

动物复杂性状的进化解析与调控

1 立项背景及意义

动物复杂性状是动物长期适应进化的结果，是动物多样性存在的主要基础，其调控的失衡是人类重大慢性病发生的内在原因，对其形成规律的认识对家养动物经济性状的改良和动物特殊功能的仿生有重要意义。因此，系统解析动物复杂性状的成因不仅是一项揭示大自然本质规律的基础性前沿科学工作，也有重大的应用前景。动物复杂性状形成机制的解析一直是一个科学难题，面临难以追溯、难以预测、难以调控的困难局面。

“动物复杂性状的进化解析与调控”先导专项试图通过大尺度、跨物种的进化比较，系统整合不同动物物种遗传因子-发育网络进化-表型适应性3个层面的全面数据，突破目前GWAS（全基因组关联分析）的框架，解析动物复杂性状的成因，我们把这种研究动物复杂性状的新思路称为eGPS（evolutionary Genotype-Phenotype Systems biology）。

专项以动物脑容量大小、高原适应和脂肪能量代谢等复杂性状产生的内在规律作为贯穿整个研究的主线，从遗传、发育调控网络进化、功能适应性3个层面，层层递进，解析动物复杂性状的成因。在遗传层面阐明驱动复杂性状进化的主要内在遗传变异力量的进化模式，然后鉴定决定动物脑容量大小、高原适应和脂肪能量代谢等动物复杂性状的遗传元件，在发育网络调控进化层面阐明这些元件如何组装，在功能研究层面主要用功能

操控实验验证前面两个层面的结果，并理解这些元件组成和相互作用的适应性意义和平衡状态破坏后导致不适应（如疾病）的机制，即分析性状进化的进化成本或代价。阐明驱动复杂性状进化的主要内在遗传变异力量的进化模式，进行eGPS模型的建立，阐明动物复杂性状的成因，实现动物复杂性状可追溯、可预测、可调控的目标。

2 已取得的进展

专项自2014年6月启动以来，取得了一系列重要的研究进展，阶段性完成了原有计划任务目标。专项首先搭建了非模式动物（猴子和蝴蝶）的大规模基因遗传操作平台，以及世界上类群最丰富的高原动物遗传资源库（>20 000样本量和复杂性状数据）；其次，突破常规方法的限制，建立了一套复杂动物基因组测序、组装和分析的新方法；另外，建立了分析动物复杂性状多层次、高维度海量数据的方法体系和模型。在此基础上，在动物脑容量大小、高原适应和脂肪能量代谢3个代表性复杂性状的进化机制解析方面取得了突破，已发表论文154篇，其中37篇发表在国际相关领域顶级期刊上，如*Science*、*Nature Genetics*、*PNAS*、*PLoS Biology*、*Neuron*等，专项代表性成果详细介绍如下。

2.1 建立了类群最丰富的高原动物遗传资源样本库和表型数据统计库

为了解析高原动物的适应性进化机制和能量代谢调

控机理，专项收集并建立了世界上迄今类群最丰富的高原动物遗传资源样本库，包括了3 000余份野生青藏高原脊椎动物（鱼类、两栖类、爬行类、鸟类和哺乳动物等）DNA样本，1 000余份青藏高原家养动物（猪、马、驴、牛、狗和牦牛等）样本，测量并统计了66个高原动物种群在不同海拔高度的27个生理指标，包括血细胞参数、血气参数和心脏重量指数等。此外，从65个地理种群收集到16 206份藏族人群血样和250份胎盘组织及105个从新生儿脐静脉原代培养细胞系。为构建eGPS的运算和分析数学模型及解析高原动物（含人类）的高原适应分子机制奠定了基础。

2.2 动物高原适应的遗传机制研究取得重要突破

除了重要保护生物学价值，金丝猴属物种不仅发展出以树叶为食的特化食性，而且占据了从低海拔到高海拔（800 m—4 500 m）的生境类型，为研究动物对高海拔环境适应性进化遗传机制提供了很好的模型。专项团队对金丝猴属物种高海拔环境适应遗传机制进行了系统分析，发现滇金丝猴中与DNA修复和氧化磷酸化过程相关的基因显著扩张，在3个高海拔金丝猴物种（滇金丝猴、怒江金丝猴和川金丝猴）中发现6个基因中有8个共有氨基酸替换，这些基因与肺功能、DNA修复和血管生成相关，并证实其中的*RNASE4*基因突变型在诱导血管生成方面具有更高的活性。这一研究以非人灵长类为研究模型，对高海拔适应这一复杂性状作了一个新的和更全面的揭示。

