Kresse G, Hafner J. Ab initio molecular dynamics for liquid metals [J]. Phys Rev B Condens Matter, 1993, 48(1): 13115-13118.

[27] Plimpton S. Fast parallel algorithms for short-range molecular dynamics [J]. Journal of computational physics, 1995, 117(1): 1-19.

[28] Lukas H, Fries S G, Sundman B. Computational Thermodynamics. The Calphad Method [M]. Cambridge: Cambridge University Press, 2007: 79-160.

[29] Otero-de-la-Roza A, Abbasi-Pérez D, Luaña V. Gibbs2: A new version of the quasiharmonic model code. II. Models for solid-state thermodynamics, features and implementation [J]. Computer Physics Communications, 2011, 182(10): 2232-2248.

[30] Tegeler M, Shchyglo O, Kamachali R D, et al. Parallel multiphase field simulations with OpenPhase [J]. Computer physics communications, 2017, 215: 173-187.

[31] Madsen G K H, Singh D J. BoltzTraP. A code for calculating band-structure dependent quantities [J]. Computer Physics Communications, 2006, 175(1): 67-71.

[32] Jain A, Ong S P, Hautier G, et al. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation [J]. APL Materials, 2013, 1(1): 011002.

[33] Chaudhary V, Katyal P, Kumar A, et al. Optical and structural properties of silica nanoparticles using material design software: MedeA [M]. Recent Trends in Materials and Devices. Springer, Cham, 2017: 65-72.

[34] Supka A R, Lyons T E, Liyanage L, et al. AFLOWπ: A minimalist approach to high-throughput ab initio calculations including the generation of tight-binding hamiltonians [J]. Computational Materials Science, 2017, 136: 76-84.

[35] Restuccia P, Levita G, Wolloch M, et al. Ideal adhesive and shear strengths of solid interfaces: A high throughput ab initio approach [J]. Computational Materials Science, 2018, 154: 517-529.

[36] Mathew K, Singh A K, Gabriel J J, et al. MPInterfaces: A Materials Project based Python tool for high-throughput computational screening of interfacial systems [J]. Computational Materials Science, 2016, 122: 183-190.

[37] Sarikurt S, Kocabaş T, Sevik C. High-throughput computational screening of 2D materials for thermoelectrics [J]. Journal of Materials Chemistry A, 2020, 8(37): 19674-19683.

[38] Gao J, Chu G, He M, et al. Screening possible solid electrolytes by calculating the conduction pathways using Bond Valence method [J]. Science China Physics, Mechanics & Astronomy, 2014, 57(8): 1526-1536.

[39] Chen L Q. Research progress in cathode materials of Li-ion battery [J]. Battery Bimonthly, 2002, 32(S1): 6-8.

[40] 肖睿娟, 李 泓, 陈立泉. 基于材料基因组方法的锂电池新材料开发 [J]. 物理学报, 2018, 67(12): 291-299.

[41] 潘 都,戚兴国,刘丽露,等. 钠离子电池正负极材料研究新进展 [J]. 硅酸盐学报, 2018, 46(4): 479-498.

[42] Wang D, Zhang X, Xiao R, et al. Electrochemical performance of Li-rich Li $[Li_{0.2}Mn_{0.56}Ni_{0.17}Co_{0.07}]O_2$ cathode stabilized by metastable Li_2SiO_3 surface modification for advanced Li-ion batteries [J]. Electrochimica Acta, 2018, 265: 244-253.

[43] Zhang T, Jiang Y, Song Z, et al. Catalogue of topological electronic materials [J]. Nature, 2019, 566(7745): 475-479.

[44] Shen Z H, Wang J J, Jiang J Y, et al. Phase-field modeling and machine learning of electric-thermal-mechanical breakdown of polymer-based dielectrics [J]. Nature communications, 2019, 10(1): 1-10.

[45] Kennedy K, Stefansky T, Davy G, et al. Rapid method for determining ternary-alloy phase diagrams [J]. Journal of Applied Physics, 1965, 36(12): 3808-3810.

[46] Hanak, Joseph J. The "Multiple-Sample Concept" in Materials Research: Synthesis, Compositional Analysis and Testing of Entire Multicomponent Systems [J]. Journal of Materials Science, 1970, 5(11): 964-971.

[47] Pirrung M C, Read J L, Fodor S P A, et al. Large scale photolithographic solid phase synthesis of polypeptides and receptor binding screening thereof: U.S. 5143854 [P]. 1992-9-1.

[48] Xiang X D, Sun X, Briceno G, et al. A combinatorial approach to materials discovery [J]. Science, 1995, 268(5218): 1738-1740.

[49] Potyrailo R, Rajan K, Stoewe K, et al. Combinatorial and high-throughput screening of materials libraries: review of state of the art [J]. ACS combinatorial science, 2011, 13(6): 579-633.

[50] Li M X, Zhao S F, Lu Z, et al. High-temperature bulk metallic glasses developed by combinatorial methods [J]. Nature, 2019, 569(7754): 99-103.

