

目录

前言	1
一、人类表型组：科学概念与大科学计划	3
(一) 表型与人类表型组	3
(二) 国际人类表型组计划	5
1、发展历程	5
2、国际人类表型组研究协作组 (IHPC)	7
3、中国国际人类表型组研究协作组 (HPCC)	10
三、人类表型组：全球布局与发展态势	12
(一) 中国：整合并推进国际人类表型组大科学计划	12
(二) 美国：人类表型组是医疗健康领域未来发展机遇	14
(三) 欧盟：完善基因组 - 表型组档案建设	15
(四) 英国：将表型组学列入生命科学发展规划	16
(五) 澳大利亚：建设全球领先的分子表型组平台	17
(六) 印度：发起沙漠表型组计划	18
四、人类表型组：年度研究进展	19
(一) 国际人类表型组计划 (一期) 进展	19
1、概况	19
2、首张人类表型组全景导航图	21
3、质量控制与标准化	22
(二) 新冠肺炎的表型组学研究	22
1、明确新冠感染后表型转化特征	23
2、评估新冠患者功能恢复并预测疾病发展	23
3、研发新冠肺炎诊断与治疗表型组学新技术	24
(三) 深入理解基因 - 表型间关联	24
1、解析亚洲热带原住民遗传起源和表型趋同演化机制	24
2、解析皮肤菌群表型助力皮肤疾病治疗	25
3、X 连锁遗传性视网膜劈裂症深度表型测量	25

- (四) 表型跨尺度关联发现疾病风险.....25
 - 1、阐明阿尔兹海默症表型关联因素.....25
 - 2、母亲压力与儿童期呼吸系统疾病发病风险相关.....25
 - 3、食物组与慢性肾脏疾病表型.....26
 - 4、过敏性皮炎严重程度的多因素分析.....26
 - 5、揭示肝脏疾病发病风险.....26
- (五) 疾病表型组学研究进展.....27
 - 1、人群深度表型测量实现健康到疾病转化的早诊.....27
 - 2、心血管疾病的表型组学研究.....27
 - 3、无创早期检测阿尔兹海默症.....28
 - 4、血浆临床表型预测坏死性软组织感染.....28
 - 5、为多种癌症患者分层管理提供策略.....28
 - 6、辅助制定呼吸系统临床治疗方案.....29
 - 7、多发硬化症认知表型助力疾病分型.....29
- (六) 新技术研发.....29
 - 1、人工智能 CARE 系统筛选眼底异常表型.....29
 - 2、基于血液液体活检的红外分子指纹癌症检测技术.....30
 - 3、自动化智能表型可塑性平台.....30
 - 4、超极化 ^{129}Xe 磁共振成像测量预测儿童哮喘发展情况.....30
- 五、人类表型组：总结与展望.....31**

前言

人类表型组研究的重要性已成为学界共识。通过精确测量人类表型并进行综合分析，可以系统解构表型之间关系，构建表型关联网络，了解并明确基因 - 表型 - 环境之间、宏观和微观表型之间的多维度和跨尺度相关性。近年来，人类表型组计划被公认为是后基因组时代生命科学的战略制高点和原始创新源，是引领生命科学研究范式新变革和生物医药产业发展新突破的关键所在。

《人类表型组研究进展报告》是中国人类表型组研究协作组（HPCC）、国际人类表型组研究协作组（IHPC）、“一带一路”人类表型组联合研究中心、上海国际人类表型组研究院、复旦大学人类表型组研究院以及中国科学院上海营养与健康研究所战略情报研究部等机构联合定期发布的系列领域进展报告。《人类表型组研究进展报告 2021》聚焦表型组国际布局和领域研究热点，总结了 2021 年国际人类表型组计划最新发展，以期为研判人类表型组学未来发展方向及关键布局提供支撑。

我国自 2016 年起就前瞻布局人类表型组国际大科学计划，2017 年“国际人类表型组计划（一期）”作为首批上海市市级科技重大专项启动。在专项支持下，金力院士领衔上海核心科研力量，凝聚国内外数十个顶尖科学家团队，牵头推动人类表型组大科学计划在国际科学界先行实质性启动。由中国、美国和英国科学家共同发起，全球 20 余个国家科学家共同参与的国际人类表型组计划在 2021 年顺利推进。在国际人类表型组研究协作组带动下，人类表型组领域正加快布局，多国成立了国家级表型组研究中心，发起了一批表型组学领域的研究计划，促进了全球人类表型组研究的发展。

中国科学家团队积极投入并引领人类表型组研究：张江复旦国际创新中心建成了全球首个跨尺度多维度人类表型精密测量平台；复旦大学人类表型组研究院建立了全球首个深度表型全景测量的千人队列；上海国际人类表型组研究院团队已研发建成全球首个人类表型组大数据处理云平台；复旦大学和中科院团队已绘制完成全球首版人类表型组导航图。

2021 年度人类表型组测量方法、质量控制与标准化工作、表型组大数

据计算与分析等关键领域都产生了突破性成果。人类表型组研究深入探索基因 - 环境 - 表型到个体健康的表型间跨尺度关联，发现疾病起源及多样性，研发并应用新技术和方法，必将创造改善人类健康的新范式。

《人类表型组研究进展报告 2021》得到了国际人类表型组研究协作组的 23 位专家、中国人类表型组研究协作组的 91 位专家的大力支持与指导，特此致谢！

指导专家：

金力、Leroy Hood、田梅、杨忠、徐萍、张梦翰、许俊

报告作者：

刘晗、陶韡烁、孙学会

工作小组：

夏鑫、王嘉雯、裴浩宇、陈佳佳、陈舜泠

一、人类表型组：科学概念与大科学计划

（一）表型与人类表型组

基因和环境相互作用决定人体特征，人体特征即表型（Phenotype）。表型组（Phenome）是生物体从胚胎发育到出生、成长、衰老乃至死亡过程中，形态特征、功能、行为、分子组成规律等所有生物、物理和化学特征的集合。



图 1 表型组定义

近年来，国际生命科学界对表型组关注与重视日渐增强，越来越多的科学家认识到人类表型组将接棒人类基因组，成为后基因组时代的战略制高点。精密测量人体表型，全景解析人类表型组，将系统解构表型之间强关联，构建表型网络，打通宏观表型与微观表型间多维度、跨尺度关联，明确表型间跨尺度关联，是拓展生命科学领域知识和应用技术有效手段，将助力生物医学研究、促进科技创新发展、保障人民生命健康。

表 1、开展表型组研究的科学界共识^[1]

时间	重要观点	来源
2003	首次提出“人类表型组计划”一词，认为当时表型检测方法可能不足以优化利用基因型数据，呼吁国际力量共建表型组数据库和分享表型组数据，深度整合表型组信息并开发新的分析方法	Nature Genetics, Vol 34 (1):15-21
2005	在提出的当代 125 个最具挑战性科学问题中，排在第三、四和十四位的问题分别是：人类基因为何如此之少、遗传变异与人类健康、什么基因改变造成了人类独特性，其核心在于解析基因与表型之间的关联	Science, Vol 309 (5734):558-558
2012	美国人类遗传学学会召开主题“为人类表型组计划做准备”的分论坛，讨论了表型和基因型数据分享的需求和重要性，认为亟需优化临床表型数据的收集系统和工具等	Human Mutation, Vol. 34 (4):661-666
2013	相对于人类基因组而言，表型组信息如个体生理和行为等特征的准确描述非常欠缺，表型相关的术语统一和资源整合对于表型研究和疾病探讨至关重要	Nature, 494 (7438):416-419
2016	美国科学基金会（National Science Foundation, NSF）将“理解生命的规律”列为建议关注的六大研究前沿领域第三位，核心内容是从基因型预测生物体表型，理解生物体和环境之间的互作机制，应对传染性疾病和气候变化带来的持续影响	Science, Vol 352 (6287):755-756 NSF 官网 (https://www.nsf.gov/)
2018	国际顶级科学家们展望了改变未来生命科学研究的六大技术，“连接基因和表型”被列入其中，尤其对于医学研究而言，需借助该技术从机制视角将遗传变异转化到治疗应用中	Nature, Vol 553 (7689):531-534.
2019	专刊《从基因型到表型》阐述了探究基因型和表型之间复杂关系的研究进展及方法，包括不同的环境因素如何影响表型变化等，促进人类健康相关的更多发现	Science, Vol 365 (6460):1394-1395
2020	二十周年特刊提出，解析复杂表型包括疾病相关的“家族相似性”是遗传学领域的最大难题之一，比如鉴定决定表型的关键信号通路上的基因等，为理解人类多基因疾病机制奠定基础	Nature Review Genetics, Vol 21 (10):581-596
2021	在聚焦未来十年精准医学的七个关键性发展领域中，包含“表型组学与环境”和“队列研究”，涉及整合多队列数据，并通过不同方法测量临床表型、环境暴露和生活习惯等	Cell, Vol 184 (6):1415-1419

近年来，深度表型测量创新加速，近五年来科技文献增长显著，多学科领军学者提出表型组学及深度表型测量技术对本领域的重要支撑作用。2020 年底英国国家医疗服务体系 NHS 代谢 / 化学病理顾问、英国国家卫生和临床技术优化研究所高血压指南发展组主席、伦敦国王学院教授在心血管领域顶级杂志 Circulation 上撰文提出：表型组学对心血管疾病预测能力更

强，表型组学检查可为心血管疾病患者治疗方案优化提供更精准依据，摆脱目前过度治疗的困境^[2]。2021 年美国牙科研究协会发布会议报告，以“下个世纪的科学：深度表型测量”为题，介绍了口腔医学领域表型组学在疾病早期预防和干预、表型与临床症状跨尺度关联、表型组测量识别和明确口腔疾病发病状况的进展，探讨深度表型测量的重要性^[3]。多尺度表型图谱为大脑细胞网络阐释^[4]、多器官蛋白质组定量^[5]、疾病免疫细胞表型分析^[6]等领域提供了理论支撑。

