

# 人类神经退行性疾病相关的三核苷酸重复 DNA 序列不稳定性机制研究进展

吕柯舜<sup>1</sup>, 潘学峰<sup>1,2,3</sup>

1. 北京理工大学生命学院, 北京 100081

2. 河北大学医学院, 保定 071030

3. 河北大学化学与环境科学学院, 保定 071002

**摘要:** 三核苷酸重复 DNA 序列扩增或缺失不稳定性与 50 多种人类神经退行性疾病有关。与疾病相关的三核苷酸重复拷贝数的增加或减少, 影响了特定基因的表达, 或因之产生具有细胞毒性的 RNA 和蛋白质已成为相关疾病的共有病理机制。现有的研究表明, 疾病相关的三核苷酸重复拷贝数的改变有可能起因于相关三核苷酸重复 DNA 序列的异常 DNA 复制、修复、重组以及基因转录。有关人类遗传学研究也提示, 发生在疾病相关的三核苷酸重复 DNA 部位的异常 DNA 复制、修复、重组或基因转录有可能在三核苷酸重复 DNA 不稳定过程中发挥着关键作用。本文根据本课题组的研究经验, 综述了近年来有关疾病相关三核苷酸重复不稳定性机制的研究进展, 包括碱基突变不稳定、重复单元的扩增和缺失不稳定, 以助更好地理解疾病相关三核苷酸重复 DNA 序列不稳定性的分子机制。

**关键词:** 三核苷酸重复; 不稳定性扩增; R-环; 非 B 型 DNA 二级结构; 神经退行性疾病

## Progress on the mechanistic research of the trinucleotide repeat instabilities underlying human neurodegenerative diseases

Kenao Lv<sup>1</sup>, Xuefeng Pan<sup>1,2,3</sup>

1. School of Life Science, Beijing Institute of Technology, Beijing 100081, China

2. School of Medicine, Hebei University, Baoding 071030, China

3. School of Chemistry and Environmental Science, Hebei University, Baoding 071002, China

**Abstract:** The expansion and deletion instabilities shown by some trinucleotide repeated DNA sequences are associated with more than 50 neurodegenerative diseases in humans. The increase or decrease of the trinucleotide repeat units underlying the diseases are not yet clearly explained using any mechanism, but has been found to affect the expression of specific genes, or produces cytotoxic RNA and protein, which has now become a common pathological

收稿日期: 2021-05-24; 修回日期: 2021-07-25

基金项目: 北京市自然科学基金(No. 5132014)和河北省医学研究计划重点项目(No. 20160051)资助[Supported by Beijing Natural Science Foundation (No. 5132014), and Hebei Medical Research Key Program (No. 20160051)]

作者简介: 吕柯舜, 在读硕士研究生, 专业方向: 分子生物学。E-mail: 2829435034@qq.com

通讯作者: 潘学峰, 教授, 研究方向: 分子药理学、分子生物学、药物递送。E-mail: xuefengpancam@aliyun.com

DOI: 10.16288/j.ycz.21-182

网络出版时间: 2021/8/20 17:26:37

URI: <https://kns.cnki.net/kcms/detail/11.1913.r.20210819.1550.005.html>

mechanism of the diseases. The ongoing studies have shown that the changes in the copy numbers of the disease-related trinucleotide repeats may result from abnormal DNA replication, repair, recombination, and gene transcription. Human genetical studies also suggest that abnormal DNA replication, repair, recombination, or gene transcription that occurred in the disease-related trinucleotide repeat DNA sites may play a key role in the trinucleotide repeat DNA instabilities. Based on the research experiences of our research group, this paper reviews the recent research progress on the mechanisms of the disease-associated trinucleotide repeat DNA instabilities including their base mutation instabilities, the amplification and deletion instabilities of the repeat units, to better understand the molecular mechanism of the disease-associated trinucleotide repeats instabilities.

**Keywords:** trinucleotide repeats; expansion instability; RNA: DNA hybrids; non-B DNA secondary structure; neurodegenerative disease

