



新闻动态

当前位置 > 首页 > 新闻动态 > 科研动态

- 综合新闻
- 头条新闻
- 科技前沿
- 科研动态
- 媒体关注
- 图片新闻
- 通知公告
- 图片展示
- 视频

成都生物所破译中华蟾蜍大型基因组揭示其“成功”的秘诀

发表日期: 2021-01-19

作者: 吕彬

文章来源:



打印 文本大小 大 中 小

中华蟾蜍 (*Bufo gargarizans*) 隶属蟾蜍科蟾蜍属, 俗名“癞蛤蟆”, 是我国最常见也最为人所熟知的两栖动物。蟾蜍在我国的传统文化中占有重要的地位, 在神话传说中, 蟾蜍是嫦娥的化身, 和玉兔一起居住在月亮上, 因此月亮也被历代文人称为“蟾宫”, 蟾蜍被称为“月精”。蟾蜍自古以来都是美好意愿的象征。中秋佳节, 各家各户焚香拜月, 祈愿蟾宫折桂, 平步青云; 家庭美满, 圆如皓月。蟾蜍也是招财进宝的吉祥物被供奉于家家户户。此外, 蟾蜍有较高的医药研究价值, 其耳后腺和背部皮肤的分泌物有较强的毒性, 成分复杂, 主要是强心甙类化合物和生物碱, 统称为蟾蜍毒素 (Bufotoxin), 有麻醉止痛的功效。

蟾蜍科 (Bufonidae) 物种的分布非常广泛, 在除了澳大利亚、马达加斯加与南极洲以外的各大洲都有分布, 可以说是最成功的两栖动物类群。甘蔗蟾蜍 (*Rhinella marina*) 在被人为引入澳洲后迅速入侵土著物种的生境, 泛滥成灾, 造成了严重的生态系统破坏。中华蟾蜍在我国分布十分广泛, 从岛屿到高原都有其足迹, 其食性广泛, 能适应各种不同的生境, 强烈的毒性也让许多潜在的捕食者 (蛇类等) 望而却步。蟾蜍广布适应性的遗传机制目前还不清楚, 通过基因组测序和组装可以帮助我们更好的了解这一重要的物种。相较其它动物, 两栖动物的基因组破译是非常困难的, 这主要是由于其基因组普遍较大, 无尾两栖类 (anurans) 基因组大小多数范围在2到8Gb之间, 蚓螈 (caecilians) 多数达到~5-10 Gb, 蝾螈 (salamanders) 更是多数高达~20-60 Gb, 基因组最大的是一种洞螈 (*Necturus lewisi*), 为137Gb。另外, 重复序列占比较高也是造成测序和组装困难的主要原因之一。因此, 目前两栖动物的基因组数量较少, 而且多数组装呈高度片段化, 质量不佳。近几年发展成熟的长片段测序技术能够跨过较长的重复区域, 结合基因组辅助组装技术有望得到高质量的两栖类基因组。

为了揭秘中华蟾蜍的适应性遗传机制, 中国科学院成都生物研究所吕彬, 江建平和傅金钟主导, 联合国内多所高校科研人员, 包括华中师大吴华, 河南师大陈晓红, 信阳师大宋晓威和西华师大廖文波共同发起了中华蟾蜍基因组测序项目, 成功破译中华蟾蜍基因组, 获得首个组装到染色体水平的蟾蜍基因组参考图谱。

研究利用了4种最新的测序组装技术来构建参考基因组, 包括PacBio长读长测序 (103X), BioNano光学图谱 (137X), 10x genomeics (130X) 和Hi-C染色体构象捕获技术 (176X), 总计产生了近3Tb的测序数据量。最终基因组组装大小约4.55Gb, 仅由747条scaffolds构成 (6,822条for热带爪蟾 *Xenopus tropicalis*), Scaffold N50达到539.8 Mb, gap占比仅为1.79%, 96%的序列可以定位到11条染色体上 (图1A和图2A)。比较发现中华蟾蜍与热带爪蟾的基因组染色体之间具有良好的共线性, 推测至少经历过1次染色体融合和3次分裂事件 (图3A)。中华蟾蜍基因组是目前两栖动物中连续性最好的参考基因组 (图1B-D), 为两栖动物的基因进化和功能以及遗传学研究提供了重要的理论支撑。

