基因重复在基因选配(gene co-option)过程中的重要作用: 对 motilin/ghrelin 荷尔蒙基因家族及其受体基因家族的 进化研究

何静^①, Irwin M. David^{①②}, 张亚平^{①③*}

① 中国科学院昆明动物研究所遗传资源与进化国家重点实验室, 昆明 650223;

② Departments of Laboratory Medicine and Pathobiology and the Banting and Best Diabetes Centre, University of Toronto, Toronto, Ontario, Canada;

③ 云南大学生物资源保护与利用实验室, 昆明 650091

* 联系人, E-mail: zhangyp@mail.kiz.ac.cn

2010-12-20 收稿, 2011-05-16 接受

国家重点基础研究发展计划(2007CB411600)、国家自然科学基金(30621092, 30623007)和云南省科技厅 (O803481101)资助

摘要 原先存在的基因经过一些改良可能获得新的属性从而承担新的功能. 在分子遗传的层次, 这个过程通常伴随基因重复以及基因重复后旁系同源基因的功能分化. 本研究探索了基因重复后荷尔蒙、受体间特异性相互作用的演化形成. 尽管对这个问题已有相关研究报道, 但针对更多个案的研究能帮助我们更好地理解此前已经提出的进化模型的普遍性, 可能还有助于发现新的进化模式. 基于生物信息学的手段, 结合比较基因组学、系统发育学的方法, 本研究揭示, ghrelin 前体基因和 motilin 前体基因是由一个祖先基因重复而来, 基因重复发生在羊膜动物与两栖动物刚刚分歧之后. 与此形成鲜明对照, 它们各自的特异性受体却呈现了很不一样的进化历史. GPR39 最先分化出来, 而后一个祖先基因经连续的基因重复分化为鱼类特异的进化支 A, GHSR 和 MLNR, 基因重复发生在射线鳍鱼(ray-finned fish)与四足动物分化之前. ghrelin/GHSR 信号系统的功能从鱼类到哺乳动物的进化过程中始终保守. motilin-MLNR 之间的特异性相互作用是荷尔蒙基因重复后, MLN 与 GHRL 基因分化之后, 在羊膜动物世系中配体与受体协同进化而形成的. 该结果提示, 在分子水平上, 基于保守的分子元件, 新的神经内分泌响应模式能够通过基因重复后的基因选配过程形成. 基因重复既节俭又极具创造力, 为生物多样性的演化提供了物质基础.

关键词

ghrelin
motilin
基因家族
配体
受体
协同进化
基因间相互作用

“进化是一个在原有基础上修修补补的过程, 这一过程是自然选择将就手头的材料而进行的”, 尽管这一基本观念秉承自达尔文的思想^[1], 然而真正让这一观念在现代进化生物学领域变得显著还得归功于诺贝尔奖得主 Jacob 一篇里程碑式文章的发表^[2]. 在分子水平, 进化过程的修修补补就更显而易见了. 随着大规模基因测序技术的发展, 已经有许多物种的

基因组被测定. 通过对这些数据的比对、分析发现, 尽管不同的多细胞生命形态上表现出非同寻常的多样性, 在分子水平它们却共享大量的基因. 通过改变基因的表达调控模式或和基因所编码蛋白质的功能, 已经存在的基因能通过基因选配(gene co-option)过程生成新的功能元件, 而这一过程通常伴随基因重复及其后重复拷贝的分化^[3]. 尽管基因选配的重要性

英文版见: He J, Irwin M D, Zhang Y P. Gene duplication plays a major role in gene co-option: Studies into the evolution of the motilin/ghrelin family and their receptors. Chinese Sci Bull, 2011, 56, doi: 10.1007/s11434-011-4614-9

已成为进化生物学的热点问题,蛋白质功能的多效性也受到了很大的关注^[4,5],但是基因重复后基因选配的发生以及新功能的演化,功能的特异性与多效性如何协调共存都还需要深入探讨。

motilin 和 ghrelin 共同组成了 motilin/ghrelin 荷尔蒙家族,二者都参与调控胃肠消化道的功能^[6]。ghrelin 是其前体基因 *preproghrelin*(*GHRL*)翻译后剪辑而来,主要由胃部分泌,经循环系统流通,传入下丘脑及后脑发挥作用,通过与其特异性受体 *GHSR* (growth hormone secretagogue receptor, 一种 G 蛋白偶联受体)结合刺激垂体生长激素(growth hormone, GH)的释放^[6]。越来越多的证据表明,除了刺激生长激素的释放,ghrelin 还能增加胃肠蠕动、增大食欲,在能量代谢平衡中促进能量的积累、增加体重,此外,ghrelin 还参与多种其他的功能^[7]。前体基因 *prepromotilin*(*MLN*)翻译后剪辑产生荷尔蒙多肽 motilin^[8]。motilin 能通过激活神经转导或直接作用于平滑肌增加胃肠消化道的蠕动。以人类或狗为模型的生理实验表明,motilin 主要在禁食状态调控胃肠的蠕动^[8]。值得注意的是,motilin 还有微弱的刺激生长激素释放的作用^[9]。1999 年,一种未知 G 蛋白偶联受体 *GPR38* 被发现是 motilin 的特异性受体,此后它被命名为 *MLNR*^[10]。不仅两种荷尔蒙,ghrelin 和 motilin,表现出结构的高度相似;它们各自的特异性受体也高度同源,蛋白质序列同源相似性达到 44%,跨膜同源区的相似性更高达 87%。尽管两对配体、受体序列上高度相似,但没有证据显示相互间存在交叉反应的可能,与之对应两种荷尔蒙的药效基团大不相同^[6]。

ghrelin 和 motilin 荷尔蒙通过结合各自特异性受体参与调控相关但并不相同的生理过程。至今关于 motilin/ghrelin 荷尔蒙及其特异性受体基因家族的进化历史没有深入研究,为了更好地理解基因重复后基因产物参与分子间特异性相互作用的动态变化规律,以及功能的分化,探索配体、受体间是否存在进化上的协同性,我们重建了荷尔蒙及其受体基因家族的进化历史,进行了深入的分子进化比较分析。本研究对理解新基因功能的早期进化具有重要意义。

1 材料与方法

从 GenBank 获取了 *GHRL*, *MLN*, *GHSR*, *MLNR* 和 *GPR39* 的已知序列。*GHSR*, *MLNR* 和 *GPR39* 是隶属于类 rhodopsin 受体基因家族 β 簇,彼此高度同源,