自 1991 年 Verkerk 等<sup>[1]</sup>发现位于人类脆性 X 染色体综合征(fragile X syndrome) *FMR-1* 基因非编码区的简单串联重复 DNA 序列为 CGG 三核苷酸重复 DNA 序列(trinucleotide repeats, TNRs), 以及 La Spada 等<sup>[2]</sup>发现位于肯尼迪氏病(Kennedy's disease, 又名 X 连锁脊髓和延髓肌萎缩, X-linked spinal and bulbar muscular atrophy)肾上腺受体基因(androgen receptor gene, *AR*)编码区内的简单串联重复 DNA 序列为 CAG 三核苷酸重复 DNA 序列, 迄今已发现 50 种以上的神经退行性疾病与三核苷酸重复 DNA 序列扩增有关。与疾病有关的三核苷酸重复 DNA 序列的不稳定性扩增可能位于基因的非编码区, 如内含子或编码序列侧翼的非翻译调控区[如脆性 X 染色体综合征、弗洛里德共济失调(Friedreich ataxia, *FRDA*)、强直性肌营养不良(myotonic dystrophy, *MD*)及脊髓小脑共济失调 8 型(spino cerebellar ataxia type 8)等]。在这种情况下, 三核苷酸重复 DNA 序列可能因扩增而变得很长, 由成千上万个重复单位组成。这些新增的三核苷酸重复 DNA 通过多种机制对受累基因产生影响, 如诱导基因沉默而造成功能丧失(loss of function), 或通过其他方式获得新功能(gain of function), 如从有义链或反义链产生“毒性”RNA<sup>[3,4]</sup>。这些毒性 RNA 干扰 mRNA 正常加工以及与特定蛋白的结合。位于编码序列内的三核苷酸重复 DNA 序列的扩增与非编码区域内发生的三核苷酸重复 DNA 序列扩增相比, 其长度往往适中<sup>[5,6]</sup>, 而且均为 CAG 重复序列, 编码聚谷氨酰胺(polyglutamine, polyQ)<sup>[7]</sup>。据此, 人们通常把三核苷酸重复 DNA 序列有关的神经退行疾病分为三类: 第一类由 CAG 重

复序列引起, 包含亨廷顿舞蹈病(Huntington's disease, *HD*)和脊髓小脑共济失调(spino cerebellar ataxia)等在内的 16 种疾病<sup>[8-10]</sup>; 第二类是在表型上更加多样化且异常扩增规模较小, 由存在于外显子中的三核苷酸重复 DNA 序列不稳定性扩增引起<sup>[11,12]</sup>; 第三类则与位于内含子等非编码区的三核苷酸重复 DNA 序列不稳定有关, 包括脆性 X 染色体综合征、强直性肌营养不良、脊髓小脑共济失调、青少年肌阵挛性癫痫(juvenile myoclonic epilepsy, *JME*)和弗里德里希共济失调<sup>[13]</sup>。除此以外, 还可依据病理过程中涉及由扩增的 CAG 三核苷酸重复 DNA 序列编码而产生的 polyQ 的影响, 分为 polyQ 疾病和非 polyQ 疾病(表 1)。其中属于 polyQ 疾病的病理共性涉及蛋白质聚集、蛋白水解切割、转录失调、自噬损伤和线粒体功能障碍等分子机制<sup>[14]</sup>。但最近的研究对 polyQ 长度决定发病率这一观点提出质疑, 转而认为 polyQ 疾病的发病率取决于多个 DNA 维持基因多态性变异<sup>[15]</sup>。

尽管对引发多种神经系统疾病的三核苷酸重复 DNA 序列异常扩增机制和表现出染色体“脆性”(fragility)现象的研究已历时数十年之久, 但迄今依然对疾病状态下三核苷酸重复 DNA 序列异常扩增的分子机制缺乏透彻理解。在相当长的一段时间里, 很多人相信疾病相关三核苷酸重复 DNA 序列在特定情况下通过形成非 B 型 DNA 结构(non-B DNA structure)影响 DNA 复制、基因转录、修复、重组等过程, 并由此影响自身在基因组内的稳定性, 出现碱基突变、重复单位的扩增或缺失<sup>[32,33]</sup>。但已报道的研究表明, 至少在细菌、酵母(*Saccharomyces*)、线

表 1 与三核苷酸重复扩增相关的疾病信息汇总

Table 1 Summary of the trinucleotide repeat expansions associated disease information