蟾蜍基因组中重复序列占比为65.5%, 大部分为长末端重复 (LTR) 序列, 这与墨西哥钝口螈 (*Ambystoma mexicanum*) 中的发现较为近似, 其基因组大小约为32Gb。推测LTR反转录转座子的转座复制

产生的大量拷贝可能是两栖类基因组巨大的重要原因。基因结构注释和比较发现越大的基因组拥有较长的基因间区和内含子，外显子长度在不同物种间较为保守(图2B)。

研究发现中华蟾蜍中具有超多的多拷贝基因，在斑马鱼，甘蔗蟾蜍和墨西哥钝口螈中也存在类似的情况。我们知道斑马鱼经历了硬骨鱼全基因组加倍事件，基因复制数目在本研究物种中确实是最多的；而墨西哥钝口螈基因组是现今完成测序的最大的动物基因组，也发生了较多的基因复制事件；甘蔗蟾蜍和中华蟾蜍隶属蟾蜍科，都是“真蟾蜍”，大多数复制都是局部发生。中华蟾蜍中的基因复制方式以串联(TD)和近端复制(PD)为主，并且TD和PD受到的正选择压力显著高于其他复制方式(图3B和C)，推测这两种复制方式可能与其环境适应性有关。

根据以上推断，我们进而分析了中华蟾蜍基因组中基因家族的扩张状况，发现有122个基因家族发生了显著的扩张(图4)，化学感受基因家族显著富集，包括嗅觉和苦味受体基因。相比而言，中华蟾蜍的中性替换速率并不高，处于中间水平，但是其祖先却经历了快速的进化过程(图4)。

和其他脊椎动物类似，两栖动物同样拥有数目庞大的嗅觉基因(800~1400 *ORs* this study)，有功能的嗅觉基因占比60%~70%，众多的基因数目说明嗅觉对两栖动物的生存是比较重要的。中华蟾蜍中共鉴定出1013个嗅觉基因(图5A和B)，765个是完整可能有功能的，占比较高(~76%)。从功能和变化趋势来说，蟾蜍感受空气中化学分子的嗅觉基因增多且功能基因比例升高，水中则相反。信息素受体在种内和种间通讯中扮演着重要的角色，研究在蟾蜍中鉴定到5个完整的信息素受体1(*VIRs*)但没发现假基因，4个完整的信息素受体2(*V2Rs*)和33个假基因(图5C)。*VIRs*负责检测空气中挥发的小分子，*V2Rs*检测水中的。多物种比较倾向支持从水到陆驱动的动态进化假说。

苦味受体能帮助动物感知食物是否有毒。苦味受体(*TAS2Rs*)在“真蟾蜍”中发生过3次急剧的扩张(图5D)，形成了130多个功能基因，分布在1, 4, 6, 8和11号染色体上。与其他无尾目物种相比，其家族成员数目是最多的，完整基因的比例也是最高的(~80%)。蟾蜍分布广且食性杂，大量的苦味受体有助于提高毒素检测的能力。无独有偶，先前关于美国白蛾入侵机制的研究(Wu et. al 2019)同样发现了苦味受体的显著扩张。

蟾蜍利用毒素进行防御，但其如何避免自中毒是个有意思的话题。我们发现细胞色素P450基因家族中起代谢解毒作用的*CYP2C*亚家族在“真蟾蜍”中发生了显著扩张(图6A)，中华蟾蜍中产生了39个基因，甘蔗蟾蜍中有35个，而其他无尾类都只有约一半的成员数。聚类分析发现了蟾蜍科特异性的*CYP2C*扩张(图6A)，分布在中华蟾蜍5号染色体上一段约3Mb的区域内(图6B)，通过PD(近端)和TD(串联复制)方式形成。这些扩张的家族成员应是在祖先物种中就已经形成了。

离子通道基因在防止自中毒过程中发挥着重要的作用。举一个著名的例子，箭毒蛙的烟碱乙酰胆碱受体只发生了1到2个位置的氨基酸替代就进化出了毒素耐受性(2017 Tarvin et. al)。我们发现一种烟碱乙酰胆碱受体*CHRNAE*在中华蟾蜍，甘蔗蟾蜍和草莓箭毒蛙基因组中都产生了2个拷贝(图7B)，通过串联复制产生，可能和抗生物碱毒性有关。此外，研究发现在“真蟾蜍”的钠钾离子泵基因的关键毒素结合区存在很多独有的氨基酸突变(图7A)，先前实验证明有的位点突变能显著鉴定毒素结合敏感性，降低程度甚至能达到1000倍以上。有意思的是，这些关键区域的氨基酸位点在中华蟾蜍和甘蔗蟾蜍中是完全一致的，说明受到较强的功能限制。进化分析表明，相对其他物种，“真蟾蜍”钠钾离子泵基因的经历了快速的进化并受到正选择压力。