彼此间的相似度远高于它们与任何已经鉴定的受体之间的相似度^[11]。以上述已知序列为模板,用 *TBlastN*^[12] 搜索黑猩猩(*Pan troglodytes*) panTro2, 红猩猩(*Pongo pygmaeus abelii*) ponAbe2, 恒河猴(*Macaca mulatta*) rheMac2, 绒猴(*Callithrix jacchus*) calJac1, 小鼠(*Mus musculus*) mm9, 大鼠(*Rattus norvegicus*) rn4, 豚鼠(*Cavia porcellus*) cavPor3, 猫(*Felis catus*) felCat3, 狗(*Canis lupus familiaris*) canFam2, 马(*Equus caballus*) equCab2, 牛(*Bos taurus*) bosTau4, 负鼠(*Monodelphis domestica*) monDom4, 鸭嘴兽(*Ornithorhynchus anatinus*) ornAna1, 鸡(*Gallus gallus*) galGal3, 蜥蜴(*Anolis carolinensis*) anoCar1, 非洲爪蟾(*Xenopus tropicalis*) xenTro2, 斑马鱼(*Danio rerio*) danRer5, 潜水艇鱼(*Tetraodon nigroviridis*) tetNig1, 河豚(*Takifugu rubripes*) fr2, 棘鱼(*Gasterosteus aculeatus*) gasAcu1, 青鳉(*Oryzias latipes*) oryLat2 基因组,查找相关荷尔蒙及受体基因同源序列,然后用 Wise2 (version 2.1.20 stable)软件预测基因的编码区。这个软件通过 DNA 和氨基酸序列的比对,能准确预测同源基因序列的编码区^[13]。新预测的受体基因反过来 Blast 同源搜索已经详细注释的哺乳动物基因组,以排除之前方法找到的非 *GHSR*, *MLNR* 和 *GPR39*,但类似 rhodopsin 的受体基因。然后用 HMMTOP^[14](一个能够准确预测蛋白质跨膜螺旋和拓扑结构的软件)检查每一个新预测的受体基因具备 G 蛋白偶联受体特异的 7 次跨膜结构。所有预测的荷尔蒙及受体基因具有保守的内含子-外显子边界(如 AG/GT)和相同的内含子相位。本研究涉及的序列在补充材料中。

基于普遍接受的物种间系统发育关系^[15,16]和物种间分歧时间,我们估算了 *GHSR* 和 *MLNR* 两种受体在脊椎动物进化过程中的大致进化速率。我们采用的物种间分歧时间如下:射线鳍鱼与四足动物,450 MYA (百万年);两栖动物与羊膜动物,360 MYA;鸟类与哺乳动物,310 MYA^[17];单孔目与真兽亚纲,225 MYA;有袋目与真兽亚纲,190 MYA;劳亚兽总目(Laurasiatheria)与灵长总目(Euarchontoglires),94 MYA;灵长目与啮齿目,87 MYA;灵长目与兔形目,87 MYA;猪与偶蹄类反刍动物,65 MYA;牛和山羊/绵羊,20 MYA;大鼠与小鼠,23 MYA;旧大陆猴与人猿超科,36 MYA;食蟹猴与恒河猴,1.92 MYA;人类与黑猩猩,5.5 MYA^[16-19]。

用 ClustalW 进行基于氨基酸序列比对的 DNA 编

码序列的比对^[20]. 荷尔蒙及受体基因家族的系统发育关系采用 PHYLIP 3.65 软件包构建^[21]. 构建系统关系时序列比对结果中存在插入/缺失的(gap)位点不予考虑. 序列比对结果经 SEQBOOT 生成数据矩阵, 然后分别用最大简约法、邻接法、最大似然法 3 种算法构建系统发育关系, 最后采用 CONSENSE 合成一致树. 基于普遍接受的物种间系统发育关系, 用 PAML^[22]中 codeml 软件估算了 *GHSR*, *MLNR* 两种受体的氨基酸进化距离, 并同时预测了可能受到正选择作用的氨基酸位点.

2VT4A^[23], 2J4YA^[24], 2Z73A^[25]与受体基因家族同源性最高, 三者的 *E* 值分别为 3×10^{-15} , 9×10^{-14} 和 7×10^{-7} . 检索 SWISS-PROT (<http://www.expasy.org/sprot/>)数据库获得 *MLNR*, 2VT4, 2J4Y 和 2Z73 的二维结构信息. 以 2VT4A^[23], 2J4YA^[24], 2Z73A^[25]为模板, 以二维结构为参考, 基于全局序列比对, 用 MODELLER (<http://www.salilab.org/modeller/>)进行同源模建预测 *MLNR* 的三维结构, 并对环区(loop)用 Accelrys Discovery Studio 2.1 (<http://accelrys.com/products/discovery-studio/>)进行优化. 用 Rasmol^[26]进行三维结构的可视化.

2 结果

2.1 脊椎动物的 *GHRL*, *MLN* 以及类似 *GHSR* 的受体基因

与此前研究过的生长激素、胰岛素、胰高血糖素不同^[27], 本研究在鱼类基因组中只发现存在 1 个拷贝的 *GHRL/MLN* 类似基因(表 S1 和 S2), 推测鱼类特异的全基因组重复后, 只有一个拷贝被保留了下来^[28]. 之前有些物种的 *GHSR* 类似基因被克隆, 但除了来自两个鱼类物种的信息外(*Acanthopagrus schlegelii* 和 *Sphoeroides nephelus*)^[29,30], 大都集中于哺乳动物类群. 应用生物信息的方法我们获得了多种脊椎动物 *GHSR* 类似受体的信息(表 S3). 不同物种生物信息的比对研究还显示, *GHRL*, *MLN*, *GHSR*, *MLNR* 和 *GPR39* 分别位于不同的基因座位, 不同物种间相关基因的上下游十分保守, 表明不同物种的对应基因是直系同源的关系. 如果比较 *GHRL* 和 *MLN* 上下游临近基因, 不存在显著相似性. 当对比人类与其他脊椎动物基因组时(UCSC, <http://genome.ucsc.edu/>), 发现与人类 *GHRL* 和 *MLN* 及其上下游临近基因同源

的区段, 在羊膜动物中分别位于不同染色体的不同基因座位; 而在与非洲爪蟾、青鳉、潜水艇鱼和斑马鱼的比对中, 却发现与人类 *GHRL* 和 *MLN* 及其上下游临近基因同源的区段, 位于同一染色体上 *GHRL* 的基因座位处(图 S1). *GHRL*, *MLN* 上下游临近区段分布不同的基因, 没有重合的同源基因. 以上结果提示, 鱼类与两栖动物的共同祖先只有一个拷贝的 *GHRL/MLN* 基因. 然后经过羊膜动物世系特异的基因组大片的重复, 产生了包含 *GHRL/MLN* 基因的 2 个区段. 其后 2 个区段中相同基因发生任一拷贝的随机丢失, 只有 *GHRL/MLN* 基因两个拷贝都保留了下来, 分别分化为 *GHRL* 和 *MLN*. 因此我们只观察到 *GHRL* 和 *MLN* 的同源相似性, 但二者的上下游却没有显著同源性. 我们没有观察到受体基因座位上下游的显著同源相似性, 也就是说基于基因组数据的比较分析没有提示全基因组的重复产生了相关受体基因, 但是也不能排除这种可能.