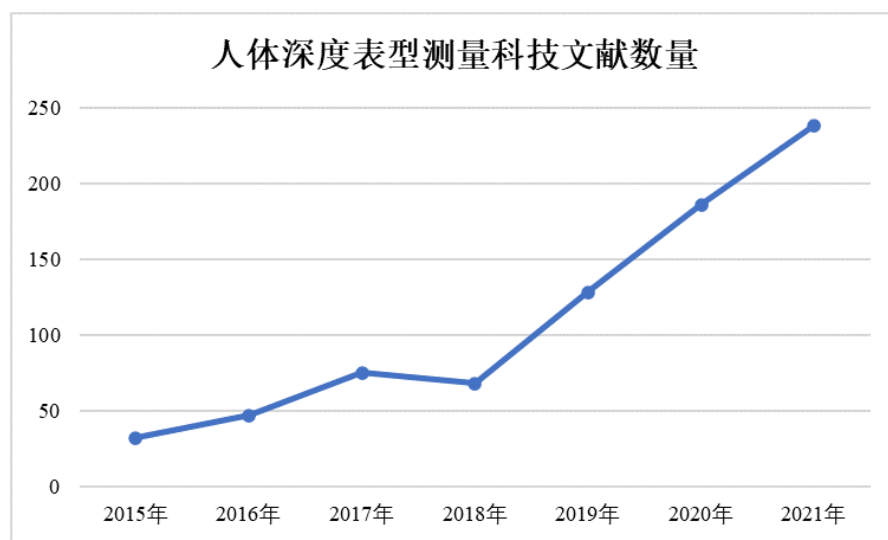


图2 人体深度表型测量（Deep phenotyping）科技文献数量在“国际人类表型组”大科学计划发起后增速显著

2021 年，人类表型组学研究（Human Phenomics）入选由中国工程院、科睿唯安公司与高等教育出版社联合发布的《全球工程前沿》，被评选为医药卫生领域 TOP10^[7]，作为具有前瞻性、先导性和探索性的前沿领域得到领域专家高度认可，是对科技未来发展有重大影响和引领作用的关键方向。

（二）国际人类表型组计划

1、发展历程

以复旦大学为代表的中国科学界自 2014 年起就开始在“测一切之可测”理念指导下筹备发起人类表型组计划。2015 年 5 月，金力院士倡议发起并组织召开了“国际人类表型组研究”香山科学会议，会上国际“代谢组学之父”英国 Jeremy Nicholson 院士、国际人类基因变异组学会主席 Richard Cotton 教授和金力院士、赵国屏院士、王辰院士等多位专家学者一致提议发起国际人类表型组计划。2016 年 5 月，复旦大学在上海组织召开了“首

届国际人类表型组大会”，金力院士在大会主旨报告中首次提出了人类表型组计划的核心任务与路线图。本次大会上，由金力院士，美国四院院士、拉斯克奖获得者、“系统生物学之父”莱诺伊·胡德（Leroy Hood）院士，英国皇家医学科学院院士、国际“代谢组学之父”杰瑞米·尼克尔森（Jeremy Nicholson）院士三位科学家共同发出的“国际人类表型组研究计划”倡议获得了国内外科学家的一致认可和积极响应，奠定了我国在组织国际人类表型组计划中的领导力与话语权。

2016年，“国际人类表型组”被国务院批准列入需要布局的重大科学基础工程。2017年11月，上海市首批市级科技重大专项对“国际人类表型组计划（一期）”予以立项支持，为这一国际大科学计划的加速启动与推进注入了关键动力。在前期中外科学界已经具备的合作基础上，复旦大学在上海市重大专项支持下，积极规划组织模式顶层设计与推进路径，加快推动构建大科学计划全球协同的网络布局。

表 2、国际人类表型组计划大事记

2015年5月	“国际人类表型组研究”香山科学会议召开，提议发起国际人类表型组计划。
2016年4月	国务院印发《上海系统推进全面创新改革试验加快建设具有全球影响力的科技创新中心方案》，将“国际人类表型组”作为重大任务，列入需要布局的重大科学基础工程。
2016年5月	复旦大学在上海组织召开了“首届国际人类表型组大会”，金力院士在大会主旨报告中首次系统的提出了人类表型组计划的核心任务与路线图。
2018年3月	上海市首批市级科技重大专项“国际人类表型组计划（一期）”正式启动。
2018年10月	“国际人类表型组研究协作组（IHPC）”在“第二届国际人类表型组研讨会”举办期间正式成立“人类表型组大科学计划”正式由科学界先行启动。
2019年9月	经上海市批准，新型社会组织和研发机构上海国际人类表型组研究院成立。
2020年10月	“第三届国际人类表型组研讨会”及“第二次国际人类表型组协作组理事会”召开，确定了人类表型组计划的三大优先推进方向。
2021年11月	“第三次国际人类表型组协作组理事会”召开，发布了全球首张人类表型组参比导航图，“人类表型组计划科研数据跨境共享与开放原则”达成共识。

人类表型组国际大科学计划绘制全球人类表型组参比图谱的构想目标，需要全球数个甚至数十个团队通过“分布式”的协同测量与数据协同分析实现，对国际科技合作的组织协调提出了极大要求。IHPC的成立标志着由中国科学家倡议发起的国际人类表型组计划已得到多国科学家的认同和参与。截至2021年，IHPC汇聚了美、英、德、澳大利亚、哈萨克、加纳等20个

国家 23 个高校科研机构，理事包括 10 位各国院士及 13 位领军科学家；HPCC 共有 91 名协作委员，其中 33 位院士，协作单位 / 机构 69 家，其中 38 家高校科研院所，27 家三甲医院，4 家国内知名企业，共同推进人类表型组国际大科学计划。

人类表型组国际大科学计划

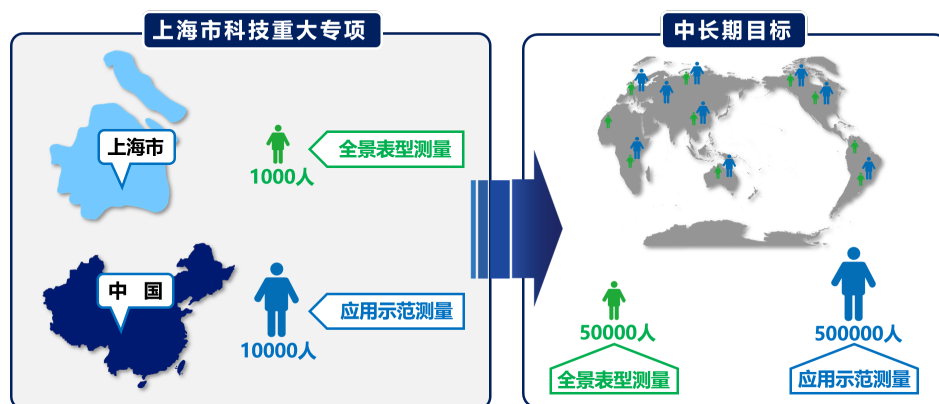
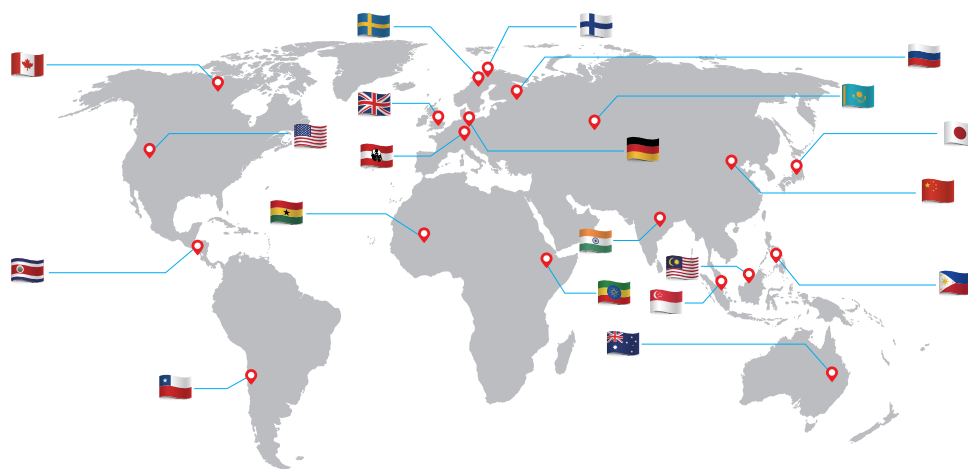


图 3 国际人类表型组计划目标

2、国际人类表型组研究协作组（IHPC）

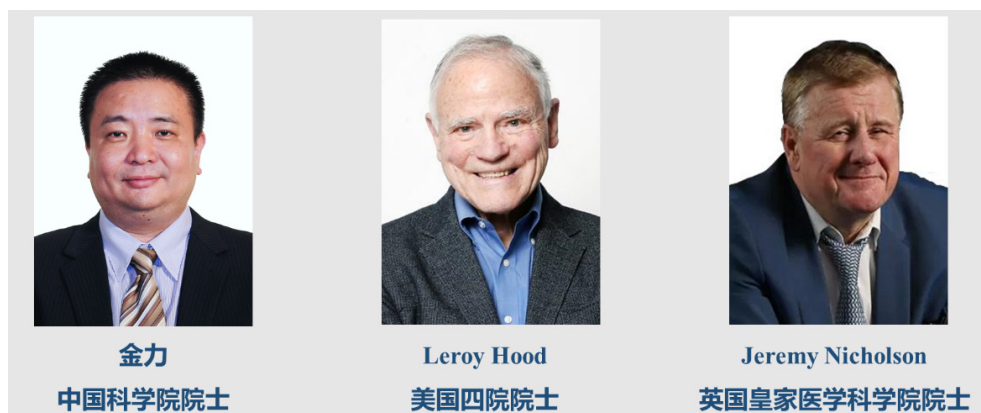


(1) IHPC 简介：

国际人类表型组研究协作组（英文名“International Human Phenome Consortium”，缩写为 IHPC）于 2018 年 10 月在上海成立。中国科学院院士、复旦大学校长金力，美国科学院、医学院、工程院院士、“国际系统生物学

之父”、美国系统生物研究所教授 Leroy E. Hood 以及“国际代谢组之父”、英国皇家医学科学院院士、澳大利亚莫道克大学教授 Jeremy Nicholson 担任 IHPC 理事会的共同主席，来自美、英、德、哈萨克、加纳等 20 个国家的 23 位相关领域顶级专家为理事会成员。理事会下设“标准与技术规范”、“知识产权、数据共享与数据安全”和“伦理与法律社会问题”三个专业委员会，秘书处设在上海国际人类表型组研究院。

共同主席：



秘书处：中国·上海国际人类表型组研究院



秘书长：



田梅
复旦大学特聘教授
复旦大学人类表型组研究院执行院长
上海国际人类表型组研究院副院长
国家杰出青年科学基金获得者
世界分子影像学会主席

目标：

着力构建国际协同创新网，凝聚全球人类表型组研究力量，全面推进人类表型组国际大科学计划。

使命：

- ※ 利用国际和多学科力量，从基因 - 环境 - 表型到人类健康层面，积极探寻微观与宏观表型之间的关联及机制。
- ※ 通过使用物理、化学和生物指标，从分子、细胞到生命体各层面发现人类特征和疾病的起源以及多样性。
- ※ 应用新发现的知识和技术来创造改善人类健康的新范式。