此外，研究发现6个与牙齿(牙釉质和牙本质)形成相关的基因在“真蟾蜍”祖先中就已经丢失(图8)。这和无齿的龟，鸟类，须鲸一致，但都是独立丢失的，说明四足动物牙齿形成机制是比较保守的。

该工作以“A large genome with chromosome-scale assembly sheds light on the evolutionary success of a true toad (*Bufo gargarizans*)”为题发表在Molecular Ecology Resources上。研究得到国家自然科学基金(31729003, 31600104, 31572245, 31872220)，四川省科技厅应用基础重点项目(18YYJC0171)和生态环境部生物多样性调查监测与评价项目(2019 - 2023)等支持。论文链接：<https://doi.org/10.1111/1755-0998.13319>

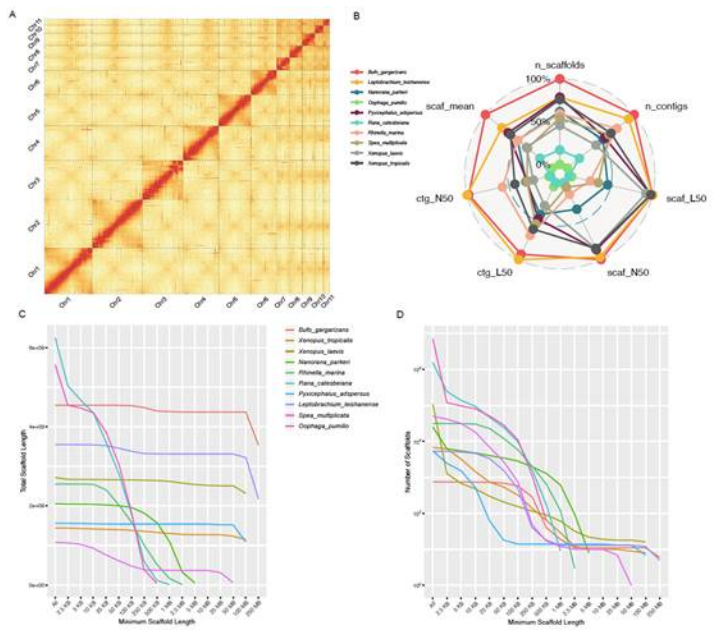


图1. (A) Hi-C热图; (B-D) 两栖无尾类基因组组装连续性比较。

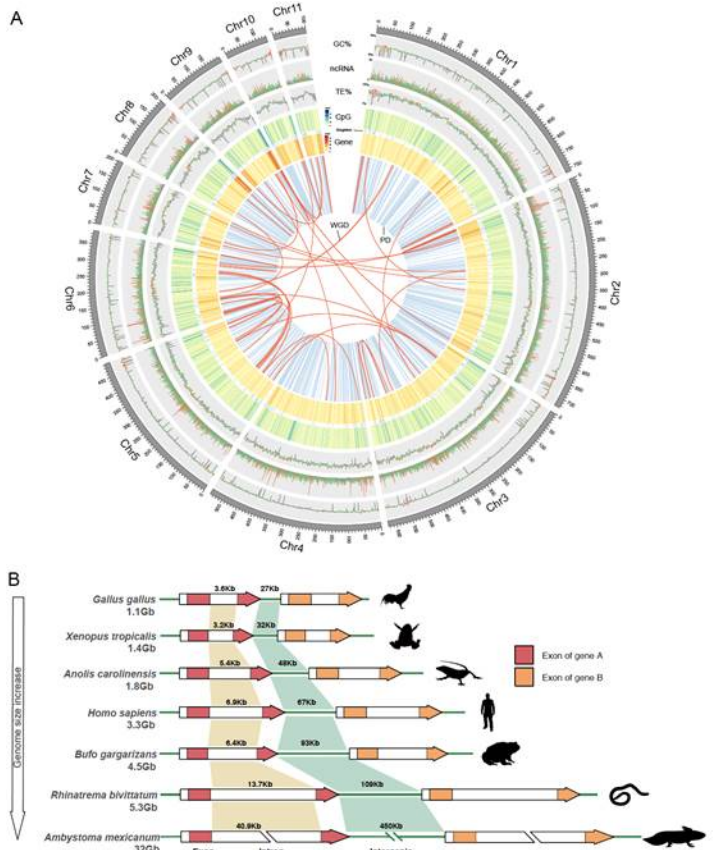


图2. (A)中华蟾蜍基因组全景; (B) 不同动物类群基因组的基因区及基因间区长度示意。

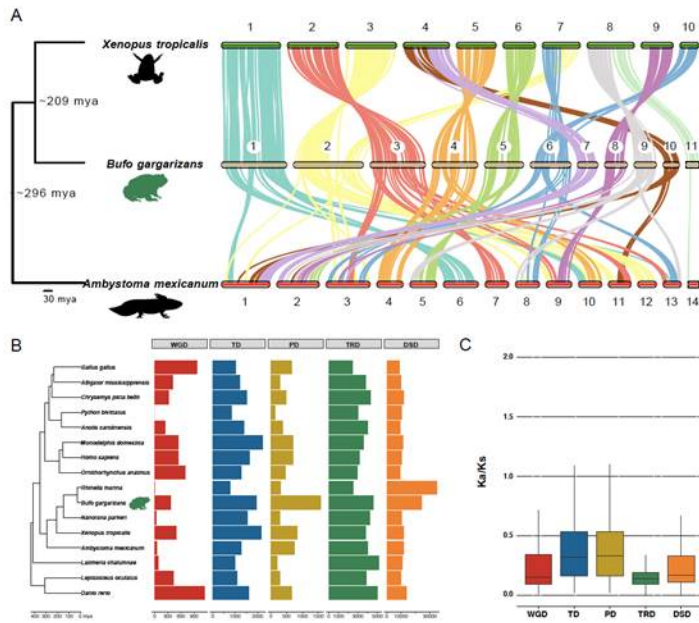


图3. (A)染色体共线性; (B)5种复制模式统计; (C)中华蟾蜍不同复制模式的选择压力比较。

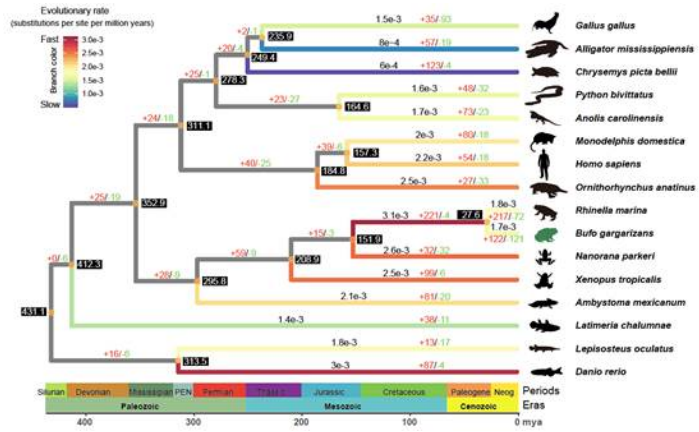


图4. 基因家族进化分析。枝颜色代表进化速率，黑框内数字为分化时间，+扩张，-收缩。

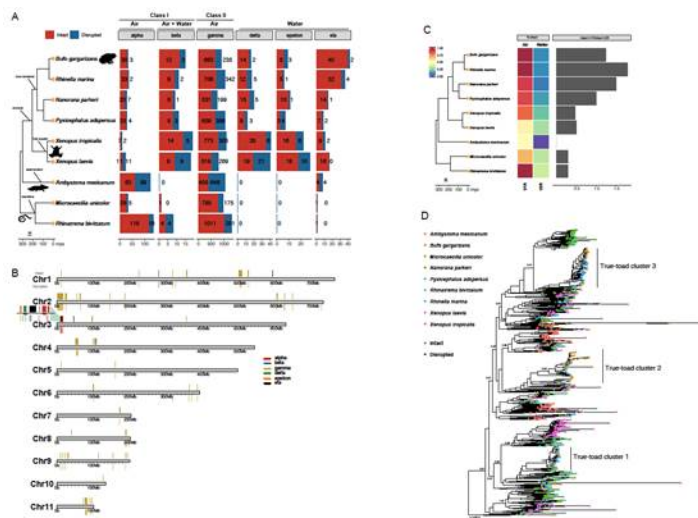