2.2 motilin/ghrelin 荷尔蒙基因家族的进化历史

尽管 *GHRL* 和 *MLN* 已经彼此分化巨大, 但还是保留了一些特征, 有助于解析它们的进化历史. 我们获得了来自脊椎动物不同门类(包括鱼类、两栖类、爬行类、鸟类和哺乳类)不同物种的 *GHRL* 基因. 但只从爬行类、鸟类和哺乳动物物种获得了 *MLN* 基因. 荷尔蒙 ghrelin 的氨基酸序列在进化过程中十分保守, 尤其是 N 端高度保守, 荷尔蒙 motilin 也遵从相同的规律. *GHRL* 和 *MLN* 在基因的整体结构、序列方面都十分近似. 二者的可读框都由 4 个外显子编码, 具有相同的外显子/内含子组织结构和相位. ghrelin 和 motilin 两种荷尔蒙多肽的编码序列跨越了第一、第二外显子, 内含子都是在第 14 个氨基酸将荷尔蒙多肽打断, 并且两者共享相同的翻译后剪辑多肽位点(图 1), 说明二者由一个祖先基因分化而来. 细致的分析研究发现, 射线鳍鱼和两栖类的 *GHRL* 基因只有一个翻译后剪辑位点, 位于信号肽剪辑位点的 C 端, 因此翻译后剪辑只能生成一种多肽荷尔蒙 ghrelin. 但羊膜动物(爬行类、鸟类和哺乳动物)的 *GHRL* 拥有 3 个翻译后剪辑位点, 位于信号肽剪辑位点的 C 端, 可能产生两种多肽荷尔蒙, ghrelin 和一种 24 个氨基酸长的 ghrelin 相关多肽(图 1), 后者在近期一个研究工作中被发现并命名为 obestatin^[31]. 所获得的 *MLN* 基因(来自不同的爬行类、鸟类和哺乳动

图1 *ghrelin* 和 *motilin* 基因结构图

以矩形表示 *ghrelin* 和 *motilin* 前体蛋白不同的结构域, 长度与氨基酸序列实际长度成比例。空心 and 实心三角分别指示信号肽剪辑位点和翻译后蛋白酶剪辑位点。蛋白结构域下方列表了不同脊椎动物门类蛋白酶剪辑位点的氨基酸序列。列表中鸟类和硬骨鱼列项中(R)表示在对应蛋白酶剪辑位点存在另外的可能性, 即单个精氨酸。“↑”标示内含子位置, 并将内含子相位显示在旁边。SP 表示信号肽。“/”表示没有蛋白酶剪辑作用位点

物)拥有 3 个翻译后剪辑位点, 位于信号肽剪辑位点的 C 端, 可能产生两种多肽荷尔蒙, *motilin* 和一种 17 个氨基酸长的 *ghrelin* 相关多肽(图 1)。

我们构建了 *GHRL/MLN* 的系统发育关系。基因的系统发育关系基本与物种的系统发育关系吻合, 除了牛蛙 *GHRL* 的系统位置显得突兀。如图 2, S2 及 S3 所示, 牛蛙 *GHRL* 与羊膜动物 *MLN*, 而不是羊膜动物 *GHRL* 聚在一起, 尽管支持率很低。此前的研究证明, 牛蛙 *GHRL* 具有所有 *GHRL* 的特征, 其 *ghrelin* 荷尔蒙多肽的第三位氨基酸为苏氨酸, 此位点翻译剪辑后还经过 8 个氨基酸酰基化修饰^[32]。基于已经发布的基因组数据只获得了部分非洲爪蟾的 *GHRL* 编码区序列(第 1 和第 3 外显子, 其序列与牛蛙 *GHRL* 十分相似, 整个基因组只发现有一个 *GHRL/MLN* 类似的拷贝), 所以非洲爪蟾的 *GHRL* 序列没能用于系统发育关系的构建。依据 *GHRL* 和 *MLN* 基因在不同门类不同物种的存在情况, 以及 *GHRL/MLN* 基因翻译后剪辑位点的分布, 推测 *MLN* 基因是由羊膜动物世系特异的基因重复而产生。这也是最简约的情况, 并将系统发育树对应进行了的定根(图 2)。当然并不能完全排除其他可能性(如基因重复发生于鱼类与四足动物分化之前), 但那意味着基因重复后经历了更多的基因丢失事件, 并平

行获得了同样的翻译后剪辑位点。

2.3 *GHSR* 类似受体基因家族的进化历史

我们构建了来自不同物种 *GHSR* 类似受体基因家族的系统发育关系, 分别采用了最大似然和邻接法, 两种算法结果基本一致。 *GHSR* 类似受体基因家族成员进化上聚为四枝, 且每一枝以很高的支持率聚为单系, 其中 3 个进化枝有哺乳动物直系同源基因代表, 也就据此进行了命名(图 3), 由此定义了 3 种类型的受体, 每种类型的受体在我们所研究的每一物种中只有一个拷贝。第 4 个进化枝是鱼类特异的, 我们命名为进化枝 A。

如图 3 所示, 受体 GPR39 最先分化, 而后 *GHSR* 类似基因先后分化为 3 个单系: 进化枝 A, *GHSR* 和 *MLNR*。Palyha 等人^[30]以河豚(*Spheroides nephelus*)为模型, 克隆了 3 种不同的 *GHSR* 类似基因(SnAF 082209, SnAF082210, SnAF082211), 在我们的系统树上它们分别聚类为 *GHSR*, *MLNR* 和进化枝 A 基因。进化枝 A 只由来自鱼类的基因组成, 没有发现鱼类以外其他脊椎动物的直系同源拷贝, 目前也不了解它们的功能。进化枝 A 的基因有一些独有的特征, 如具有很长的第 2 和第 3 环区, 大体上比 *GHSR*