(2) 积极布局国际人类表型组领域治理**2019 年：成立上海国际人类表型组研究院**

2019 年，经上海市政府批准，由复旦大学等单位发起成立的上海国际人类表型组研究院正式建立，金力院士任首任院长。研究院定位为：人类表型组计划的战略智库，表型组标准化工作的创新中心，面向全球科学家的数据共享与信息管理平台，推动国际、国内合作的服务协调机构，生命科学与生物医药领域的高水平研发机构和高质量产业化机构。作为大科学计划的组织枢纽，上海国际人类表型组研究院于 2019 年底起正式承担了国际人类表型组研究协作组（IHPC）和中国人类表型组研究协作组（HPCC）的秘书处职能，确保人类表型组计划组织推进的两大协同机制与核心网络稳定运营。研究院协同国内外科研机构、团队与资源网络，共同实质性推进人类表型组国际大科学计划，推动生命科学与生物医药领域的原创性、交叉性和颠覆性研究；面向健康医疗设备、新型诊断试剂、原创新药、健康管理产品等相关领域，全面开展人类表型组相关产学研合作，加速精准医疗与精准健康管理等产业发展。

2020 年：加强国际协同机制组织化建设，明确人类表型组大科学计划优先推进的三大方向

国际人类表型组研究协作组致力于构建全球协同创新网络，加强大科学计划国际协同机制组织化建设。IHPC 第二次理事会正式确定将国际协作组秘书处设于上海国际人类表型组研究院，并设置“标准与技术规范”、“伦理、法律与社会规范”、“数据共享”三个专业分委会。

在 IHPC 秘书处的组织推动下，第二次理事会就进一步推进人类表型组国际大科学计划的优先发展方向初步达成共识，认为近期应优先聚焦“新冠

肺炎和其他重大疾病的表型组学研究”、“表型组研究技术体系与科研基础设施构建”以及“表型组学研究中的标准操作程序”三大方向。

2021 年：人类表型组科研数据跨境共享原则

在 IHPC 第三次理事会上，经过各国科学家的充分探讨与协商，就人类表型组大科学计划下科研数据跨境共享与开放的基本原则，在兼顾安全和开放的前提下达成了六点共识，包括采纳“FAIR (Findable, Accessible, Interoperable, and Reusable)”原则，即“寻获性、可及性、互用性和复用性”；安全合规；数据标准化原则，IHPC 将致力于研究和建立国际人类表型组科研数据标准化体系；不同类型、级别的数据适用不同的分享与开放；多模式分场景共享；鼓励各国科研团队通过科研实践探索国际科研合作与数据跨境共享的最佳实践。

(3) 引领国际人类表型组研究领域发展

IHPC 理事在各国积极引领人类表型组研究，中国、美国、澳大利亚、印度等国纷纷发起人类表型组相关科研计划并建设研究平台。在国际人类表型组计划引领下，分子表型组学、早期癌症筛查、多组学生物标志物筛选等多方面已取得突破性成果，促进了 IHPC 框架下国际科技合作。

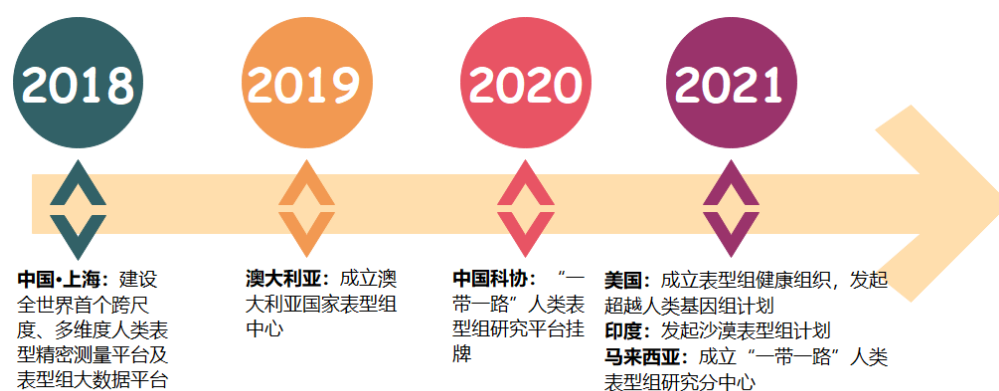


图 4 国际人类表型组计划引领下多国发起人类表型组相关科研计划并建设研究平台

3、中国国际人类表型组研究协作组（HPCC）

中国人类表型组研究协作组（HPCC）由复旦大学金力院士、中国医学科学院王辰院士和中国科学院大学徐涛院士担任共同组长。HPCC 旨在整合国内高校科研院所、医疗机构和产业界等在人类表型组学领域的优势，发展表型测量技术、制定相关标准和规范、实现中国团队在人类表型组研究领域的重大突破、形成主导全球协作的中国合力，共同推动人类表型组国际大科

学计划的实施。

HPCC 现有 91 名协作委员，其中 33 位院士；协作单位 / 机构 69 家，其中 38 家高校科研院所，27 家三甲医院，4 家国内知名企业。近年来，在国内各协作单位的共同努力下，尤其是 HPCC 各位委员的大力支持下，HPCC 整合国内人类表型组研究力量、推进我国人类表型组领域原始创新和协同攻关并取得了显著进展。

三、人类表型组：全球布局与发展态势

2020 年至 2021 年，多国竞相布局人类表型组领域，人类表型组研究出现在多国政策性文件和智库报告中，成为生命科学热点领域。多个人类表型组相关科技组织成立，多国启动了大型表型组学研究计划。

（一）中国：整合并推进国际人类表型组大科学计划

1、推进区域表型组研究机构落地与发展

2021 年，重庆两江新区与上海国际人类表型组研究院、重庆医科大学签订战略合作框架协议，将共同推动人类表型组学前沿研究在中国西部区域布局落地与纵深发展，助力重庆和两江新区打造中国西部人类表型组研究与转化重镇。

重庆两江新区、重庆医科大学与上海国际人类表型组研究院三方将充分发挥两江新区在政策落地、产业集聚、投资环境、人才吸引等方面的资源优势，结合重庆医科大学、上海国际人类表型组研究院的人才、智力及学术优势，拓展科研资源，进一步提升区域创新能力，加速科技成果转化。三方将共同建设对标上海的人类表型组精密测量与研究センター，从而进一步优化人类表型组大科学计划的协同网络和产学研联盟。

2、打造表型组学学术期刊 Phenomics



图 5 2021 年度 Phenomics 期刊封面

2021 年 1 月，上海国际人类表型组研究院与施普林格·自然出版集团联

合创办 Phenomics (《表型组学》) 期刊, 聚焦表型组学这一新兴交叉学科研究, 旨在搭建该领域的全球学术交流载体, 推动表型组学理论创新和学科发展; 致力于发展成为服务于表型组学这一特定的生命科学研究领域的国际高水平期刊。

Phenomics 期刊创刊后发展迅速, 经国际学界协商确定为 IHPC 会刊、人类表型组国际大科学计划的旗舰期刊。期刊聚焦表型各方面研究, 包括分子水平的蛋白质组和代谢组研究, 细胞水平的细胞特征及器官水平的各种器官研究, 基因组结构和调控网络机制, 表型关联与疾病风险和干预措施等, 开设论著、综述、实验方法等文章类型, 吸引了大量国内外表型组学领域的优秀科学家投稿。2021 年在线出版了 6 期, 共计 24 篇文章, 作者分布于中、美、德、英、澳等 20 多个国家, 累计下载量超过 5 万次。

3、构建人类表型组标准体系

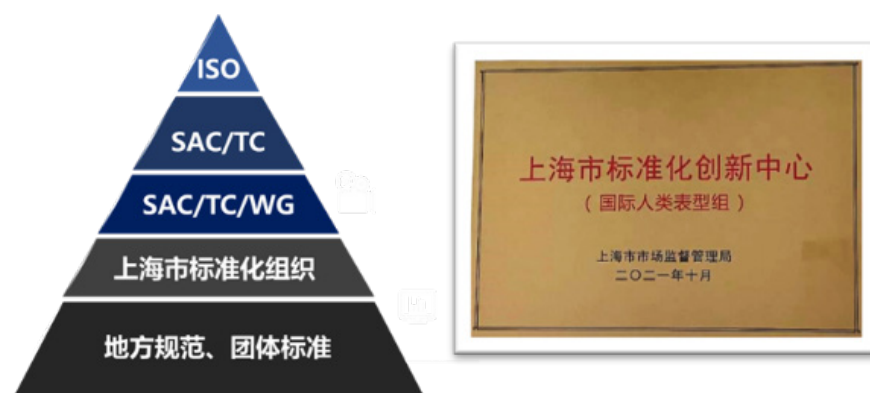


图 6 上海市标准化创新中心 (国际人类表型组) 于 2021 年 10 月获批

上海国际人类表型组研究院持续推动建设引领全球人类表型组标准化研究的创新枢纽。2022 年 10 月, 研究院获批 “上海市标准化创新中心 (国际人类表型组)”, 成为上海市首批 6 家新型标准化技术组织单位之一。

上海市标准化创新中心 (国际人类表型组) 聚焦探索国际大科学计划的前进方向, 研制人类表型组标准, 推进 “标准领跑”、“质量领跑”, 发展上海地区在人类表型组影像表型、皮肤表型的团体标准, 推动技术标准走向生物医药产业应用的发展方向, 促进上海生物医药企业主动实施高标准、追求高质量, 倡导行业标准和质量领先, 持续提升上海市生物医药企业的市场竞争力。上海市标准化创新中心 (国际人类表型组) 也将推动 “上海标准” 创新性发展, 充分发挥上海市遗传学会标准化技术委员会、上海市生物信息

学会标准化技术委员会、上海人类学学会标准化技术委员会等 3 个学会标准化技术委员会的平台优势，加大标准宣贯和培训力度，提升上海生物医药企业对团体标准的认知普及度。

4、中国科协“一带一路”人类表型组联合研究中心

2020 年，中国科协国际科技组织合作平台——“一带一路”人类表型组联合研究中心获批建设。该中心由复旦大学牵头发起，得到来自新加坡、菲律宾、马来西亚、蒙古国、俄罗斯等 9 个国家 11 家高校与科研机构的积极响应与参与。当年 12 月 13 日，联合研究中心在上海正式揭牌。作为人类表型组计划聚焦“一带一路”地区优先布局落地的重要战略之一，“一带一路”人类表型组联合研究中心将努力成为“一带一路”区域生命科学前沿国际协同与交流的重要枢纽，引领构建“一带一路”区域科技创新联合体，为打造健康“一带一路”，守护各国人民生命健康做出应有贡献。“一带一路”人类表型组联合研究中心的主体部分设立于张江复旦国际创新中心内。2021 年马来西亚“一带一路”人类表型组联合研究中心分中心在 UCSI 大学设立并揭牌。

（二）美国：人类表型组是医疗健康领域未来发展机遇

1、表型组学及多组学研究被纳入美国发展战略规划

2020 年至 2021 年期间，美国多份政策性文件和有极高影响力的智库报告都将表型组学和多组学作为生物医学重点发展的关键领域。

2020 年 1 月，美国国家科学、工程院与医学院发布了《护航生物经济》（Safeguarding the Bioeconomy）。这份美国最高水平学术机构发布的政府决策参考的报告中多次提及“表型”和“多组学”，提出更大规模人群发现基因组 - 表型组之间重要关系将促进医学进步，并指出投资临床表型数据基础设施建设和数据相关领域研发的意义。