[51] Li M X, Sun Y T, Wang C, et al. Data-driven discovery of a universal indicator for metallic glass forming ability [J]. Nature materials, 2022, 21(2): 165-172.

[52] Liu Y, Wang J, Xiao B, et al. Accelerated development of hard high-entropy alloys with data-driven high-throughput experiments. Journal of Materials Informatics, 2022, 2(1): 3.

[53] Regalado E L, Pereira T, Shevlin M, et al. Organic chemistry. Nanomole-scale high-throughput chemistry for the synthesis of complex molecules [J]. Science (New York, NY), 2014, 347(6217): 49-53.

[54] Perera D, Tucker J W, Brahmbhatt S, et al. A platform for automated nanomole-scale reaction screening and micromole-scale synthesis in flow [J]. Science, 2018, 359(6374): 429-434.

[55] Gregoire J M, McCluskey P J, Dale D, et al. Combining combinatorial nanocalorimetry and X-ray diffraction techniques to study the effects of composition and quench rate on Au-Cu-Si metallic glasses [J]. Scripta Materialia, 2012, 66(3-4): 178-181.

[56] Mccluskey P J, Vlassak J J. Combinatorial nanocalorimetry [J]. Journal of Materials Research, 2010, 25(11): 2086-2100.

[57] Kubota H, Takahashi R, Kim T W, et al. Combinatorial synthesis and luminescent characteristics of $RECa_4O(BO_3)_3$ epitaxial thin films [J]. Applied surface science, 2004, 223(1-3): 241-244.

[58] Wei T, Xiang X D, Wallace-Freedman W G, et al. Scanning tip microwave near-field microscope [J]. Applied Physics Letters, 1996, 68(24): 3506-3508.

[59] Gao C, Duewer F, Xiang X D. Quantitative microwave evanescent microscopy [J]. Applied physics letters, 1999, 75(19): 3005-3007.

[60] Xiang X D, Takeuchi I. Combinatorial Materials Synthesis [M]. New York: Dekker, 2003: 318-350.

[61] Wu H, Fan G, Huang M, et al. Deformation behavior of brittle/ductile multilayered composites under interface constraint effect [J]. International Journal of Plasticity, 2017, 89: 96-109.

[62] Michael J R, Nakakura C Y, Garbowski T, et al. High-Throughput SEM via Multi-Beam SEM: Applications in Materials Science [J]. Microscopy & Microanalysis, 2015, 21(S3): 697-698.

[63] Takeuchi I, Yang W, Chang K S, et al. Monolithic multichannel ultraviolet detector arrays and continuous phase evolution in $Mg_xZn_{1-x}O$ composition spreads [J]. Journal of Applied Physics, 2003, 94(11): 7336-7340.

[64] Reeves W H, Skryabin D V, Biancalana F, et al. Transformation and control of ultra-short pulses in dispersion-engineered photonic crystal fibres [J]. Nature, 2003, 424(6948): 511-515.

[65] Jin Z, Fukumura T, Kawasaki M, et al. High throughput fabrication of transition-metal-doped epitaxial ZnO thin films: A series of oxide-diluted magnetic semiconductors and their properties [J]. Applied Physics Letters, 2001, 78(24): 3824-3826.

[66] 吴圣川，吴正凯，康国政，等. 先进材料多维多尺度高通量表征研究进展 [J]. 机械工程学报，2021，57(16): 37-65.

[67] 王海舟，汪　洪，丁　洪，等. 材料的高通量制备与表征技术 [J]. 科技导报，2015，33(10): 31-49.

[68] 李姿昕，张　能，熊　斌，等. 材料科学数据库在材料研发中的应用与展望 [J]. 数据与计算发展前沿，2020，2(2): 78-90.

[69] 牛程程，李少波，胡建军，等. 机器学习在材料信息学中的应用综述 [J]. 材料导报，2020，34(23): 23100-23108.

[70] Liang Y, Chen M, Wang Y, et al. A universal model for the formation energy prediction of inorganic compounds [J]. arXiv preprint arXiv: 2108.00349, 2021.

[71] Liu Y, Niu C, Wang Z, et al. Machine learning in materials genome initiative: A review [J]. Journal of Materials Science & Technology, 2020, 57: 113-122.

[72] Islam N, Huang W, Zhuang H L. Machine learning for phase selection in multi-principal element alloys [J]. Computational Materials Science, 2018, 150: 230-235.

[73] Jiang Y, Guo J, Liu Y, et al. A Universal Deep Learning Framework based on Graph Neural Network for Virtual Co-Crystal Screening [J]. ChemRxiv preprint: 2021.

[74] 郑伟达，张惠然，胡红青，等. 基于不同机器学习算法的钙钛矿材料性能预测 [J]. 中国有色金属学报，2019，29(4): 803-809.

[75] Liu Y, Wu J, Yang G, et al. Predicting the onset temperature (Tg) of Ge_xse_{1-X} glass transition: a feature selection based two-stage support vector regression method [J]. Sci-

ence Bulletin, 2019, 64(16): 1195-1203.