图2 *motilin/ghrelin* 荷尔蒙基因家族系统发育进化关系

用不同门类多种脊椎动物 *ghrelin* 和 *motilin* 前体蛋白序列构建的系统发育树, 基于最大似然算法 Jones-Taylor-Thornton 模型, 每一结点支持率显示在旁边。GHRL 即 *ghrelin* 前体蛋白, MLN 即 *motilin* 前体蛋白。MLN 以粗体显示

和 *MLNR* 的同源环区长 100 个氨基酸左右. 此外, 不同物种 *MLNR* 的第 2 环区也呈现很大变异, 哺乳动物的 *MLNR* 第 2 环区较之非哺乳动物及 *GHSR* 的同源区段长约 40 个氨基酸. 研究表明, 受体基因这些环区前后两端的氨基酸位点对于它与荷尔蒙的结合相互作用十分重要^[33], 但环区长出部分的功能意义目前还不清楚.

2.4 *GHSR* 类似基因的进化速率

以鱼类为外群, 基于普遍接受的物种间系统发育关系^[15,16]和物种分化时间, 结合 PAML 的 *codeml* 计算进化枝上的氨基酸 P 距离^[22], 我们估算了 *GHSR* 和 *MLNR* 在脊椎动物中的进化速率. *GHSR* 在脊椎动物中进化相对匀速, 速率大约在 $0.139 \times 10^{-9} \sim 0.71 \times 10^{-9}$ 氨基酸位点⁻¹ a⁻¹ 的区间内(图 S4(a)). 而 *MLNR*, 在羊膜动物分化之前的进化枝经历了插曲式的快速进化(图 S4(b))所示, 被标记为 a-b, 进化速率

高达 2.907×10^{-9} 氨基酸位点⁻¹ a⁻¹, 而其他进化枝呈现相对的匀速. 蛋白质氨基酸序列的快速进化可能是两种进化驱动力的结果, 达尔文正选择或者功能限制的放松. 我们采用 PAML^[22]不同模型来检验 *MLNR* 基因在脊椎动物进化过程中所经历的选择压力. 多种模型的检验都显示正选择作用的存在(表 1). 进一步采用枝-位点模型检测进化枝 a-b 上可能经历正选择作用的氨基酸位点. 结果显示, *MLNR* 在羊膜动物祖先进化的早期有 7.5% 的氨基酸位点经历了正选择作用下的快速进化. 我们将这些受到正选择作用的位点映射到同源模建的 *MLNR* 三维结构上, 以便进一步研究这一正选择作用的生物学意义. 之前基于定位突变探索多肽荷尔蒙与其特异性受体相互作用的研究显示, 胞外环区于受体、配体相互作用具有重要作用^[34]. 除此之外, 有一些多肽荷尔蒙与受体结合时, 受体跨膜区会形成一个结合口袋, 因而位于口袋内表面的一些氨基酸也参与受体、配体的

图3 多种脊椎动物64个 *GHSR* 类似受体的系统发育关系

用多种脊椎动物64个 *GHSR* 类似受体的氨基酸序列基于最大似然法构建的系统发育关系(我们同时也用邻接法进行了系统发育关系构建, 两种方法结果基本一致). 支持率显示在结点旁边. 人类的 NMUR1 和 NMUR2 作为外群来定根. Hs, *Homo sapiens*; Pt, *Pan troglodytes*; Rm, *Macaca mulatta*; Mf, *Macaca fascicularis*; Mm, *Mus musculus*; Rn, *Rattus norvegicus*; Oc, *Oryctolagus cuniculus*; Cf, *Canis familiaris*; Bt, *Bos taurus*; Oa, *Ovis aries*; Ss, *Sus scrofa*; Et, *Echinops telfairi*; Md, *Monodelphis domestica*; Ora, *Ornithorhynchus anatinus*; Gg, *Gallus gallus*; Ac, *Anolis carolinensis*; Xt, *Xenopus tropicalis*; Dr, *Danio rerio*; Tr, *Takifugu rubripes*; Tn, *Tetraodon nigroviridis*; Sn, *Spheroides nephelus*; As, *Acanthopagrus schlegelii*; Ga, *Gasterosteus aculeatus*; Ol, *Oryzias latipes*.

结合相互作用^[34]。如图 4 所示, 在我们的受体三维模型中, 受到正选择作用的氨基酸大多聚集在环区和配体结合的口袋内部, 提示正选择作用是由配体、受体相互作用驱动的。因为这一正选择作用下的快速进化, 发生在两栖动物分歧之后, 鸟类与哺乳动物分化之前的短时间内, 此后受体的进化又恢复匀速; 与此同时 *GHRL/MLN* 基因经历基因重复产生了 2 个拷贝, 推测这一 *MLNR* 的快速进化是 *MLNR* 与其新配体 motilin 相互作用、协同进化的结果。

3 讨论

我们重建了 motilin/ghrelin 荷尔蒙基因家族的进化历史。基于进化的简约原则, 我们推断祖先 *GHRL/MLN* 基因编码蛋白只有一个信号肽 C 端的翻译后剪辑位点; 两栖动物分歧后, 羊膜动物分化之前, 该基因演化出了另外两个信号肽 C 端的翻译后剪辑位点, 然后经基因重复产生了 *GHRL* 和 *MLN*。但受体基因家族的进化历史与之不同, 大约 450 MYA, 射线鳍鱼与四足动物分化之前, 受体基因经历先后多次基因重复。 *GPR39* 最先分化出来, 然后进化枝 A, 最后形成 *GHSR* 和 *MLNR* (图 3)。进化枝 A 的基因只在鱼类中被保留, 其他几种受体在辐鳍亚纲的鱼 (actinopterygians) 和四足动物中一直保守。

受体-配体间是特异性紧密相互作用的, 那么预料中受体-配体间应该彼此协同进化, 也就意味着受体、配体应该同时经基因重复产生^[35,36]。但本研究中, *MLNR* 的进化历史与其受体并不一致。受体的进化历史表明, 射线鳍鱼与四足动物分化之前, 大约 450MYA, *GHSR* 和 *MLNR* 已经过基因重复产生。 ghrelin 在从射线鳍鱼到哺乳动物的进化过程中一直保守, 主要参与下丘脑生长激素的释放^[37-39]。与之