2021 年 3 月，顶级科技期刊 Cell 刊登美国精准医学 2030 学界倡议，美国国立卫生研究院（NIH）主席 Francis Collins 和 All of US 研究项目的首席执行官 Joshua Denny 在文中提出，未来十年精准医学将持续改变医疗健康。在七大加速实现精准医疗的关键领域中，“电子健康档案（EHR）作为表型和基因研究的来源”、“表型组学与环境”和“纵向大队列共享中，相关表型判断并建立联盟”位列其中三项。

2021 年，美国国会在《无尽前沿法案》、《2021 年战略竞争法案》、《2021 年迎接中国挑战法案》、《2021 年美国创新与竞争法案》等多部方

案审议稿中均将多组学研究列入生物医学领域重点关注和发展的关键技术，高度重视多组学研究对保持美国在科技领域优势的重大意义。其中《2021 年生物经济研究与发展法案》定义了多组学（Omics），并提及国家工程生物研究与发展计划需持续支持数据库与相关工具的开发，包括支持筹备相关组学数据库进行工程生物学研究并提高该类数据库效用；法案规定国家海洋及大气管理局必须建立多组学研究和相关生物信息科学的支持项目，以推进可持续生物经济和发展下一代提高生态系统管理、监督、经营、评估和预测的工具和产品。

2021 年 5 月，美国大西洋理事会下的“新兴技术与数据的地缘政治影响委员会”发布《新兴技术与数据的地缘政治影响》为美国未来十年数字与科技重点战略发展方向提供导向性政策建议。该报告中提到基因组技术以及多组学技术将影响全球卫生与健康领域，列为需要发展的关键技术，体现了其战略地位。

2、成立表型组健康组织并发起超越人类基因组计划

2021 年 10 月，国际人类表型组计划共同发起人、IHPC 共同主席、美国四院院士 Leroy Hood 宣布表型组健康（Phenome Health）组织成立。表型组健康组织是一个非营利组织，旨在推动科学和数据驱动的方法来优化个体的大脑和身体健康。表型组健康组织于同年发起了超越人类基因组计划（Beyond The Human Genome, BHG），希望通过在十年内分析美国多族群 100 万人的基因组和表型，了解并优化个人健康轨迹，有望在医疗保健领域实现从以疾病为中心转向以预防和健康为中心的愿景。

（三）欧盟：完善基因组 - 表型组档案建设

欧盟是最早启动人类表型组相关知识库建设的地区。1996 年启动了专注于罕见疾病及相关药物等信息的 Orphanet 知识库建设，协调并制定罕见疾病表型标准术语。2008 年德国启动 Human Phenotype Ontology (HPO) 项目，采用本体工程学、计算机科学对来自医学文献的表型信息进行结构化归纳，对表型相关词汇建立分层关系，并对描述表型的词汇及其语义相互关系进行准确定义。

2013 年起，欧洲分子生物实验室（European Molecular Biology Laboratory, EMBL）下属生物信息研究所和基因组调控中心（Centre for Genomic Regulation, CRG）共同发起并管理欧洲人类基因组 - 表型组档案（European Genome-phenome Archive, EGA）计划。该计划在严格协议下，

为生物医学研究项目产生的个人可识别基因、多组学和表型数据提供永久的存档和分发服务。

2020 年 EMBL、德国癌症研究中心 (German Cancer Research Center, DKFZ) 等多家机构成立联盟, 将德国人类基因组 - 表型组档案 (German Human Genome-Phenome Archive, GHGA) 作为德国国家研究数据基础设施 (National Research Data Infrastructure Germany, NFDI) 提案的一部分正式启动。GHGA 作为欧洲人类基因 - 表型档案 (EGA) 在德国的分支, 旨在提供基础设施, 以公开和公正原则处理多组学数据, 并保证个人数据安全。不同于其他欧洲的基础设施, GHGA 可满足德国特有的法律, 使德国的研究人员能够参与制定未来国际数据交流标准, 在国际研究组织中发挥主导作用。

(四) 英国：将表型组学列入生命科学发展规划

1、表型组学被列入生命科学发展战略十年规划

2020 年 7 月英国政府颁布了建设科技强国的战略政策文件《英国研究与发展路线图》(UK Research and Development Roadmap)。路线图中提出了七大战略重点任务, 其中包括建立世界顶尖的基础设施和机构。英国政府将投资现有的基础设施建设及英国研究与创新署 (UK Research and Innovation, UKRI) 2019 年底发布的《英国研究与创新基础设施》(The UK's research and innovation infrastructure) 中提出的六大需投资的板块, 其中生物科学、健康和食物板块将生物影像、结构生物学和多组学技术, 以及大规模深度人类表型组测量作为重点建设领域。

2020 年 9 月, 英国政府颁布了政策文件《基因组英国》(Genome UK), 以创造世界最先进的基因组医疗卫生生态系统为愿景, 强调基因组学基础设施结合深度表型测量和临床表型数据的重要意义。英国现有的主要基因组学基础设施如 Genomics England、NHS 的基因组检测服务、英国生物样本库 (UK Biobank)、英国国家健康研究所生物资源库 (BioResource) 在基因组数据基础上, 都会结合深度表型测量和临床表型数据收集来进行个性化医疗和研究, 以改善患者结果。未来十年, 英国政府会支持对于含有独特信息的患者队列的发现、招募、表型测量和采样。

2021 年 7 月, 英国政府商务能源与产业战略部和下属的生命科学办公室发布了政策文件《生命科学愿景》(Life Sciences Vision), 宣布英国关于生命科学领域未来十年的发展规划。为实现领先战略目标, 将对临床研究、

基因组和健康数据的持续布局，其中，基因组学方面将继续支持和加强基因组研究基础设施建设，将世界级纵向队列中的“组学”资产整合起来，例如 UK Biobank、NIHR BioResource 和 Genomics England 等，并持续研发改善诊断、患者分级、临床试验和个性化治疗的多组学新兴技术。

2、启动全球最大的多模态成像研究计划

在投资既有优势的分子表型组学研究的同时，英国还启动了影像表型组学研究计划。英国 UK Biobank 启动了全球最大的多模态成像研究，计划收集 100,000 名参与者的脑部、心脏和腹部扫描影像表型，同时推动成像和分析的科学进步，开辟诊断、治疗和潜在预防痴呆症、心脏病和癌症等疾病的新方法。该计划将在健康与疾病期间捕获人体的大量图像。2021 年 UK Biobank 宣布已搜集到 50,000 名参与者的成像扫描，是目前全球最大规模的影像学研究，将助力理解生活方式、遗传因素、成像衍生表型以及这些因素与疾病风险之间的关联。

（五）澳大利亚：建设全球领先的分子表型组平台

1、建成澳大利亚国家表型组中心

澳大利亚国家表型组中心（ANPC）于 2019 年正式成立，国际人类表型组计划共同发起人、IHPC 共同主席、英国皇家医学科学院杰里米·尼克森（Jeremy Nicholson）院士作为执行主席，研究基因、环境和生活方式对人类和动物健康的复杂相互作用和影响，是西澳大利亚最主要的健康研究合作项目之一，也是西澳大利亚健康转化网络（Western Australian Health Translation Network，简称 WAHTN）的核心平台。ANPC 目前是全球最全面的代谢组学平台之一，建有全球最大的核磁共振及质谱分析平台，通过核磁共振光谱和前沿数据建模带动了西澳大利亚代谢组学发展，在代谢组学分析和建模方法领域具备引领优势，是代谢表型研究的全球表型组中心网络（International Phenome Centre Network, IPCN）的协调中心。

2、启动多项分子表型组学研究计划

2021 年 ANPC 在西澳大利亚州政府支持下，启动了多项分子表型组学研究计划，关注多项健康挑战，如影响全球并带来健康危机的 COVID-19 分子表型组分析国际合作计划、老年人健康保障计划、疾病预防生物标记物识别筛选计划、儿童代谢与免疫疾病分子表型研究计划等。ANPC 实验室根据从世界各地收集的样本，在了解 COVID-19 感染者代谢表型变化方面取得了巨大进展。

（六）印度：发起沙漠表型组计划

IHPC 理事、曾任德里基因组学和整合生物研究所（CSIR-IGIB）的首席科学家的 Mitali Mukerji 教授，2021 年加入印度理工学院（IIT）焦特布尔分校，协调启动了沙漠表型组计划 Thar DESIGNS 计划（Desert EcoSystem Innovations Guided by Nature and Selection，自然和选择指导下的沙漠生态系统创新）。该计划旨在通过医学、工程、环境和生命科学的跨学科框架开展生态系统表型组研究，通过物联网设备、数据分析框架、机器学习和领域知识，将在地理时空尺度上推断环境、表型和基因型之间的联系。

四、人类表型组：年度研究进展

根据 Web of Science 核心合集的 SCI 期刊论文引用数据和生命健康领域热点，我们总结并摘选了人类表型组研究领域 2021 年度研究进展及典型案例。

（一）“国际人类表型组计划（一期）”项目进展

1、概况

基于 Web of Science 核心集 SCI 期刊论文和会议论文数据，在上海市首批市级科技重大专项对“国际人类表型组计划（一期）”支持下，多领域表型组学相关研究取得进展。截至 2021 年底，中国、美国、法国、英国、荷兰等 23 个国家的 194 家研究机构共发表科技论文 791 篇，呈增长趋势。发表的科技论文除临床与生命科学外，还涵盖了化学、农业环境与生态、电气工程和计算机科学、物理、社会科学、工程与材料科学、地球科学、数学以及人文等多学科，展示了表型组学研究领域学科融合和综合发展的趋势。参与科研攻关的包括大学、科研机构、医院以及健康产业公司，促进了基础科研与医疗服务、健康产业的结合与转化应用。