类型	疾病名称	基因	三核苷酸重 复 DNA 序列	基因 结构	染色体 定位	正常 重复数	异常 重复数	首次报 道时间	参考 文献
PolyQ 三核苷 酸扩增 性疾病	脊髓小脑共济失调 1 型	<i>ATXN1</i>	CAG	ORF	6p22.3	6~9	41~83	1993	[16]
	脊髓小脑共济失调 2 型	<i>ATXN2</i>	CAG	ORF	12q24.12	14~32	33~200	1996	[17,18]
	脊髓小脑共济失调 3 型	<i>ATXN3</i>	CAG	ORF	14q32.12	12~40	55~86	1994	[19]
	脊髓小脑共济失调 6 型	<i>CACNA1A</i>	CAG	ORF	19p13.13	4~18	21~33	1997	[20]
	脊髓小脑共济失调 7 型	<i>ATXN7</i>	CAG	ORF	3p14.1	7~17	38~120	1997	[21]
	脊髓小脑共济失调 12 型	<i>PPP2R2B</i>	CAG	5'UTR	5q32	7~41	>51	1999	[22]
	脊髓小脑共济失调 17 型	<i>TBP</i>	CAG	ORF	6q27	25~44	47~63	2001	[23]
	亨廷顿舞蹈病	<i>HTT</i>	CAG	ORF	4p16.3	6~35	36~250	1993	[24]
	脊髓延髓肌萎缩	<i>AR</i>	CAG	ORF	Xq12	9~36	38~62	1991	[25]
非 polyQ 三核苷 酸扩增 性疾病	易碎性 X 综合征	<i>FMR1</i>	CGG	5'UTR	Xq27.3	6~53	>230	1991	[26,27]
	易碎性 X 智力低下(马丁-贝尔综合征)	<i>FMR2</i>	CGG	5'UTR	Xq28	6~53	>200	1993	[28]
	弗洛里德共济失调	<i>FXN</i>	GAA	内含子	9q21.11	7~34	>70	1996	[29]
	强直性肌营养不良	<i>DMPK</i>	CTG	3'UTR	19q13.32	5~37	>50	1992	[30]
	脊髓小脑共济失调 8 型	<i>SCA8</i>	CTG	ORF	13q21.33	16~37	90~250	1999	[31]

虫(*Caenorhabditis elegans*)或人类细胞系(human cell lines)等模式生物中疾病相关的三核苷酸重复 DNA 序列所能表现的不稳定性多为“变短”(缺失)<sup>[34]</sup>, 而不是患者基因中所表现的“变长”(扩增)。更为遗憾的是, 很多利用模式生物所得到的研究结果也通常自相或者相互矛盾, 因此很难对现有的任何一种机制达成共识, 很难真正明确疾病状态下三核苷酸重复 DNA 序列体内扩增机制及与疾病发生和发展相关的病理机制<sup>[35,36]</sup>。基于本课题组多年的研究经验, 并结合人类基因组内不同位点处三核苷酸重复 DNA 序列以纯合(pure repeats)和非纯合(interrupted repeats)形式分布所表现出的稳定性差异, 本文尝试把疾病相关的三核苷酸重复 DNA 序列的稳定性分为碱基突变不稳定(base mutation instability)和重复单位扩增和缺失不稳定(repeats expansion instability), 并针对疾病状态下这些不稳定性发生的分子机制进行了系统论述。

### 1 三核苷酸重复DNA序列不稳定诱发的细胞毒性与疾病病理关联

三核苷酸重复 DNA 序列疾病的致病原因可分为基因功能丧失和所编码产物功能的改变。与基因

功能丧失有关的症候多以常染色体隐性或者 X 连锁隐性方式外显; 而起因于基因编码产物功能改变的症候则一般以常染色体显性或者 X 连锁显性遗传方式外显<sup>[37]</sup>。三核苷酸重复不稳定性扩增会诱导细胞产生代谢改变, 从而损害特定组织的特定功能, 其机制主要包括蛋白质功能的丧失、RNA 或蛋白质水平上的毒性增加<sup>[38]</sup>。位于可编码区的重复序列扩增会导致转录缺陷, 从而使功能基因失活丧失作用。这种转录缺陷通常表现为改变基因启动子的甲基化模式<sup>[39,40]</sup>、改变 mRNA 的剪接模式<sup>[41]</sup>, 或者直接改变蛋白表达水平<sup>[42]</sup>。较长的重复序列会阻碍相应基因的转录, 因为它需要更长的转录时间以及资源, 会使相应基因产物减少<sup>[43]</sup>。例如, 重复序列可能会形成 R-环(RNA-DNA 杂交体), 这种异常的二级结构会直接影响基因转录。在 *FRDA* 中, 由于 *FXN* 基因中的 GAA 重复序列形成了 R 环结构, 导致 RNA 聚合酶无法通过从而阻碍基因的正常转录<sup>[29,44,45]</sup>。越来越多的证据表明, 某些重复序列自身会表现出或强或弱的细胞毒性。转录形成的 RNA 会参与多价碱基配对, 从而导致其凝胶化, 最终聚集汇成核灶<sup>[46]</sup>。这些 RNA 会螯合某些结合蛋白, 通过破坏核孔复合物, 损害 mRNA 的转运<sup>[47]</sup>。Thomas 等<sup>[48]</sup>发现虽然组蛋白甲基转移酶与富含 GC 的 RNA 体外结合能力