对应, ghrelin-*GHSR* 系统进化上十分保守, *GHSR* 从射线鳍鱼到哺乳动物的进化过程中一直匀速而缓慢(图 S4(a))。 motilin 远比它的受体 *MLNR* 晚出现, 两栖动物分歧之后, 羊膜动物分化之前才经基因重复而来。自从起源伊始, *MLNR* 就受到持续的功能限制(图 3), 即使在其特异性配体演化出现之前都没有丝毫的功能限制放松的迹象。那么 motilin-*MLNR* 特异性相互作用是如何演化的? 由此生成了不同于 ghrelin-*GHSR* 另外的神经内分泌信号通路。

生物进化的过程中, 先前存在的基因经过一番改良能被用来承担新的功能, 这种改良是在祖先基因已有的结构和功能属性基础上进行的, 依赖于发生改良之前所具备的遗传物质资源和进化条件^[2,4,5]。这方面的研究近来有了新的进展。以另外的荷尔蒙-受体为模型, 研究发现, 如果受体结合配体的属性不明确, 具有与多种不同类型配体结合的潜能, 而新配体的出现暗合这种潜能(新配体是受体能够接受的类型), 随即新的分子间相互作用形成^[40]。 Ohno^[41] 的模型认为, 基因重复产生冗余拷贝, 因此有了选择压力的放松, 而这正是新功能演化的第一步。 Ohno 的模型基于在任何层次进化都遵从消极权衡(negative trade-off)的假设, 也就是说新功能的演化不得不以牺牲原先的功能为代价^[42]。越来越多的证据表明, 基因变得多功能化, 并不一定要以有损初始功能为折中^[4,42]。本研究中, 每种受体以很高的支持率独成单系, 提示受体之间已经发生了很大分化(图 3)。有研究表明, 鱼类的 *GHSR* 基因能被生长激素促分泌素(growth hormone secretagogues, GHSs)激活, 但 GHSs 和哺乳动物的 motilin 都不能激发鱼类的 *MLNR*^[30]。综上所述, *GHSR* 和 *MLNR* 基因重复后就发生了功能分化, 例如 *MLNR* 经历了亚、或者新功能化, 获得了

图 4 *MLNR* 的三维结构模型

用同源建模的方法构建了 *MLNR* 的三维结构。 *MLNR* 以带状模式显示, 不同结构域以颜色区分。我们检测到在羊膜动物祖先世系受到正选择作用的位点, 以红色、球棒模式突出显示在 *MLNR* 的三维结构上。(a)-(d)为每旋转 90°的侧面图; (e)为从胞外区看的俯视图

与 *GHSR* 不同的时空差异表达。当然探究 *motilin* 演化出现之前 *MLNR* 的配体及其所受到的功能限制是困难的, 但依据目前的信息我们可以肯定那时的 *MLNR* 绝不会具备了与 *motilin* 结合的潜能, 因为不论 *GHSR* 还是鱼类 *MLNR* 都不具备 *motilin* 激动的可能^[30]。那么 *motilin* 与 *MLNR* 间的特异性相互作用, 只能是在 *motilin* 演化出现之后才形成的。与之对应, 我们观察到 *MLNR* 配体结合环区、口袋部位的氨基酸与此同时经历了正选择作用下的快速进化。尽管 *ghrelin* 和 *motilin* 氨基酸序列很相似, 但它们与各自受体结合的模式是很不一样的, 也没有证据表明存在两对配体-受体交叉反应的可能^[6]。而且这两种多肽荷尔蒙的效用基因也不相同, *ghrelin* 具有非常独特的效用基因, 就是其第 3 位丝氨酸上的八氨基酰基

(牛蛙 *ghrelin* 中是第 3 位苏氨酸上的八氨基酰基), 研究证明失去这八氨基酰基修饰, *GHRL* 活性大大减损^[6,32]。也就是说荷尔蒙基因重复后, 新配体 *motilin* 功能的演化是以牺牲初始功能为代价的。

本研究有助于深入了解在分子水平基因重复后, 新功能演化的初期重复基因所经历的变化。基因不是孤立存在的, 而是复杂相互作用网络的组份。进化中网络各部分的功能不会始终静止不变, 网络系统能通过局部的改变而实现新的功能。基因功能的多样化为网络局部功能的优化、新功能的演化打下了基础。基因重复及其后冗余拷贝的功能分化在基因选配过程中意义重大。基因重复既节俭又极具创造力, 重复生成冗余拷贝, 在基因差异表达或新功能演化方面都增加了可塑性, 为生物多样性的演化提供了物质基础^[3]。

参考文献

- 1 Darwin C. On the Various Contrivances by Which British and Foreign Orchids Are Fertilized by Insects and on the Good Effects of Intercrossing. London: Murray, 1886
- 2 Jacob F. Evolution and tinkering. *Science*, 1977, 196: 1161–1166
- 3 True J R, Carroll S B. Gene co-option in physiological and morphological evolution. *Annu Rev Cell Dev Biol*, 2002, 18: 53–80
- 4 Khersonsky O, Roodveldt C, Tawfik D S. Enzyme promiscuity: Evolutionary and mechanistic aspects. *Curr Opin Chem Biol*, 2006, 10: 498–508
- 5 Tokuriki N, Tawfik D S. Protein dynamism and evolvability. *Science*, 2009, 324: 203–207
- 6 Peeters T L. Ghrelin: A new player in the control of gastrointestinal functions. *Gut*, 2005, 54: 1638–1649
- 7 Murray C D, Kamm M A, Bloom S R, et al. Ghrelin for the gastroenterologist: History and potential. *Gastroenterology*, 2003, 125: 1492–1502
- 8 Walsh J H, Dockaray G J. Gut Peptides: Biochemistry and Physiology. New York: Raven Press, 1993. 261–303
- 9 Samson W K, Lumpkin M D, Nilaver G, et al. Motilin: A novel growth hormone releasing agent. *Brain Res Bull*, 1984, 12: 57–62
- 10 Feighner S D, Tan C P, McKee K K, et al. Receptor for motilin identified in the human gastrointestinal system. *Science*, 1999, 284: 2184–2188
- 11 Holst B, Holliday N D, Bach A, et al. Common structural basis for constitutive activity of the ghrelin receptor family. *J Biol Chem*, 2004, 279: 53806–53817
- 12 Altschul S F, Madden T L, Schaffer A A, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res*, 1997, 25: 3389–3402
- 13 Birney E, Clamp M, Durbin R. GeneWise and genomewise. *Genome Res*, 2004, 14: 988–995
- 14 Tusnady G E, Simon I. The HMMTOP transmembrane topology prediction server. *Bioinformatics*, 2001, 17: 849–850
- 15 Murphy W J, Eizirik E, O'Brien S J, et al. Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science*, 2001, 294: 2348–2351
- 16 van Rheede T, Bastiaans T, Boone D N, et al. The platypus is in its place: Nuclear genes and indels confirm the sister group relation of monotremes and Therians. *Mol Biol Evol*, 2006, 23: 587–597
- 17 Hedges S B, Kumar S. Genomic clocks and evolutionary timescales. *Trends Genet*, 2003, 19: 200–206
- 18 Hayasaka K, Fujii K, Horai S. Molecular phylogeny of macaques: Implications of nucleotide sequences from an 896-base pair region of mitochondrial DNA. *Mol Biol Evol*, 1996, 13: 1044–1053
- 19 Springer M S, Murphy W J, Eizirik E, et al. Placental mammal diversification and the Cretaceous-Tertiary boundary. *Proc Natl Acad Sci USA*, 2003, 100: 1056–1061
- 20 Thompson J D, Higgins D G, Gibson T J. ClustalW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res*, 1994, 22: 4673–4680