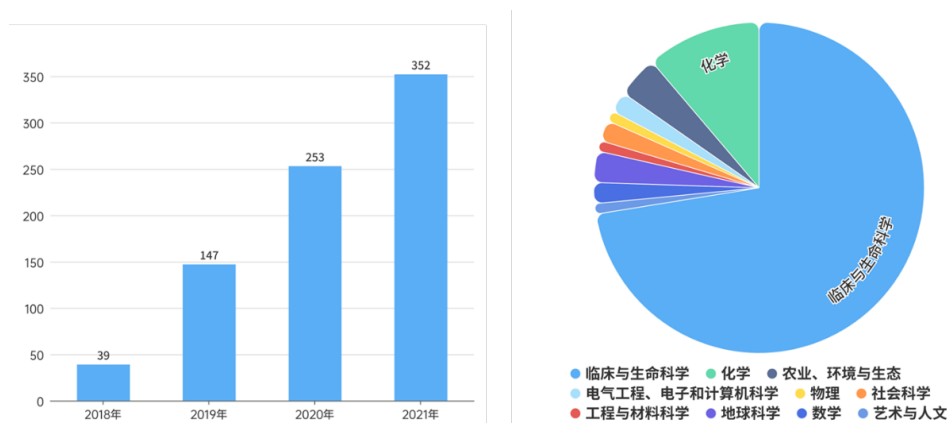


图7 左：上海市级科技重大专项“国际人类表型组计划（一期）”历年发表论文情况；右：发表论文各学科占比。

2021 年度一期发表论文前十位的主题分别是：分子与细胞生物学、基因组研究、神经退行性疾病、肺癌、免疫学、炎症性肠病和感染、生殖生物

学、质谱法、肝炎、心律失常、皮肤病学、精神病学、神经科学，综合占比接近当年一期发表论文总数的一半，说明基因 - 表型关联及相关机制、疾病的防诊治和表型测量技术是当前国际人类表型组计划重点关注和着力研究的主题。

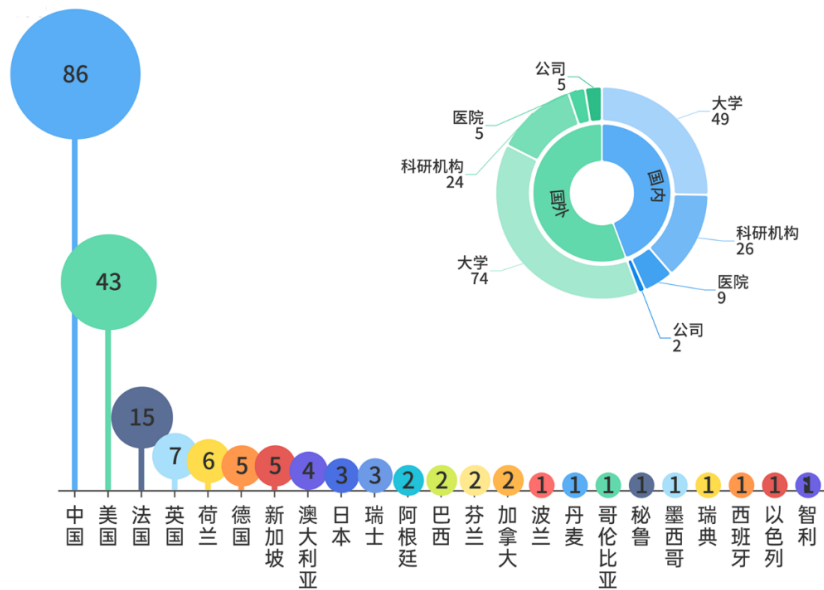


图 8 左：“国际人类表型组计划（一期）”发表论文署名单位国别情况；
右上：论文署名国内外机构占比情况

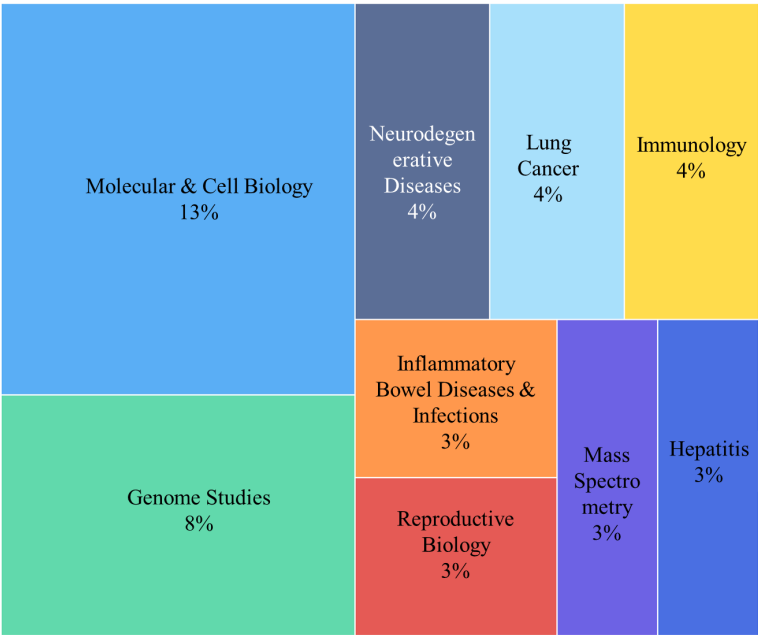


图 9 2021 年上海市级科技重大专项“国际人类表型组计划（一期）”发表论文占比前十位领域及分布情况。

2、首张人类表型组全景导航图

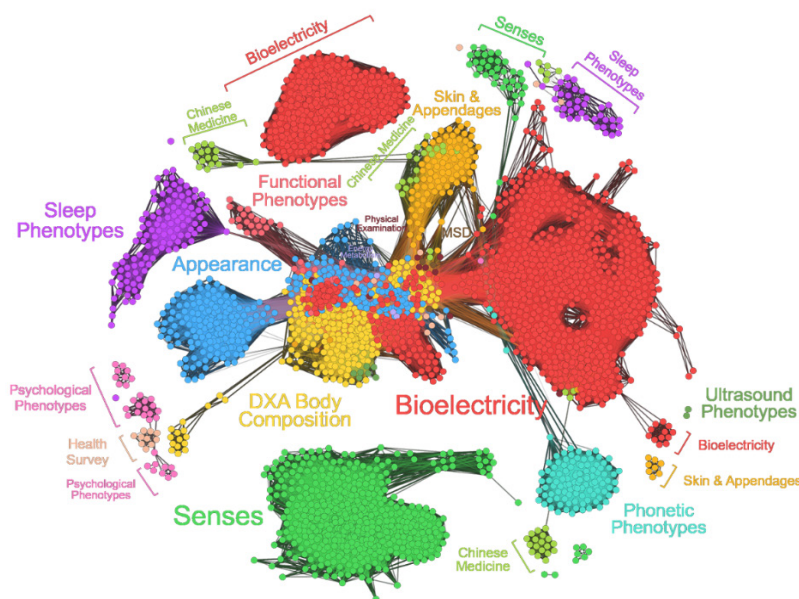


图 10 首张人类表型组全景导航图（示意图）

2021 年 11 月，在国际人类表型组研究协作组 IHPC 第三次理事会上，全球首张人类表型组导航图宣布初步绘制完成。依托在张江复旦国际创新中心建成的世界首个跨尺度、多维度人类表型精密测量平台和自主研发的全过程、自动化表型组大数据平台，中国科学家从 2020 年起开展了“上海自然人群健康表型核心队列研究”。1000 位常住上海的 20-60 岁志愿者完成了在张江平台 2 天 1 夜、每人超 2 万个指标的全景表型测量，使科学家首次获得了自然人群样本贯通宏观至微观尺度 20 余个领域类别的海量表型基线大数据，数据总量超过了 3PB。经过多学科团队对现有数据的协同分析攻关，科学家发现了不同表型间超过 150 万个强关联，其中约 39% 为跨尺度关联，且大部分关联是科学界首次看到。通过这些强关联形成的网络，中国科学家初步绘制了基于上海核心队列的全球首张人类表型组参比导航图^[8]。目前，面向科研用户的导航图数据库网站已经上线公测，多支科学团队正在从现有 1.0 版的导航图中，筛选具有重大科学意义和应用价值的强关联，开展进一步科研攻关。

这张“导航图”将为科学家未来进一步解析复杂生命系统的机理与奥秘指明新的方向、提供有价值的线索，是具有战略性意义的生命科学原始创新源，有望引领新一轮健康科技与生物产业变革。

3、质量控制与标准化

长期以来，由于缺乏合适的参考物质，无法产生标准化的 DNA 测序数据集并对不同测序流程或不同算法的性能进行全面测试和评估，制约了肿瘤基因组学方法的落地实施和广泛使用。2021 年 Nature Biotechnology 专题发布了国际组学数据质量控制联盟（MAQC）第四期（亦为测序质量控制第二期 SEQC2）的最新研究成果。该项目历时 5 年多，全球 100 多个单位的 300 多位科学家参与，通过建立和使用一系列基因组参考物质，MAQC 对众多新一代测序平台和生物信息学方法在体细胞和胚系突变、拷贝数变异、肿瘤靶向测序和液态活检、单细胞测序等多个应用场景的性能进行了系统研究^[9-13]。Nature Biotechnology 同期配发了主编 Andrew Marshall 撰写的社论，认为 SEQC2 对新一代测序平台和分析方法进行了迄今为止最为全面客观的评估，为基础研究、临床和监管部门阐明了相关的技术规范和质量控制指标^[14]。此外，同期斯坦福大学教授 Marc Salit 和 FDA 代理主任 Janet Woodcock 博士评论“MAQC 高质量工作为监管和临床提供了基础，推动了整个领域的发展，提升了人们对基因组数据可靠性的信心，直接促进了监管科学的发展，证明监管科学可以促进新技术的快速推广和临床使用，正支撑着精准医学领域的许多药物和诊断设备的上市申报并希望 MAQC 的成功转化将在更多领域得到复制^[15]。”

复旦大学人类表型组研究院石乐明教授是 MAQC 项目核心推动者之一，作为共同通讯作者领导其中两篇论文的研究与撰写。石乐明教授及团队长期从事多组学质量控制研究，研发的多组学标准物质被包括欧洲转化医学基础设施（EATRIS）在内的全球 100 多家科研机构使用。

2021 年由石乐明教授牵头起草的国际标准 ISO/TS 22690:2021《基因组信息学 高通量基因表达数据可靠性评估》（Genomics informatics — Reliability assessment criteria for high throughput gene-expression data）于 2021 年 10 月发布，该标准规定了高通量基因表达数据的可靠性评估标准，适用于从基因芯片、新一代测序的基因表达数据的准确性、复现性、可比性的评估应用。

（二）新冠肺炎的表型组学研究

2021 年，新冠肺炎大流行造成亟待解决的全球人类健康问题，了解新冠肺炎的传播途径与生物学后果，成为生命健康领域研究热点。在 2020 年 11 月召开的第三届国际人类表型组研讨会上，各国科学家高度认同，应把

聚焦新冠肺炎疫情开展表型组学研究作为人类表型组大科学计划启动后实质推进的首要优先方向之一。人类表型组研究对表型数据动态跨尺度的采集，揭示了新冠感染后表型转化，助力新冠肺炎诊断、病情发展预测和新型技术的研发。

1、明确新冠感染后表型转化特征

国际人类表型组计划共同发起人、澳大利亚国家表型组中心执行主席 Jeremy Nicholson 院士团队利用新型核磁共振波谱和质谱技术等代谢表型分析方法并综合数据分析 COVID-19 感染后生物标记物和代谢途径，明确脂蛋白、糖蛋白、脂质和其他关键代谢产物与炎性细胞因子之间相关性，建立个体从健康状态转变为新冠感染的紊乱病理生理状态表型转化模型，揭示急性后新冠肺炎综合征（post-acute COVID-19 Syndrome, PACS）相关持续症状和多器官损害^[16]，并提出患者恢复过程高度多变，在生化特征恢复正常之前，不能认为患者“完全恢复”^[17,18]。