利用eGPS的研究体系，专项团队确定了几个与高原适应密切相关的新的基因，并对这些基因相关的体内和体外功能进行了研究。例如：将青藏高原的羚羊、牦牛、鼢鼠和鼠兔的基因组和转录组数据进行了分析，并与它们对应的平原物种组学数据进行了类比，在青藏高原哺乳动物的一个代谢相关酶基因中发现了高原物种特有的突变，在体外细胞实验中发现突变蛋白的酶活性显著高于野生型蛋白。构建了遗传修饰小鼠，发现携带该突变的小鼠在低氧下存活力显著长于野生型小鼠，其主

要脏器的携氧量也比野生型小鼠高，表明该基因可能是高原动物适应相关的重要基因。

专项还系统性地解析了青藏高原鱼类、两爬类和藏鸡的高原适应分子机制。除了在基因组上的发现，专项在高原牦牛等动物的微生物菌群中也发现了高原特异的适应性类群，eGPS的方法学还被应用在脑的形态学分析。研究发现，磷脂组重塑在人脑功能的进化过程中起着十分重要的作用，人脑中脂质的富集相比非神经系统组织的脂质富集要高5倍。

2.3 动物能量代谢机制的演化

羊是肉、奶和毛纺纤维的主要来源，并代表了一类利用专门的消化器官（瘤胃）对植物食材进行初步消化的特殊动物。专项从40个不同的绵羊组织中收集并获得了高质量的基因组和转录组数据，并发现了许多参与脂质代谢的基因发生了扩增或不同的组织表达模式，这可能是响应皮肤的脂质代谢屏障变化以及羊毛合成的结果（图1）。



图1 羊脂代谢特殊途径

大熊猫喜欢吃竹子，主要通过缩小一些重要器官的大小和减少日常运动来保持能量平衡。专项组成员测量了5只圈养和3只野生大熊猫的能量代谢情况，结果表明，大熊猫在*DUOX2*基因编码序列上的终止密码子突变，导致其外周血中的甲状腺激素偏低、对能量的需求低，因而从形态、行为、生理和遗传等角度，对大熊猫

主要靠竹子维持生存和维持能量的机制进行了全面的解析（图2）。

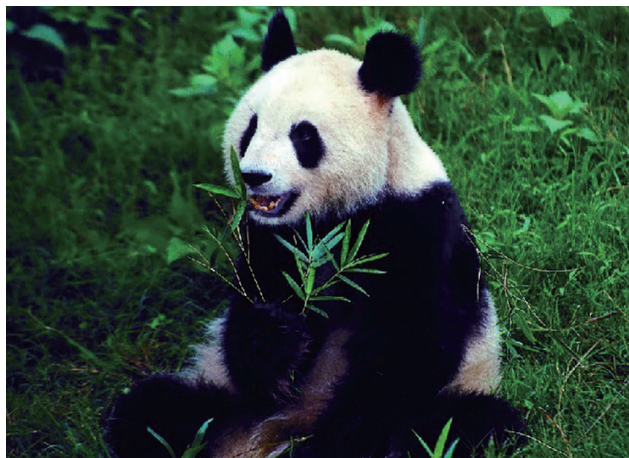


图2 大熊猫异常低能量代谢

3 独创性

（1）建立了eGPS研究的新型研究模式和计算方法。利用跨物种和跨组织的大规模数据，专项团队成员开发出新的基于动物eGPS研究的通用定量模型和计算方法（基因型表型的进化和系统生物学）。专项开发了基于动态网络（Dynamic Network Biomarkers, DNB）生物标志物的新理论框架，用来量化从个性状转化到另一个性状的关键节点；提出了部分相互独立信息的概念（Part Mutual Information, PMI），PMI的方法能有效地从已有的组学数据中重新构建大规模生物信息网络，如基因调控网络和蛋白质相互作用信息网。这些新的方法和模型已应用于对藏族人群的高原适应机制的解析工作。

（2）搭建非模式动物遗传修饰研究平台。全面突破了猴子、树鼩转基因和基因编辑技术，获得了世界上迄今数量最多的大脑发育转基因和基因敲除猴（图3），建立了专项迫切需要的关键技术平台。成功建立了利用慢病毒和CRISPR-Cas9技术生产转基因和基因敲除猴的技术体系和实验平台，并批量构建了40余只转基因或基因敲除猴，并通过精原细胞替换在世界上首次成功实现树

鼩基因编辑。同时，也成功利用CRISPR-Cas9技术对蝴蝶中的基因进行了编辑。



图3 转基因和基因敲除猴

（3）构建了组学数据库系统。建立了一套具有中国自主知识产权的组学数据库系统（Genome Sequence Archive, GSA），且已获得国际多家学术期刊认可，GSA数据库系统在国内实现了“零”的突破，在国际上将逐步打破国际核酸序列联盟（INSDC）对组学数据的垄断，并使得中国在国际组学数据管理与共享方面占据一席之地。因此，具有重要的社会价值和影响。同时，通过对数据与知识的积累，未来通过对组学大数据的深度挖掘与利用，可以在生物育种、人口健康、重大疾病等多个方面深入应用，并将产生重要的经济价值。