- 21 Felsenstein J. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle. 1993
- 22 Yang Z. PAML: A program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci, 1997, 13: 555–556
- 23 Warne T, Serrano-Vega M J, Baker J G, et al. Structure of a beta1-adrenergic G-protein-coupled receptor. Nature, 2008, 454: 486–491
- 24 Standfuss J, Xie G, Edwards P C, et al. Crystal structure of a thermally stable rhodopsin mutant. J Mol Biol, 2007, 372: 1179–1188
- 25 Murakami M, Kouyama T. Crystal structure of squid rhodopsin. Nature, 2008, 453: 363–367
- 26 Sayle R A, Milner-White E J. RASMOL: Biomolecular graphics for all. Trends Biochem Sci, 1995, 20: 374
- 27 Roch G J, Wu S, Sherwood N M. Hormones and receptors in fish: Do duplicates matter? Gen Comp Endocrinol, 2009, 161: 3–12
- 28 Taylor J S, Braasch I, Frickey T, et al. Genome duplication, a trait shared by 22000 species of ray-finned fish. Genome Res, 2003, 13: 382–390
- 29 Chan C B, Cheng C H. Identification and functional characterization of two alternatively spliced growth hormone secretagogue receptor transcripts from the pituitary of black seabream *Acanthopagrus schlegelii*. Mol Cell Endocrinol, 2004, 214: 81–95
- 30 Palyha O C, Feighner S D, Tan C P, et al. Ligand activation domain of human orphan growth hormone (GH) secretagogue receptor (GHS-R) conserved from Pufferfish to humans. Mol Endocrinol, 2000, 14: 160–169
- 31 Zhang J V, Ren P G, Avsian-Kretchmer O, et al. Obestatin, a peptide encoded by the ghrelin gene, opposes ghrelin's effects on food intake. Science, 2005, 310: 996–999
- 32 Kaiya H, Kojima M, Hosoda H, et al. Bullfrog ghrelin is modified by n-octanoic acid at its third threonine residue. J Biol Chem, 2001, 276: 40441–40448
- 33 Matsuura B, Dong M, Miller L J. Differential determinants for peptide and non-peptidyl ligand binding to the motilin receptor. Critical role of second extracellular loop for peptide binding and action. J Biol Chem, 2002, 277: 9834–9839
- 34 Gether U. Uncovering molecular mechanisms involved in activation of G protein-coupled receptors. Endocr Rev, 2000, 21: 90–113
- 35 Fryxell K J. The coevolution of gene family trees. Trends Genet, 1996, 12: 364–369
- 36 Irwin D M. Evolution of hormone function: Proglucagon-derived peptides and their receptors. BioScience, 2005, 55: 583–591
- 37 Kaiya H, Kojima M, Hosoda H, et al. Amidated fish ghrelin: Purification, cDNA cloning in the Japanese eel and its biological activity. J Endocrinol, 2003, 176: 415–423
- 38 Kaiya H, Kojima M, Hosoda H, et al. Identification of tilapia ghrelin and its effects on growth hormone and prolactin release in the tilapia, *Oreochromis mossambicus*. Comp Biochem Physiol B Biochem Mol Biol, 2003, 135: 421–429
- 39 Unniappan S, Lin X, Cervini L, et al. Goldfish ghrelin: Molecular characterization of the complementary deoxyribonucleic acid, partial gene structure and evidence for its stimulatory role in food intake. Endocrinology, 2002, 143: 4143–4146
- 40 Bridgham J T, Carroll S M, Thornton J W. Evolution of hormone-receptor complexity by molecular exploitation. Science, 2006, 312: 97–101
- 41 Ohno S. Evolution by Gene Duplication. New York: Springer, 1970
- 42 Kondrashov F A. In search of the limits of evolution. Nat Genet, 2005, 37: 73–76

补充材料

图 S1 基因组的比较分析. 对比人类和其他脊椎动物 *GHRL* 和 *MLN* 基因座位相关区段. 人类 *GHRL* 和 *MLN* 基因座位相关区段显示在最上面, 其他物种对应的同源区段分别显示在下面, 每一区段对应的基因组位置标示于左列

图 S2 *motilin/ghrelin* 荷尔蒙基因家族系统发育关系, 基于最大节约法构建

图 S3 *motilin/ghrelin* 荷尔蒙基因家族系统发育关系, 基于邻接法构建

图 S4 *GHSR*, *MLNR* 在脊椎动物进化过程中, 进化速率的比较

表 S1 来自脊椎动物不同物种的 *ghrelin* 前体基因

表 S2 来自脊椎动物不同物种的 *motilin* 前体基因

表 S3 来自脊椎动物不同物种的 *GHSR*, *MLNR* 和 *GPR39*

本文的以上补充材料见网络版 csb.scichina.com. 补充材料为作者提供的原始数据, 作者对其学术质量和内容负责.

图 S1 基因组的比较分析. 对比人类和其他脊椎动物 GHRL 和 MLN 基因座位相关区段. 人类 GHRL 和 MLN 基因座位相关区段显示在最上面, 其他物种对应的同源区段分别显示在下面, 每一区段对应的基因组位置标示于左列.