慕尼黑工业大学的 Andreas Pichlmair 教授联合蛋白质组学资深教授 Matthias Mann 等专家，在 Nature 上发布了冠状病毒感染宿主细胞的分子机制研究。研究团队通过对肺源性人类细胞系感染 SARS-CoV-2 和 SARS-CoV 前后进行转录组、蛋白质组、泛素化修饰组和磷酸化修饰组学研究，分析了这两种冠状病毒蛋白质与宿主蛋白质的相互作用，阐明冠状病毒独特以及共有的关键蛋白质^[19]。

2、评估新冠患者功能恢复并预测疾病发展

国际人类表型组计划共同发起人、美国系统生物学研究所 Leroy Hood 院士团队参与的解析轻度和中度 COVID-19 之间免疫表型综合图谱项目，揭示轻度、中度感染之间免疫表型区别。该研究团队对新冠患者血浆代谢物和蛋白质水平综合分析并提出血浆代谢物可区别新冠临床严重程度，为新冠患者疾病发展预测提供了免疫代谢机制数据及分析方法^[20]。来自美国拉根研究所、麻省理工学院、哈佛医学院、系统生物学研究所和麻省总医院的研究人员在国际顶级学术期刊 Cell 发布多组学 COVID-19 研究论文，对 139 名不同疾病严重程度的 COVID-19 患者的临床测量结果、免疫细胞和血浆多组学数据进行综合分析，鉴定出轻度 COVID-19 和中度 COVID-19 之间的转变点，在此过程中炎症信号的升高伴随着特定类代谢物和代谢过程的丧失。在中度 COVID-19 患者的这种应激的血浆环境中，出现了多种特征性免疫细胞表型，并随着疾病严重程度而增强。研究人员将超过 120,000 个免疫特征浓

缩为免疫反应轴，与主要的血浆组成变化，血液凝固的临床指标以及轻度和中度疾病之间的急剧过渡保持一致^[21]。为 COVID-19 患者干预治疗区间提出了建议。

粤港澳大湾区精准医学研究院沈侠团队和哈佛医学院 Vadim Gladyshev 等团队合作报道加速衰老的因素导致 COVID-19 风险增加，Notch 信号传导通路和 B 细胞免疫表型可能起了重要作用，针对这些因素降低生物年龄的干预措施可能会降低 COVID-19 的风险^[22]。中山大学第五医院严志祥团队揭示了新冠患者感染后肠道分子表型组学变化，并提出受干扰的肠道分子谱可能影响新冠疾病发展^[23]。

3、研发新冠肺炎诊断与治疗表型组学新技术

瑞典乌普萨拉大学 Carreras-Puigver 团队将细胞形态绘制和基于抗体的病毒感染检测相结合，设计出病毒感染细胞的分类图像分析方法，可通过细胞表型特征进行针对病毒感染的非靶向药物筛选^[24]。东南大学附属中大医院放射科主任、东南大学医学院副院长居胜红教授团队在《表型组学》(Phenomics) 期刊上发表封面文章，利用深睿医疗 Dr. Wise 肺部疾病智能解决方案（新冠肺炎增强版）进行肺部病灶的检测、分割，并由专家进行矫正提取，得到新冠肺炎肺部病灶的影像组学定量化特征，该研究结果对新冠肺炎分型进行补充，有助于严重病例的识别与程度评估^[25]。

（三）深入理解基因 - 表型间关联

基因组与表型组之间的关联是生命健康领域公认的科学难题，也是当今科研最热门的研究领域之一。探索表型的产生机制，将推动人群遗传关系、健康管理、精准医学等多领域发展。2021 年多篇表型组学科技论文为深入理解基因 - 表型关联打下基础。

1、解析亚洲热带原住民遗传起源和表型趋同演化机制

复旦大学徐书华教授团队和“一带一路”人类表型组联合研究中心马来西亚分中心、马来西亚 USCI 大学 Boon-Peng Hoh 教授团队合作成果 2021 年底在 Molecular Biology and Evolution 期刊在线发表。研究团队通过全球热带地区近 1000 例原住民的基因型数据，系统研究了人群的遗传关系、祖源构成和迁徙历史，并以深肤色为例探索人群适应性进化的遗传机制，基于文献和数据库报道的 1057 个色素相关基因和其中包含的 103 个功能效应明确的，对其在全球人群中的遗传变异模式进行分析，归纳出几种可能的演化模型，并发现了基因水平的趋同进化特征，为理解现代人类表型多样性

产生和演化机制研究提供了参考模型和典型案例^[26]。

2、解析皮肤菌群表型助力皮肤疾病治疗

上海复旦大学和深圳华大生命科学研究院的研究人员构建了相对完整的人类皮肤微生物基因组 (iHSMGC)，分析了皮肤菌群的组成和功能特点、抗生素抗性基因的来源和分布以及与国人皮肤特性相关的痤疮丙酸杆菌 (C) 和奥斯陆分枝杆菌 (M) 为优势菌的微生物表型^[27]。该成果发布于 Microbiome 期刊上，将有助于更好地了解皮肤微生物组稳态的形成和皮肤病的治疗。

3、X 连锁遗传性视网膜劈裂症深度表型测量

X 连锁遗传性视网膜劈裂症 (X-Linked retinoschisis, XLRS) 是由 RS1 基因的致病性变体引起的遗传性视网膜疾病 (IRD)，约占所有儿童期发病的 IRD 的 5%。英国的研究团队对分子确诊的 132 名经儿童和成人组成的 XLRS 队列进行深度表型测量研究，描述他们的遗传和临床特征，研究基因型 - 表型相关性，并建立了最佳矫正视力与年龄、光学相干层析 (OCT) 特征、自体荧光 (FAF) 特征和并发症之间的纵向临床相关性。研究结果表明，XLRS 具有高度的表型差异，但因中心凹和最佳矫正视力可保留到成年晚期，所以能进行更准确的预测^[28]。

(四) 表型跨尺度关联发现疾病风险

通过对人类表型组系统精密的跨尺度测量，进而积累起海量的生物信息大数据，再对这些数据进行分析建模，能够发现基因 - 表型 - 环境以及宏观 - 微观表型之间的关联机制，将助力理解健康 - 疾病间转化过程，支撑精准医学发展，最终达到对疾病精准地诊断、干预、调控和治疗。2020-2021 年期间，多种疾病的多维度、跨尺度表型关联被发现，对人类健康和疾病精准防治有重要意义。

1、阐明阿尔兹海默症表型关联因素

美国西奈山伊坎医学院和英国伦敦国王学院的团队在顶级临床神经学期刊 Annals of Neurology 上报道了多种风险因素与阿尔茨海默 (AD) 表型的因果关系，如较高水平胆固醇和低密度脂蛋白与神经斑块集聚、海马体体积减少有因果关系；舒张压与脉压增加引发血管性脑损伤；吸烟与脑皮层厚度降低相关；教育和长时间睡眠与脑皮层厚度提升有关，可延缓 AD 发作时间^[29]。该研究成果为降低 AD 发作风险，延缓老年痴呆症发作时间提供了理论支撑。

2、母亲压力与儿童期呼吸系统疾病发病风险相关

美国北卡教堂山分校与威斯康星大学的研究团队进行了幼儿呼吸表型形成的研究,将产前和出生(至3岁)母亲的压力分数与孩子从小到大(至10岁)时呼吸表型、呼吸道病毒感染情况等比对,发现母亲的早期压力、抑郁与儿童呼吸系统疾病和中度喘息表型呈正相关,可能影响幼儿气道发育、肺部发育与改变气道张力机制^[30]。该研究扩展了婴儿和儿童时期呼吸道疾病机制范畴,并为母婴期疾病预防提供了新方向。

3、食物组与慢性肾脏疾病表型

瑞典斯德哥尔摩卡罗琳斯卡学院研究团队在 Nature Reviews Nephrology 期刊发表综述,提出食物组(foodome)与代谢类疾病风险相关,长期摄入不健康饮食将提高患病风险,饮食疗法有望靶向慢性肾脏疾病中尿毒症表型如表观遗传改变、肠道菌群失调、线粒体功能障碍、炎症、氧化应激、过早衰老等表型^[31]。该综述整合了目前食物疗法的基础及临床研究,填补了食物组与代谢性疾病临床管理研究和开发的不足,为疾病精准分型和分级管理提供了建议。

4、过敏性皮炎严重程度的多因素分析

瑞士苏黎世大学、德国波恩大学、慕尼黑工业大学医学院等多家环境医学和皮肤过敏研究机构联合撰文,通过对367位青少年和成人过敏性皮炎患者长达四年多的前瞻性纵向研究,利用深度表型测量采集过敏性皮炎表型,采用湿疹面积和严重程度指数(EASI)对患者按严重程度组进行分层,并采用基于交叉验证优化的机器学习梯度推进方法以及回归分析模型,分析得到了130个与过敏性皮炎严重程度相关的因素,如血清IgE>1708IU/ml和嗜酸性粒细胞值大于6.8%、患者发病年龄(12-21岁或52岁以上)、主动吸烟或被动吸烟暴露以及斑秃等因素与重度过敏性皮炎相关^[32]。该研究发现过敏性皮炎严重程度关联有助于深入了解疾病发生发展机制,助力个性化预防、治疗策略的制定。

5、揭示肝脏疾病发病风险

复旦大学人类表型组研究院陈兴栋研究员团队在国际权威胃肠肝病学期刊 Clinical Gastroenterology & Hepatology 发布代谢相关脂肪肝病发病风险因素研究成果,利用基因型数据,构建了脂肪肝相关的多基因风险评分(GRS),并发现较高的GRS能够显著放大代谢性脂肪肝病对于肝脏相关疾病的风险^[33]。这是首次利用大型人群队列证实了代谢相关脂肪性肝病这一最新疾病表型的远期健康效应,为该疾病2020年的更名(由非酒精性脂肪

肝改名为代谢相关脂肪性肝病) 提供了理论依据。该团队还在国际权威流行病学期刊 *International Journal of Epidemiology* 上发布了肝癌发病预测模型, 利用全球疾病负担 (Global Burden of Disease) 数据和贝叶斯年龄 - 时期 - 队列模型预测了不同病因导致的肝癌的发病率和发病人数^[34]。

(五) 疾病表型组学研究进展

疾病人群和医学健康人群的深度表型测量, 将疾病表型与健康人表型进行详细测量、量化和数据搜集, 对表型组数据归类和特征分析, 将绘制人 / 人群健康基线, 及时发现健康状态偏离, 发现疾病早期生物标志物, 助力疾病早期诊断及患者管理。