4 对产业的意义

疾病可视为一类特殊的复杂性状，专项建立的eGPS研究体系对于充分挖掘人类疾病相关大数据，揭示相关疾病发生发展的调控网络，提供新的诊断和治疗靶点和方案都具有重要价值。专项建立了非模式动物（猴子、树鼩等）转基因遗传操作及功能验证平台，将为后续的大脑容量和结构复杂化这一复杂性状的功能解析提供资源保障，对人类疾病模型的创建和药物研发具有重要意义。

5 对我国学科、产业推进、人才培养等未来部署的建议

以基因组学为代表的组学技术的迅猛发展，产生了海量的数据，为系统解析动物复杂性状的成因提供了条件。本专项建立的 eGPS 技术，为系统挖掘、整合海量组学数据、揭示复杂性状的调控网络提供了重要平台。从学科角度，在基因组时代的大背景下，“遗传、发育与进化”的综合研究领域已成为国际研究热点之一，取得了一系列重要成果。近年来我国在该学科领域科研总体

水平正在迅速提高，并受到国际同行的高度关注。当前是遗传发育与进化研究的关键时期，也是取得重大原创性突破的最佳时机，我国科学家在这方面有很好的积累及特色，且在某些方向已达到世界一流水平。建议集中我国在相关领域的优秀研究力量，围绕生物适应进化领域最重大的核心科学问题和国家最迫切的重大需求，开展全基因组水平上生物适应性的分子机制研究，提高我国生命科学基础创新能力，增强我国科技整体竞争力。

（依托单位：中科院昆明动物所）

专家点评

这是一个创新性非常强的项目，旨在解析跨物种间的动物复杂性状的进化机制。通过整合最前沿的各种组学分析手段的 eGPS 方法，研究团队在动物脑容量大小、高原适应和脂肪能量代谢 3 个代表性复杂性状的进化机制解析方面取得了重大突破，已发表论文 154 篇，其中 37 篇发表在国际相关领域顶级期刊上，如 *Science*、*Nature Genetics*、*PNAS*、*PLoS Biology*、*Neuron* 等。这些研究成果将为我们解析生理、形态等性状的遗传学基础提供科学依据，同时为人类将来靶向治疗肥胖相关的人类疾病研究提供基础科学数据。更重要的是，该专项将为动物复杂性状包括疾病发生的遗传机制解析提供前所未有的方法学平台，特别是在整合跨物种、多组学数据的模型和方法方面进行了重点部署，提出了一些新的方法和模型，具有重要意义。如能进一步整合、发展成为完整的理论系统和方法体系，对生命科学其他领域的研究也将具有重要的借鉴意义。

点评专家

李 蓬 中科院院士，发展中国家科学院院士，清华大学生命科学院教授，国家自然科学基金委生命学部主任，长江学者特聘教授，长江学者创新团队负责人，国务院特殊津贴专家，“973”计划首席科学家，国家杰出青年基金获得者。主要研究脂代谢的调控机制和代谢

性疾病如肥胖症，糖尿病和脂肪肝的发生发展机理。在国际顶尖杂志如 *Cell*、*Nature Medicine*、*Nature Cell Biology*、*Cell Metabolism*、*Genes & Development* 等上发表 60 余篇有影响力的文章。

专家点评

通过前两年的努力奋斗，该专项整体已经取得了卓越的成果。包括发表了一些高水平的学术论文，建立了一些独创性的理论、方法、技术，如建立了 eGPS 研究的新型研究模式和计算方法、搭建了非模式动物遗传修饰研

究平台、构建了组学数据库系统（Genome Sequence Archive, GSA）。专项在非模式动物体中使用 eGPS 方法在能量平衡和脂肪储存调控机制和情况方面已经有了开创性发现，3 篇重要的论文发表在 *Science* 上，1 篇发表在 *PLoS Genetics* 上。主要发现包括猎豹捕食活动和耗能方面的证据，大熊猫在低能量的竹子饮食下存活的适应性机制，羊皮肤里新的脂代谢通路，在昆虫和其他节肢动物中脂信号的新的遗传学研究。

点评专家

杨洪远 澳大利亚新南威尔士大学教授。主要从事代谢性疾病发生和发展机制的研究，包括脂肪发育，脂肪细胞和肝细胞中的脂肪积累、分泌和分解的分子机制与

胰岛素敏感性关系，以及细胞内脂质运输等领域。近年来在 *Cell*、*Cell Reports*、*Diabetes*、*Journal of Cell Biology* 等各种期刊发表文章 68 篇。