图 S2 motilin/ghrelin 荷尔蒙基因家族系统发育关系, 基于最大节约法构建.

支持率显示在对应的结点旁边. 基于进化的最简约原理进行了树的定根, 其他的进化模式 (定根方式) 需要更多的基因重复和丢失事件, 因而不是最简约的.

图 S3 motilin/ghrelin 荷尔蒙基因家族系统发育关系，基于邻接法构建。

支持率显示在对应的结点旁边。基于进化的最简约原理进行了树的定根，其他的进化模式（定根方式）需要更多的基因重复和丢失事件。

图 S4 GHSR, MLNR 在脊椎动物进化过程中，进化速率的比较。

枝长表示氨基酸进化速率 ($\times 10^{-9}$ /氨基酸位点/年)。

表 S1 来自脊椎动物不同物种的 ghrelin 前体基因.

Species	References	Accession No.	Chr	Strand	Start	End
<i>Homo sapiens</i>	human	NM_016362	3	-	10303388	10306864
<i>Pan troglodytes</i>	chimpanzee	this work	3	-	10580209	10584506
<i>Pongo pygmaeus abelii</i>	Orangutan		3	+	59626696	59631012
<i>Macaca mulatta</i>	macaca	AY372274	2	+	50735050	50738523
<i>Callithrix jacchus</i>	Marmoset	this work	Contig2633	-	287048	291681
<i>Otolemur garnettii</i>	bush baby	this work				
<i>Cebus paella</i>	brown capuchin	AB365873				
<i>Aotus trivirgatus</i>	douroucoul	AB365874	6	-	113667047	113669423
<i>Hylobates lar</i>	common gibbon	AB365870				
<i>Pongo pygmaeus</i>	Bornean orangutan	AB365869				
<i>Macaca fuscata</i>	Japanese macaque	AB365871				
<i>Papio hamadryas</i>	hamadryas baboon	DQ987858				
<i>Saimiri sciureus</i>	common squirrel monkey	AB365872				
<i>Cavia porcellus</i>	guinea pig	this work	87	+	545525	549998
<i>Mesocricetus auratus</i>	golden hamster	EU863658				
<i>Mus musculus</i>	mouse	NM_021488	4	-	149622686	149628082
<i>Rattus norvegicus</i>	rat	NM_021669				
<i>Meriones unguiculatus</i>	gerbil	AF442491				
<i>Oryctolagus cuniculus</i>	rabbit	this work				
<i>Felis catus</i>	cat	AB089201				
<i>Canis familiaris</i>	dog	NM_001003052	20	+	11095912	11100316
<i>Equus caballus</i>	horse	XM_001491134				
<i>Rangifer tarandus</i>	reindeer	EU099031				
<i>Capra hircus</i>	goat	AB089200				
<i>Ovis aries</i>	sheep	NM_001009721				
<i>Bos taurus</i>	cattle	NM_174067	22	-	55704870	55708242
<i>Bubalus bubalis</i>	water buffalo	EF583468				
<i>Ailuropoda melanoleuca</i>	giant panda	EU375449				
<i>Sus scrofa</i>	pig	NM_213807				
<i>Kogia breviceps</i>	whale	AY455981				
<i>Myotis lucifugus</i>	microbat	this work				
<i>Erinaceus europaeus</i>	hedgehog	this work				
<i>Dasypus novemcinctus</i>	armadillo	this work				
<i>Echinops telfairi</i>	tenrec	this work				
<i>Loxodonta africana</i>	elephant	this work				
<i>Monodelphis domestica</i>	opossum	this work	6	-	195751382	195753698
<i>Gallus gallus</i>	chicken	NM_001001131	12	-	20095854	20097175
<i>Meleagris gallopavo</i>	turkey	AY333783				
<i>Anas platyrhynchos</i>	duck	AY338466				
<i>Anser sp.</i>	goose	AY338465				
<i>Dromaius novaehollandiae</i>	emu	AY338467				
<i>Trachemys scripta elegans</i>	turtle	AB161458				
<i>Rana catesbeiana</i>	bullfrog	AB058510				
<i>Rana esculenta</i>	edible frog	AM055941				
<i>Xenopus tropicalis</i>	Xenopus	this work				
<i>Danio rerio</i>	zebrafish	AM055940	6	+	27301754	27302348
<i>Carassius auratus</i>	goldfish	AF454389				
<i>Ictalurus punctatus</i>	catfish	AB196449				
<i>Oreochromis mossambicus</i>	Mozambique tilapia	AB077764				
<i>Oreochromis niloticus</i>	Nile tilapia	AB104859				

<i>Oncorhynchus mykiss</i>	trout	AB101443
<i>Acanthopagrus schlegelii</i>	bream	AY643808
<i>Anguilla japonica</i>	eel	AB062427

表 S2 来自脊椎动物不同物种的 motilin 前体基因.

Species		References	Accession No.	Chr	Strand	Start	End
<i>Homo sapiens</i>	human		NM_002418	6	-	33871258	33876918
<i>Pan troglodytes</i>	chimpanzee		XM_001171525	6	-	34513313	34518751
<i>Pongo pygmaeus abelii</i>	Orangutan	this work		6	-	33808003	33817684
<i>Macaca mulatta</i>	Macaca		NM_001032807	4	-	33505628	33510986
<i>Callithrix jacchus</i>	Marmoset			Con- tig4531	-	137634	143381
<i>Otolemur garnettii</i>	bush baby	this work					
<i>Cavia porcellus</i>	guinea pig		AF323752				
<i>Oryctolagus cuniculus</i>	rabbit		X63860				
<i>Felis catus</i>	cat		NM_001009278				
<i>Canis familiaris</i>	dog		XM_538869	12	-	6134852	6146848
<i>Ovis aries</i>	sheep		NM_001009439				
<i>Bos taurus</i>	cattle		NM_173938	23	-	8242398	8246432
<i>Sus scrofa</i>	pig		NM_214235				
<i>Equus caballus</i>	horse		AF047520				
<i>Myotis lucifugus</i>	microbat	this work					
<i>Erinaceus europaeus</i>	hedgehog	this work					
<i>Dasypus novemcinctus</i>	armadillo	this work					
<i>Echinops telfairi</i>	tenrec	this work					
<i>Loxodonta africana</i>	elephant	this work					
<i>Monodelphis domestica</i>	opossum	this work		2	-	26165031 7	26165758 1
<i>Ornithorhynchus anatinus</i>	platypus		XM_001510612				
<i>Anolis carolinensis</i>	Lizard	this work		106	+	1790153	1796000
<i>Gallus gallus</i>	chicken	Huang et al. 1999		26	+	4394282	4397580

References

Huang Z, Depoortere I, De Clercq P, Peeters T (1999) Sequence and characterization of cDNA encoding the motilin precursor from chicken, dog, cow and horse. Evidence of mosaic evolution in prepromotilin. *Gene* 240:217-26

表 S3 来自脊椎动物不同物种的 GHSR, MLNR 和 GPR39.