1、人群深度表型测量实现健康到疾病转化的早诊

系统生物研究所 Leroy Hood 院士团队从商业性健康项目参与者血浆样本中蛋白组学变化发现了肺癌、乳腺癌、胰腺癌和造血系统恶性肿瘤的转移癌生物标记物, 将转移癌确诊时间提前 10 个月以上^[35]。该团队对于纵向多组学测量数据进行降维分析后, 得出四种典型表型特征揭示健康多组学状态, 通过参与者纵向数据在该模型上的移动可监测个体身体状况, 实现健康向疾病转变的发现^[36]。这些研究利用深度表型测量为健康与疾病之间的关系提供见解, 是科技创新与数据驱动优化健康解决方案的成功案例。

2、心血管疾病的表型组学研究

心血管疾病是近二十年来增长迅速的疾病, 造成了全球疾病负担。影像技术的可靠性和超声系统的便携性与日俱增, 为心血管疾病表型组检测提供了坚实基础, 使心血管疾病发病和病情发展预测成为可能。2021 年度, 心血管领域表型组学研究在表型检测方法、疾病跨尺度关联及患者精准医学管理等多领域有多篇高引用综述发表, 体现了表型组学优化疾病精准医学防治方案的重大潜力。

光学相干断层扫描 (Optical coherence tomography, OCT) 使动脉粥样硬化斑块的体内诊断成为可能, 动脉粥样硬化斑块的血管内成像可识别急性冠脉综合征患者两种主要光学表型: 斑块侵蚀与斑块破裂^[37]。在结构表型、功能表型、组织特征、血流动力学和分子代谢表型层面的心脏影像表型的标准化提取方案正式发布^[38]。在临床实践中, 人工智能增强心电图在心血管疾病管理中已处于起步阶段, 正在持续发展^[39]。

哈佛医学院麻省总医院介入心脏病专家、临床试验项目主任张益京博士团队在 *Nature Reviews Cardiology* 上刊发综述, 基于动脉粥样硬化斑块侵

蚀患者表型和分子特征，提出 OCT 血管内成像有助于对斑块进行高分辨率活检，可区分斑块侵蚀区和斑块破裂，对斑块侵蚀患者可使用药物治疗，避免过度医疗^[37]，集合临床表型、影像表型、代谢表型、环境暴露、营养摄取等多个复杂干预措施的心脏康复，将助力改善心血管疾病患者结局，减轻家庭乃至国际心血管疾病负担^[37]。

3、无创早期检测阿尔兹海默症

澳大利亚国家表型组中心通过综合代谢表型检测平台——核磁共振（NMR）及超高效液相色谱-质谱联用（UHPLC-MS），以简单无创的方式来检测人尿液中的代谢表型，发现 23 个代谢标志物可用于辨别阿尔兹海默症的早期症状，并预测早期临床疾病进程^[35]。

4、血浆临床表型预测坏死性软组织感染

瑞典斯德哥尔摩卡罗琳斯卡研究所联合多所欧洲科研院所在 The Journal of Clinical Investigation 期刊上刊发血浆表型预测坏死性软组织感染方面的科研论文。通过前瞻性多中心坏死性软组织感染患者队列确定与临床表型相关的生物标志物血栓调节蛋白（Thrombomodulin），确认 G-CSF、S100A8、IL-6 是脓毒症休克预测标志物，并在患者血浆样本中得到验证^[40]。该研究对坏死性软组织感染的早期诊断、预后和治疗方法具有极大的价值。

5、为多种癌症患者分层管理提供策略

英国巴斯大学和巴西圣保罗大学科研团队在 Frontiers in Immunology 刊发综述，提出整合表型学方法到乳腺癌临床管理及肿瘤学研究，将目前临床研究中较少被列入病因和预后指标的衰老、炎症、生活方式等进行客观测量并整合入乳腺癌治疗和护理方法，将提升乳腺癌治疗效果^[41]。美国西奈山伊坎医学院精准免疫研究所所长 Miriam Merad 团队通过免疫细胞表型分析建立非小细胞癌早期肺癌免疫反应细胞图谱，并提出了肿瘤详细分型的肺癌激活模块（Lung cancer activation module, LCAM），可作为衡量抗原特异性抗肿瘤免疫激活指标^[6]，助力肺癌精准免疫治疗。新加坡国立大学研究团队采用综合代谢组学方法和不同液相色谱法结合提取式高分辨率质谱仪进行肺癌的血浆代谢谱分析，提出 Palmitic acid, Heptadecanoic acid, 4-Oxoproline, Tridecanoic acid, and Ornithine 是鉴定罹患肺癌最有效的代谢物，同时 Palmitic acid, Heptadecanoic acid, Ornithine, Pentadecanoic acid, and Acylcarnitine C8:1 被认为是肺癌分型的有效指标^[42]。该研究团队基于此提出高性能的逻辑回归模型，可用于诊断癌症、癌症分期和癌症类型，

未来可用于研发微创、高性能、高通量和低成本的肺癌筛查试验方法。系统生物学研究田强教授科研团队针对肿瘤免疫微环境的研究，为微卫星稳定性结直肠癌肝转移中肿瘤免疫检查点的化疗调节提供证据，为结直肠癌患者化学 - 免疫治疗设计提供了策略^[43]。

6、辅助制定呼吸系统临床治疗方案

匹兹堡大学哮喘研究所所长 Sally Wenzel 在美国呼吸与危重症医学杂志发表综述,提出整合临床表型特征改善重度成人哮喘结局,对于咳嗽、胸闷、痰分泌、呼吸急促等临床表型及食管反流、肥胖、声带功能障碍等并发症的检测对哮喘类型与严重程度有重要参考意义,为临床哮喘病人诊断与治疗提供了临床免疫表型检测指标及策略^[44]。反向触发(reverse triggering)呼吸模式失调为患者呼吸表型将助力呼吸窘迫综合征患者临床精准分类,对制定危重患者的个体化肺保护性机械通气策略至关重要^[45]。波士顿大学、波士顿医疗中心与韩国首尔成均馆大学联合开发出可预测罹患哮喘风险与量化儿童哮喘严重程度的表型计算方法^[46]。哮喘严重程度的表型计算方法可避免临床医生因不同指南的哮喘严重程度定义无法统一而导致治疗方案混乱的问题,提示电子健康记录数据在患者临床不良事件预测上可能具有特殊意义。

7、多发硬化症认知表型助力疾病分型

意大利佛罗伦萨大学科研团队通过核磁共振成像,在 1,212 名多发性硬化症患者和 196 名健康对照者的横向研究中,结合认知评估方法,确定了多发性硬化症 5 种认知表型,每种分型都具有特定的临床和磁共振影像学特征^[47],为临床诊断提供了新的思路,为治疗方法选择提供了理论依据,也将助力患者认知康复治疗方案。

(六) 新技术研发

1、人工智能 CARE 系统筛选眼底异常表型

中山大学中山眼科中心林浩添教授团队在国际顶级期刊 The Lancet Digital Health 发布可以识别 14 种常见眼底异常的眼底疾病人工智能诊断专家—CARE,并在全国 35 家不同地区不同级别的医疗机构进行前瞻性临床真实世界验证。实验结果表明,CARE 适用于不同医疗场景、不同种族、不同眼底照相仪器的临床真实世界环境,且准确率媲美眼底专科医生^[48]。此外,CARE 的单卷积神经网络结构消耗更少的计算资源,在普通家用电脑就可以离线运行,减少对连接服务器网络的依赖,适用于医疗资源稀缺和网络条件差的老少边穷地区的眼底疾病筛查。

2、基于血液液体活检的红外分子指纹癌症检测技术

慕尼黑大学的研究团队通过“Onco-IR-Phenotyping”的高通量表型测量技术，用傅里叶变换红外透射光谱检测 1927 名参与者的血清和血浆，再经过机器学习来训练二元分类算法模型检测效率与疾病的相关性。团队发现光谱特征在不同的癌症类型之间是不同的，且红外指纹检测效率会受疾病严重程度的影响，机器学习算法使用的信号强度会随着疾病进展阶段的增加而增加（肿瘤体积、转移扩散或肿瘤分级评分）^[49]，为癌症识别诊断提供了低成本、互补的分析工具。

3、自动化智能表型可塑性平台

辛辛那提大学与北卡罗来纳大学教堂山分校合作发布可用于运动伤害的自动化智能表型可塑性平台 Intelligent Phenotypic Plasticity Platform (IP3)，通过表型分析对运动员健身前后的表型可塑性进行索引，从而推动个性化精准治疗方案制定^[50]。这是第一个基于表型可塑性的精准医疗平台，也是第一个多类型的精准运动医学平台。

4、超极化 ¹²⁹Xe 磁共振成像测量预测儿童哮喘发展情况

来自辛辛那提儿童医院医学中心的团队，对 37 名哮喘青少年进行了 ¹²⁹Xe 磁共振成像，结果与哮喘的严重程度、就医次数以及口服皮质类固醇的使用频率均有关联^[51]。通常儿童的哮喘严重程度的判定较为困难，¹²⁹Xe 磁共振成像产生的临床表型具有成为生物标志物的可能性，有助于在早期筛选出病情风险高的儿童并辅助改善预后的恢复效果。

五、人类表型组：总结与展望

2021 年是国际人类表型组计划发展与初步收获的一年。人类表型组研究范式在生命科学前沿探索中重要作用凸显：首张人类表型组导航图发布；表型组领域标准创制取得了突破进展；表型组研究助力提升新冠肺炎、癌症、心血管疾病、免疫疾病等疾病的风险预警、早期诊断及治疗管理方案；表型组 - 基因组关联机制研究取得多项成果；多项表型测量技术及技术系统研发问世。人类表型组学入选 2021 医药卫生领域全球工程前沿 TOP10，在生命科学领域的重要性成为学界共识。

人类表型组研究的关键领域也在不断扩大：更精密的人类表型组测量平台、表型组学队列研究、质量控制与标准化工作、表型组大数据计算与分析将为人类表型组研究深度赋能。大型队列研究积累了生命科学大数据，大数据推动了医疗保健研究的进步，使得大范围数据间关联研究成为可能，为人类表型组研究的跨域式发展提供了机遇。通过人类表型组研究进行表型精准测量，探索基因 - 环境表型到个体健康的表型间跨尺度关联，发现疾病起源及多样性，研发新技术和方法对疾病进行精细诊断、合理治疗和有效预防，将创造改善人类健康的新范式。相信随着人类表型组国际大科学计划不断推动领域国际合作与交流，必将引领生命科学研究范式变革和生物医药产业发展突破，实现服务人类健康的目标。

