科学家构建藏族人群全基因组水平的适应性遗传变异图谱

作者：writer 来源：中国科学院

本文原地址：https://www.iikx.com/news/progress/6375.html

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

科学家构建藏族人群全基因组水平的适应性遗传变异图谱。8月7日，《国家科学评论》(National Science Review)在线发表了中国科学院上海营养与健康研究所——马普计算生物学研究所徐书华团队的研究成果“ Prioritizing natural selection signals from the deep-sequencing genomic data suggests multi-variant adaptation in Tibetan highlanders ”。该研究通过分析深度基因组测序数据和藏族表型数据，构建了藏族人群全基因组水平的适应性遗传变异图谱，首次系统地将藏族人群基因组中与适应性进化相关的功能性变异呈现出来。

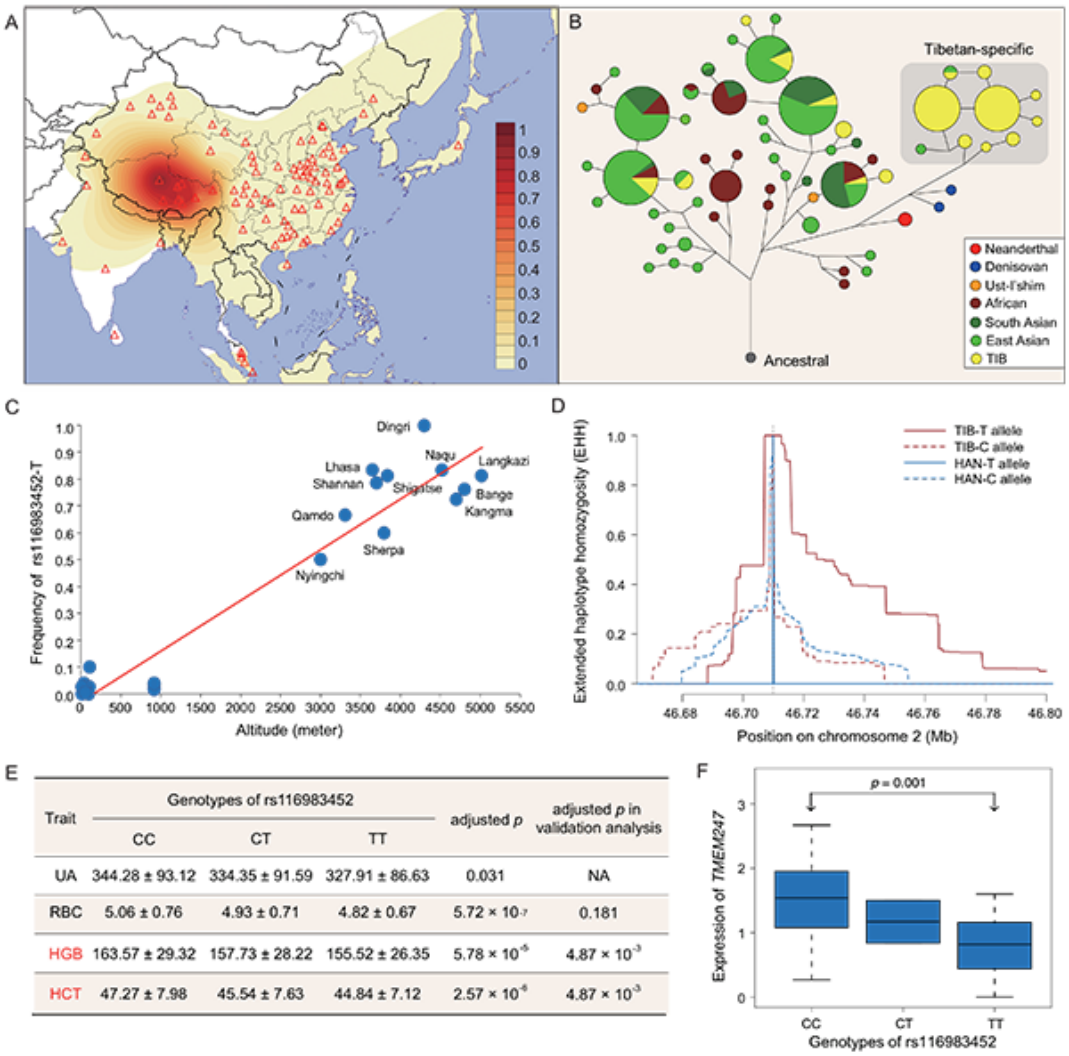
经过国内外近十年的密集研究，人们对藏族高原适应的遗传学基础有了一些初步的认识;其中EPAS1目前领域里普遍认为是藏族适应高原的关键基因，特别是由于在其他高原物种中也发现EPAS1的适应性进化的迹象，因而备受关注。但是迄今未能确定EPAS1基因中与藏族高原适应的功能性变异。这为理解人类在青藏高原上的适应性进化机制留下未解难题。由此也引发一些更根本性的关键问题：(1)有多少基因和遗传变异驱动了人类适应青藏高原的演化?(2)EPAS1在人类高原适应性中发挥作用的关键因素是否并不在EPAS1本身而是否另有出处?(3)是否还有其他基因比EPAS1对人类适应高原起到更关键的作用?