Species	Gene name	Chromosome	Strand	Start	End	Accession No.
<i>Homo sapiens</i>	HsGHSR	3	-	173645648	173648897	NM_198407
<i>Macaca mulatta</i>	RMGHSR	2	+	114582732	114585987	
<i>allithrix jacchus</i>	CjGHSR	Contig14511	+	12719	15958	
<i>Mus musculus</i>	MmGHSR	3	+	27270719	27273841	NM_177330
<i>Rattus norvegicus</i>	RnGHSR	2	+	113269623	113272744	NM_032075
<i>Ovis aries</i>	OaGHSR					NM_001009760
<i>Equus caballus</i>	EcGHSR	19	-	8641752	8645153	
<i>Bos taurus</i>	BtGHSR	1	+	97145780	97149019	
<i>Sus scrofa</i>	SsGHSR					NM_214180
<i>Canis lupus familiaris</i>	CfGHSR					NM_001099945
<i>Monodelphis domestica</i>	MdGHSR	7	-	197143200	197148178	
<i>Gallus gallus</i>	GgGHSR	9	+	20849992	20853507	NM_204394
<i>Anolis carolinensis</i>	AcGHSR	41	-	2644100	2652562	

<i>Xenopus tropicalis</i>	XtGHSR	scaffold_3	-	5090737	5099490	
<i>Danio rerio</i>	DrGHSR	2	-	21202474	21204624	
<i>Takifugu rubripes</i>	TrGHSR	scaffold_158	+	477561	479140	
<i>Tetraodon nigroviridis</i>	TnGHSR	10	-	9005119	9006568	
<i>Sphoeroides nephelus</i>	SnAF082209					AF082209
<i>Oryzias latipes</i>	OIGHSR	22	+	4983325	4984710	
<i>Acanthopagrus schlegelii</i>	AsGHSR					AY151040
<i>Gasterosteus aculeatus</i>	GaGHSR	GroupXV	-	10065093	10066414	
<i>Oncorhynchus mykiss</i>	OmGHSR					NM_001124594
<i>Homo sapiens</i>	HsMLNR	13	+	48692475	48694511	NM_001507
<i>Macaca mulatta</i>	RMMLNR	17	+	28043497	28045523	
<i>Oryctolagus cuniculus</i>	OcMLNR					NM_001099967
<i>Bos taurus</i>	BtMLNR	12	+	18236897	18238949	
<i>Canis lupus familiaris</i>	CfMLNR					NM_001128064
<i>Monodelphis domestica</i>	MdMLNR	4	-	309213170	309216380	
<i>Ornithorhynchus anatinus</i>	OraMLNR	Ultra336	-	5619346	5621406	
<i>Gallus gallus</i>	GgMLNR	1	+	173409608	173412826	NM_001127313
<i>Anolis carolinensis</i>	AcMLNR	223	-	175569	177166	
<i>Xenopus tropicalis</i>	XtMLNR	scaffold_167	+	1326643	1332588	
<i>Takifugu rubripes</i>	TrMLNR	scaffold_19	-	1396136	1397880	
<i>Tetraodon nigroviridis</i>	TnMLNR	16	-	5823817	5825573	
<i>Sphoeroides nephelus</i>	SnAF082210					AF082210
<i>Oryzias latipes</i>	OIMLNR	13	-	6217128	6224588	
<i>Gasterosteus aculeatus</i>	GaMLNR	GroupI	+	14404722	14407314	
<i>Danio rerio</i>	DrMLNR	15	+	46421916	46426447	
<i>Takifugu rubripes</i>	TrORF	scaffold_65	-	1256015	1258010	
<i>Tetraodon nigroviridis</i>	TnORF	NA	-	36940175	36943223	
<i>Sphoeroides nephelus</i>	SnAF082211					AF082211
<i>Oryzias latipes</i>	OIORF	8	+	23316228	23322619	
<i>Gasterosteus aculeatus</i>	GaORF	groupXI	+	2456033	2459723	
<i>Homo sapiens</i>	HsGPR39	2	+	132891086	133119646	NM_001508
<i>Pan troglodytes</i>	PtGPR39	2b	+	136509776	136743058	
<i>Macaca mulatta</i>	RMGPR39					XM_001116212
<i>Callithrix jacchus</i>	CjGPR39	Contig176	-	667554	901381	
<i>Cavia porcellus</i>	CpGPR39	3	-	54216357	54430055	
<i>Mus musculus</i>	MmGPR39	1	+	127573914	127769458	XM_913690
<i>Rattus norvegicus</i>	RnGPR39	13	+	38182156	38400275	NM_001114392
<i>Canis familiaris</i>	CfGPR39	19	+	38405073	38604858	
<i>Bos taurus</i>	BtGPR39	2	+	41141520	42316428	
<i>Equus caballus</i>	EcGPR39	18	+	16647271	16836311	
<i>Sus scrofa</i>	SsGPR39					NM_001128487
<i>Monodelphis domestica</i>	MdGPR39	4	+	129061182	129350619	
<i>Gallus gallus</i>	GgGPR39	7	+	31310727	31373166	NM_001080105
<i>Anolis carolinensis</i>	AcGPR39	704	+	1622	210619	
<i>Xenopus tropicalis</i>	XtGPR39	scaffold_2963	+	7655	8505	
		scaffold_2636	-	11546	12208	
<i>Takifugu rubripes</i>	TrGPR39					NM_001105235

<i>Tetraodon nigroviridis</i>	TnGPR39	2	+	18411390	18417125	
<i>Oryzias latipes</i>	OIGPR39	21	+	18691482	18725822	
<i>Gasterosteus aculeatus</i>	GaGPR39	groupXVI	+	6491812	6504200	
<i>Danio rerio</i>	DrGPR39	9	-	51273961	51276465	NM_001128771
<i>Homo sapiens</i>	HsNMUR1	2	-	232098001	232101972	NM_006056
<i>Homo sapiens</i>	HsNMUR2	5	-	151751948	151764867	NM_020167