[76] Segler M H S, Preuss M, Waller M P. Planning chemical syntheses with deep neural networks and symbolic AI [J]. Nature, 2018, 555(7698): 604-610.

[77] Sun W, Zheng Y, Yang K, et al. Machine learning - assisted molecular design and efficiency prediction for high-performance organic photovoltaic materials [J]. Science advances, 2019, 5(11): eaay4275.

[78] Ren F, Ward L, Williams T, et al. Accelerated discovery of metallic glasses through iteration of machine learning and high - throughput experiments [J]. Science advances, 2018, 4(4): eaaq1566.

[79] 徐永林,王香蒙,李 鑫,等. 基于高通量计算及机器学习的新材料带隙预测 [J]. 中国科学: 技术科学, 2019, 49(01): 44-54.

[80] Koinuma H, Takeuchi I. Combinatorial solid-state chemistry of inorganic materials [J]. Nature materials, 2004, 3(7): 429-438.

[81] Takahashi A, Kumagai Y, Miyamoto J, et al. Machine learning models for predicting the dielectric constants of oxides based on high-throughput first-principles calculations [J]. Physical Review Materials, 2020, 4(10): 103801.

[82] Zhang Y, Wen C, Wang C, et al. Phase prediction in high entropy alloys with a rational selection of materials descriptors and machine learning models [J]. Acta Materialia, 2020, 185: 528-539.

[83] Wen C, Zhang Y, Wang C, et al. Machine learning assisted design of high entropy alloys with desired property [J]. Acta Materialia, 2019, 170: 109-117.

[84] Lee W J, Kwak H S, Lee D, et al. Design and synthesis of novel oxime ester photoinitiators augmented by automated machine learning [J]. Chemistry of Materials, 2022, 34(1): 116-127.

[85] Chia C, Martis J, Jeffrey S S, et al. Neural network-based model of photoresist reflow [J]. Journal of Vacuum Science & Technology B, Nanotechnology and Microelectronics: Materials, Processing, Measurement, and Phenomena, 2019, 37(6): 061604.

[86] Tate M P, Cutler C, Sakillaris M, et al. How to design a good photoresist solvent package using solubility parameters and high-throughput research [C]. Advances in Patterning Materials and Processes XXXI. International Society for Optics and Photonics, 2014, 9051: 90510R.

[87] Tanaka N, Watanabe K, Matsuoka K, et al. Analysis of dissolution kinetics of poly (4-hydroxystyrene) with different molecular weight distributions in alkaline aqueous solution using machine learning [C]. Advances in Patterning Materials and Processes XXXVIII. International Society for Optics and Photonics, 2021, 11612: 116120R.

[88] Terrell E J, Higgs C F. A particle-augmented mixed lubrication modeling approach to predicting chemical mechanical polishing [J]. Journal of Tribology, 2009, 131(1): 012201.

[89] Zhang J, Jiang Y, Luo H, et al. Prediction of material removal rate in chemical mechanical polishing via residual convolutional neural network [J]. Control Engineering Practice, 2021, 107: 104673.

[90] 何 平,罗 萌,韩欣玉,等. 基于 pso-Svm 模型的 cu Cmp 抛光液组分优化 [J]. 微电子学, 2020, 50(5): 694-698.

[91] Zabula A V, Dey S, Robinson J R, et al. Screening of molecular lanthanide corrosion inhibitors by a high-throughput method [J]. Corrosion Science, 2020, 165: 108377.

[92] Omata K, Watanabe Y, Umegaki T, et al. Catalyst development for methanol synthesis using parallel reactors for high-throughput screening based on a 96 well microplate system [J]. Journal of the Japan Petroleum Institute, 2003, 46(5): 328-334.

[93] Putt K S, Pugh R B. A high-throughput microtiter plate based method for the determination of peracetic acid and hydrogen peroxide [J]. Plos one, 2013, 8(11): e79218..

[94] Dong X, Shin Y C. Multi-scale genome modeling for predicting fracture strength of silicon carbide ceramics [J]. Computational Materials Science, 2018, 141: 10-18.

[95] Kang J, Wang C, Li D, et al. Nanoscale crosslinking in thermoset polymers: a molecular dynamics study [J]. Physical Chemistry Chemical Physics, 2015, 17(25): 16519-16524.

[96] Gao G, Zhang S, Wang L, et al. Developing highly tough, heat-resistant blend thermosets based on silicon-containing arylacetylene: A material genome approach [J]. ACS Applied Materials & Interfaces, 2020, 12(24): 27587-27597.

[97] 俞文杰, 李卫民, 朱 雷, 等著. 集成电路材料基因组技术 [M]. 北京: 电子工业出版社, 2022: 20-21.

[98] 朱宇波, 母志强, 陈玲丽, 等. 掺钪 AlN 材料特性及 HBAR 器件性能研究 [J]. 压电与声光, 2022, 44(2): 299-303.

[99] 冯 黎, 朱 雷. 中国集成电路材料产业发展现状分析 [J]. 功能材料与器件学报, 2020, 26(03): 191-196.