图5. 化学感受基因家族进化。(A)嗅觉受体类型和数量; (B)嗅觉受体不同类型在蟾蜍染色体上的分布; (C)信息素受体的类型; (D)苦味受体的扩张。

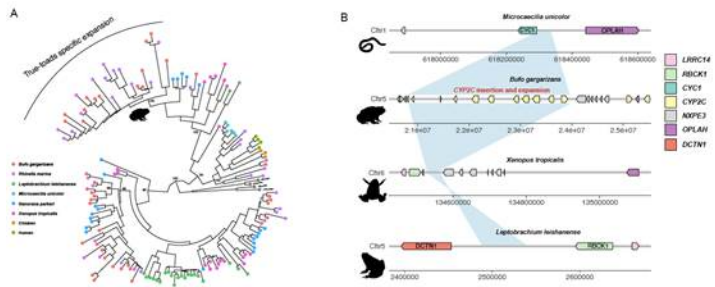


图6. P450 CYP2C基因家族在真蟾蜍中发生支系特异性的扩张(A)和分布(B)。

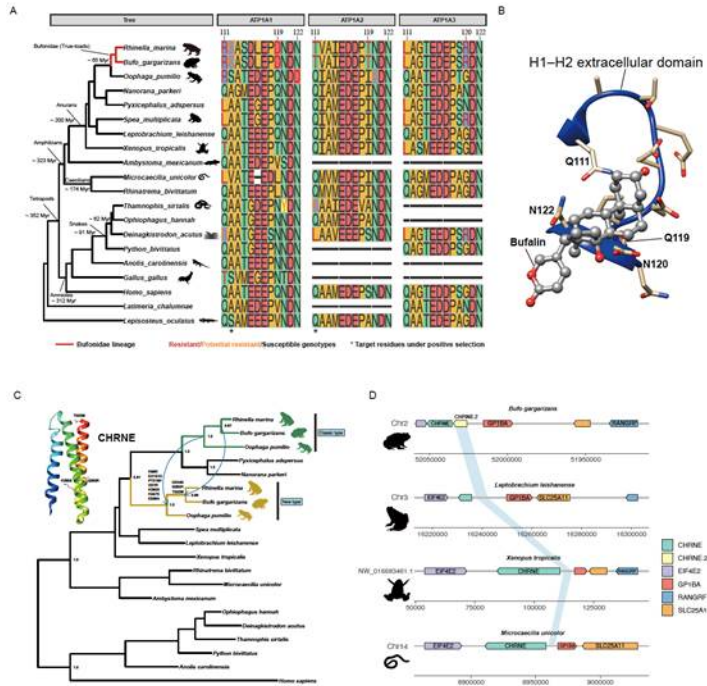


图7. 离子通道基因进化与毒素耐受性。(A-B)钠钾离子泵；(C-D)烟碱乙酰胆碱受体。

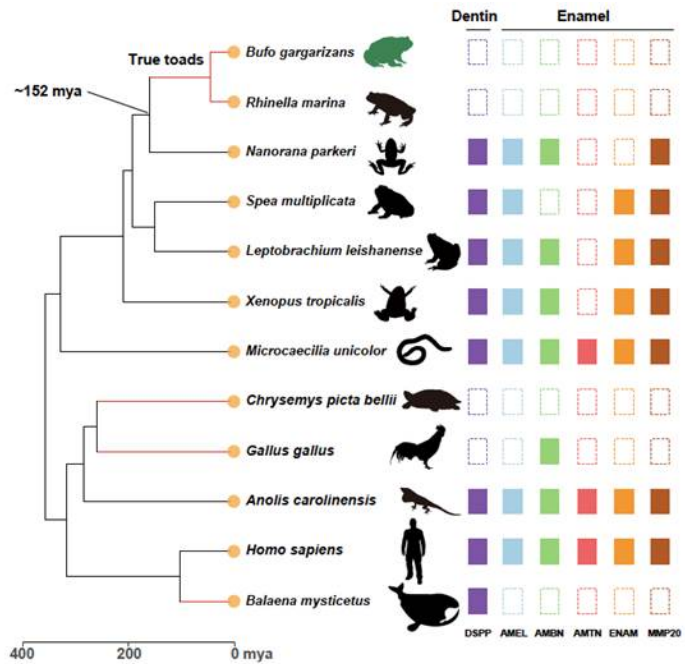


图8. 牙齿发育基因的多次独立丢失。



邮政编码: 610041 地址: 中国四川省成都市人民南路四段九号
中国科学院成都生物研究所 版权所有
蜀ICP备05005370号-1