参考文献

- [1] 金力. 寻找健康“密钥”: 从基因组到表型组 [J]. 新华文摘, 2022, (8): 5.
- [2] WIERZBICKI A S. Phenomics, Not Genomics, for Cardiovascular Risk Assessment [J]. Circulation, 2020, 142(9): 821-3.
- [3] WRIGHT J T, HERZBERG M C. Science for the Next Century: Deep Phenotyping [J]. J Dent Res, 2021, 100(8): 785-9.
- [4] NETWORK B I C C. A multimodal cell census and atlas of the mammalian primary motor cortex [J]. Nature, 2021, 598(7879): 86-102.
- [5] JIANG L, WANG M, LIN S, et al. A Quantitative Proteome Map of the Human Body [J]. Cell, 2020, 183(1): 269-83 e19.
- [6] LEADER A M, GROUT J A, MAIER B B, et al. Single-cell analysis of human non-small cell lung cancer lesions refines tumor classification and patient stratification [J]. Cancer Cell, 2021, 39(12): 1594-609 e12.
- [7] ENGINEERING P G O G E F O C A O. Engineering Frontiers 2021 [M]. Higher Education Press, 2021.
- [8] TIAN M, LIU H, CHEN S, et al. Report on the 3rd Board Meeting of the International Human Phenome Consortium [J]. Phenomics, 2022.
- [9] CHEN W, ZHAO Y, CHEN X, et al. A multicenter study benchmarking single-cell RNA sequencing technologies using reference samples [J]. Nat Biotechnol, 2021, 39(9): 1103-14.
- [10] DEVESON I W, GONG B, LAI K, et al. Evaluating the analytical validity of circulating tumor DNA sequencing assays for precision oncology [J]. Nat Biotechnol, 2021, 39(9): 1115-28.
- [11] FANG L T, ZHU B, ZHAO Y, et al. Establishing community reference samples, data and call sets for benchmarking cancer mutation detection using whole-genome sequencing [J]. Nat Biotechnol, 2021, 39(9): 1151-60.
- [12] FOOX J, TIGHE S W, NICOLET C M, et al. Performance assessment of DNA sequencing platforms in the ABRF Next-Generation Sequencing Study [J]. Nat Biotechnol, 2021, 39(9): 1129-40.

- [13] XIAO W, REN L, CHEN Z, et al. Toward best practice in cancer mutation detection with whole-genome and whole-exome sequencing [J]. *Nat Biotechnol*, 2021, 39(9): 1141-50.
- [14] Sequencing benchmarked [J]. *Nat Biotechnol*, 2021, 39(9): 1027.
- [15] SALIT M, WOODCOCK J. MAQC and the era of genomic medicine [J]. *Nat Biotechnol*, 2021, 39(9): 1066-7.
- [16] HOLMES E, WIST J, MASUDA R, et al. Incomplete Systemic Recovery and Metabolic Phenoreversion in Post-Acute-Phase Nonhospitalized COVID-19 Patients: Implications for Assessment of Post-Acute COVID-19 Syndrome [J]. *J Proteome Res*, 2021, 20(6): 3315-29.
- [17] LODGE S, NITSCHKE P, KIMHOFFER T, et al. Diffusion and Relaxation Edited Proton NMR Spectroscopy of Plasma Reveals a High-Fidelity Supramolecular Biomarker Signature of SARS-CoV-2 Infection [J]. *Anal Chem*, 2021, 93(8): 3976-86.
- [18] NICHOLSON J K. Molecular Phenomic Approaches to Deconvolving the Systemic Effects of SARS-CoV-2 Infection and Post-acute COVID-19 Syndrome [J]. *Phenomix*, 2021, 1(4): 143-50.
- [19] STUKALOV A, GIRAULT V, GRASS V, et al. Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV [J]. *Nature*, 2021, 594(7862): 246-52.
- [20] LEE J W, SU Y, BALONI P, et al. Integrated analysis of plasma and single immune cells uncovers metabolic changes in individuals with COVID-19 [J]. *Nat Biotechnol*, 2022, 40(1): 110-20.
- [21] SU Y, CHEN D, YUAN D, et al. Multi-Omics Resolves a Sharp Disease-State Shift between Mild and Moderate COVID-19 [J]. *Cell*, 2020, 183(6): 1479-95 e20.
- [22] YING K, ZHAI R, PYRKOV T V, et al. Genetic and phenotypic analysis of the causal relationship between aging and COVID-19 [J]. *Commun Med (Lond)*, 2021, 1: 35.
- [23] HE F, ZHANG T, XUE K, et al. Fecal multi-omics analysis reveals diverse molecular alterations of gut ecosystem in COVID-19 patients [J]. *Anal Chim Acta*, 2021, 1180: 338881.
- [24] RIETDIJK J, TAMPERE M, PETTKE A, et al. A phenomics approach for antiviral drug discovery [J]. *BMC Biol*, 2021, 19(1): 156.

- [25] HUANG S, WANG Y, ZHOU Z, et al. Distribution Atlas of COVID-19 Pneumonia on Computed Tomography: A Deep Learning Based Description [J]. *Phenomics*, 2021, 1(2): 62-72.
- [26] DENG L, PAN Y, WANG Y, et al. Genetic Connections and Convergent Evolution of Tropical Indigenous Peoples in Asia [J]. *Mol Biol Evol*, 2022, 39(2).
- [27] LI Z, XIA J, JIANG L, et al. Characterization of the human skin resistome and identification of two microbiota cutotypes [J]. *Microbiome*, 2021, 9(1): 47.
- [28] GEORGIU M, FINOCCHIO L, FUJINAMI K, et al. X-Linked Retinoschisis: Deep Phenotyping and Genetic Characterization [J]. *Ophthalmology*, 2022, 129(5): 542-51.
- [29] ANDREWS S J, FULTON-HOWARD B, O'REILLY P, et al. Causal Associations Between Modifiable Risk Factors and the Alzheimer's Phenome [J]. *Ann Neurol*, 2021, 89(1): 54-65.
- [30] RAMRATNAM S K, LOCKHART A, VISNESS C M, et al. Maternal stress and depression are associated with respiratory phenotypes in urban children [J]. *J Allergy Clin Immunol*, 2021, 148(1): 120-7.
- [31] MAFRA D, BORGES N A, LINDHOLM B, et al. Food as medicine: targeting the uraemic phenotype in chronic kidney disease [J]. *Nat Rev Nephrol*, 2021, 17(3): 153-71.
- [32] MAINTZ L, WELCHOWSKI T, HERRMANN N, et al. Machine Learning-Based Deep Phenotyping of Atopic Dermatitis: Severity-Associated Factors in Adolescent and Adult Patients [J]. *JAMA Dermatol*, 2021, 157(12): 1414-24.
- [33] LIU Z, SUO C, SHI O, et al. The Health Impact of MAFLD, a Novel Disease Cluster of NAFLD, Is Amplified by the Integrated Effect of Fatty Liver Disease-Related Genetic Variants [J]. *Clin Gastroenterol Hepatol*, 2022, 20(4): e855-e75.
- [34] LIU Z, XU K, JIANG Y, et al. Global trend of aetiology-based primary liver cancer incidence from 1990 to 2030: a modelling study [J]. *Int J Epidemiol*, 2021, 50(1): 128-42.
- [35] MAGIS A T, RAPPAPORT N, CONOMOS M P, et al. Untargeted longitudinal analysis of a wellness cohort identifies markers of metastatic cancer years prior to diagnosis [J]. *Sci Rep*, 2020, 10(1): 16275.
- [36] ZIMMER A, KOREM Y, RAPPAPORT N, et al. The geometry of clinical labs and wellness states from deeply phenotyped humans [J]. *Nat Commun*, 2021, 12(1): 3578.

- [37] FAHED A C, JANG I K. Plaque erosion and acute coronary syndromes: phenotype, molecular characteristics and future directions [J]. *Nat Rev Cardiol*, 2021, 18(10): 724-34.
- [38] WANG C, LI Y, LV J, et al. Recommendation for Cardiac Magnetic Resonance Imaging-Based Phenotypic Study: Imaging Part [J]. *Phenomics*, 2021, 1(4): 151-70.
- [39] SIONTIS K C, NOSEWORTHY P A, ATTIA Z I, et al. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management [J]. *Nat Rev Cardiol*, 2021, 18(7): 465-78.
- [40] PALMA MEDINA L M, RATH E, JAHAGIRDAR S, et al. Discriminatory plasma biomarkers predict specific clinical phenotypes of necrotizing soft-tissue infections [J]. *J Clin Invest*, 2021, 131(14).
- [41] ARANA ECHARRI A, BERESFORD M, CAMPBELL J P, et al. A Phenomic Perspective on Factors Influencing Breast Cancer Treatment: Integrating Aging and Lifestyle in Blood and Tissue Biomarker Profiling [J]. *Front Immunol*, 2020, 11: 616188.
- [42] QI S A, WU Q, CHEN Z, et al. High-resolution metabolomic biomarkers for lung cancer diagnosis and prognosis [J]. *Sci Rep*, 2021, 11(1): 11805.
- [43] JABBARI N, KENERSON H L, LAUSTED C, et al. Modulation of Immune Checkpoints by Chemotherapy in Human Colorectal Liver Metastases [J]. *Cell Rep Med*, 2020, 1(9): 100160.
- [44] WENZEL S E. Severe Adult Asthmas: Integrating Clinical Features, Biology, and Therapeutics to Improve Outcomes [J]. *Am J Resp Crit Care*, 2021, 203(7): 809-21.
- [45] BAEDORF KASSIS E, SU H K, GRAHAM A R, et al. Reverse Trigger Phenotypes in Acute Respiratory Distress Syndrome [J]. *Am J Respir Crit Care Med*, 2021, 203(1): 67-77.
- [46] PEER K, ADAMS W G, LEGLER A, et al. Developing and evaluating a pediatric asthma severity computable phenotype derived from electronic health records [J]. *J Allergy Clin Immunol*, 2021, 147(6): 2162-70.
- [47] DE MEO E, PORTACCIO E, GIORGIO A, et al. Identifying the Distinct Cognitive Phenotypes in Multiple Sclerosis [J]. *JAMA Neurol*, 2021, 78(4): 414-25.
- [48] LIN D, XIONG J, LIU C, et al. Application of Comprehensive Artificial intelligence Retinal Expert (CARE) system: a national real-world evidence study [J]. *Lancet Digit Health*, 2021, 3(8): e486-e95.

- [49] HUBER M, KEPESIDIS K V, VORONINA L, et al. Infrared molecular fingerprinting of blood-based liquid biopsies for the detection of cancer [J]. *Elife*, 2021, 10.
- [50] KIEFER A W, ARMITANO-LAGO C N, SATHYAN A, et al. The Intelligent Phenotypic Plasticity Platform (IP(3)) for Precision Medicine-Based Injury Prevention in Sport [J]. *Methods Mol Biol*, 2022, 2393: 877-903.
- [51] LIN N Y, ROACH D J, WILLMERING M M, et al. Xe-129 MRI as a measure of clinical disease severity for pediatric asthma [J]. *J Allergy Clin Immun*, 2021, 147(6): 2146-+.