带着这些问题，徐书华团队与中科院昆明动物研究所、温州医科大学、复旦大学、西藏民族大学等多家单位的研究人员合作，在全基因组水平对藏族人群的高原适应性变异进行了系统性梳理，充分利用深度基因组测序数据的优势，构建了藏族人群全基因组尺度上的适应性遗传变异图谱，鉴定了有相对明确功能的关键遗传变异，包括63个错义突变、7个失活性变异、1298个进化保守性变异，以及509个基因表达数量性状变异;这些分布在基因组范围的功能性变异不一定都与藏族人群的高原适应直接相关，但是大多数都与藏族人群的适应性演化密切相关。高原适应涉及到一系列复杂性状——涉及到的基因可能比医学中研究的一些复杂疾病更为错综繁杂。研究团队进一步发展了一个新统计量(FIS)对鉴定出的适应性遗传变异的相对重要性进行加权排序，发现排在首位的并不是通常认为的EPAS1，而是位于EPAS1下游的一个跨膜蛋白编码基因TMEM247;尤其是发现藏族人群基因组中的TMEM247存在一个高频关键错义变异(rs116983452)，可能对藏族人群高原适应具有重要贡献和意义。论文新发现的TMEM247基因关键突变(rs116983452)导致平原人群中高频存在的丙氨酸(Ala)(野生型)与青藏高原人群特有的缬氨酸(Val)(突变型)之间的显著分化，其中94%的藏族人都携带突变型，而在世界其他现代人群体中的频率非常低或者完全缺失，是迄今为止在青藏高原人群基因组中发现的最高频的错义突变。有趣的是，在西伯利亚丹尼索瓦洞穴中发现的距今约5万年的一个古人基因组也携带了这个变异，并且为纯合状态。该研究通过计算推断藏族人群中携带TMEM247-rs116983452适应性变异的序列可追溯至距今约6万年前，这意味着这个藏族特异的高频突变可能继承自早期进入高原的具有古人类血统的祖先并传承至今。

事实上，人类征服青藏高原的历程悠久而曲折。徐书华团队之前的研究表明，青藏高原人群的遗传起源可追溯至距今4-6万年前的旧石器时代中晚期，早期进入青藏高原的人类族群间发生广泛的基因交流，并与后期进入青藏高原的族群发生进一步遗传混合，最终形成了一个包含现代智人和早期智人多个谱系(包括考古学已经发现的阿尔泰尼安德特人和丹尼索瓦人等，以及其他未知古人类)的遗传构成极其复杂的混合人群。这个过程中，一些曾经帮助人类适应高原环境的古人类基因片段得以保留下来，因高原极端环境的自然选择作用，在现今高原人群中积累到较高的频率。TMEM247-rs116983452-T就是一个典型的例子。

分析表明，TMEM247-rs116983452-T的频率与人群居住地海拔呈显著正相关，提示与人类在青藏高原的适应可能有密切关系。进一步结合基因表达与多项生理生化表型及体质人类学特征，对高原藏族人群的适应性遗传变异进行了系统性评估;发现TMEM247-rs116983452-T与TMEM247及EPAS1的表达水平都有密切相关，并可能对藏族人群低氧环境下的血红蛋白和红细胞水平等高原适应性性状产生重要的调控。对于平原人群而言，长期暴露于低氧环境中将诱发红细胞增生以提高血液携氧能力，但最终可能过度增生而导致“ 红细胞增多症 ”。相比之下，高原世居藏族人群的红细胞和血红蛋白水平保持在相对较低的水平，TMEM247基因的功能突变可能就是产生这种保护性机制的重要遗传因素之一。通过统计模型分析，研究团队发现TMEM247-rs116983452对藏族高原适应性表型的解释度高于EPAS1的变异位点，但二者之间可能存在一定的相互作用，体现了高原适应的复杂性和多基因相互作用效应。该研究提供的这张基因组适应性变异图谱为后续进一步全面深入研究藏族适应高原的遗传基础和分子机制锁定了目标，为揭开人类征服高原极端环境的演化之谜开拓了新的视野。

该工作由营养与健康所、马普计算生物学所邓恋、张超、苑锴，以及博士研究生高扬(上海科技大学)、潘雨闻等在徐书华的指导下，与中科院昆明动物所、温州医科大学、复旦大学、西藏民族大学等多家单位的研究人员合作完成，得到中科院先导专项、国家自然科学基金委、上海市科委和国家重点研发计划等多项基金的资助。



Trait	Genotypes of rs116983452			adjusted *p*	adjusted *p* in validation analysis
	CC	CT	TT		
UA	344.28 ± 93.12	334.35 ± 91.59	327.91 ± 86.63	0.031	NA
RBC	5.06 ± 0.76	4.93 ± 0.71	4.82 ± 0.67	5.72×10^{-7}	0.181
HGB	163.57 ± 29.32	157.73 ± 28.22	155.52 ± 26.35	5.78×10^{-5}	4.87×10^{-3}
HCT	47.27 ± 7.98	45.54 ± 7.63	44.84 ± 7.12	2.57×10^{-6}	4.87×10^{-3}

图：藏族人群适应性变异TMEM247-rs116983452-T的群体遗传解析和功能统计学分析

更多 科学进展 请访问 https://www.iikx.com/news/progress/

本文版权归原作者所有，请勿用于商业用途，爱科学iikx.com转发