较弱,但受到体内解旋酶和 RNA 结合蛋白等蛋白因子的调节会形成高不溶性复合物。蛋白质毒性增加也会导致细胞毒性。例如, CAG 重复序列会翻译成多聚谷氨酰胺片段(polyQ 片段), polyQ 片段错误折叠形成  $\beta$ -折叠结构会表现出异常的极性从而非特异性结合各种调节蛋白,最终在神经元细胞中汇集成毒性包涵体<sup>[49]</sup>。

## 2 三核苷酸重复 DNA 序列不稳定性的分子机制

人类基因组中含有大量的串联重复 DNA 序列<sup>[50]</sup>,主要分布在染色体的着丝粒、端粒和某些基因的调节区域<sup>[51]</sup>。已知的大多数串联重复 DNA 序列并不会出现急剧扩增,只有部分三核苷酸重复 DNA 序列会易于大规模扩增导致疾病(表 1)。体外研究发现,与疾病相关的三核苷酸重复 DNA 序列通常会形成分子内结构影响基因转录、翻译和与某些特定蛋白结合。比如,部分三核苷酸重复 DNA 序列易形成 R 环阻碍基因转录,导致编码蛋白表达的降低。而富含 GC 的三核苷酸重复 DNA 序列则会发生超甲基化,从而沉默基因(例如 FXS、FRAXE)或者采用高度稳定的分子内折叠结合 RNA 结合蛋白(例如 DM1、FXTAS、SCA6、SCA3)<sup>[52]</sup>。在亨廷顿氏病和肌强直性营养不良中重复序列还会通过复制叉停滞引发三核苷酸重复扩增疾病<sup>[53,54]</sup>。针对这些现象,研究者们提出了不同的分子机制模型试图解释三核苷酸重复 DNA 序列的扩增机制。

### 2.1 非 B 型 DNA 二级结构介导的扩增机制

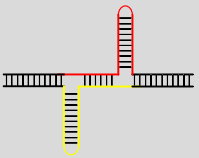
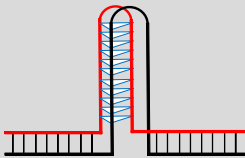
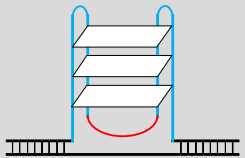
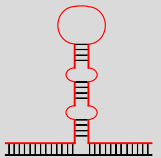
除经典的 B 型构象(B conformation)之外, DNA 分子还能在特定条件下形成非 B 型二级结构(non-B DNA secondary structure)<sup>[55]</sup>,与三核苷酸重复不稳定性有关的非 B 型 DNA 二级结构见表 2。表 2 列出了三核苷酸重复 DNA 序列所能形成的“slipped DNA”(滑脱 DNA)、H-DNA、G4 偶联体和含错配对的发卡结构等主要形式(表 2)。这些非 B 型 DNA 二级结构一旦出现在 DNA 复制、修复、重组和转录过程中则有可能介导三核苷酸重复 DNA 序列的碱基突变不稳定和重复单位扩增和缺失不稳定的发生。

#### 2.1.1 DNA 滑脱复制机制

DNA 滑脱复制(slippage DNA replication)模型见图 1。此模型的建立是基于疾病相关的三核苷酸重复 DNA 序列在复制或修复过程中常伴随局部的缺失或插入(deletion/insertion, D/I)(图 1)<sup>[56,57]</sup>。大量的体外研究证据表明,体外复制三核苷酸重复 DNA 序列时, DNA 复制酶会与重复 DNA 序列模板频繁发生“脱离-识别”,使得新合成链的 3'端有机会沿模板链“回缩”,然后再与重复序列中的另外一组重复单位配对,重新进行 DNA 复制(图 1)。这种 DNA 复制使得三核苷酸重复 DNA 的某些区段会被更多或者更少的复制<sup>[57]</sup>。尽管 DNA 滑脱模型可以通过体外实验得以验证,但该模型无法解释为什么只有部分重复序列会出现急剧不稳定性扩增,而其他序列不容易发生;也无法解释扩增的不稳定性是指数增加而非线性增加。

表 2 三核苷酸重复 DNA 形成的非 B 型 DNA 二级结构

Table 2 Schematic diagram of non-B DNA secondary structures adopted by trinucleotide repeats DNA

非 B 型 DNA 二级结构	Slipped DNA	H-DNA	G4 偶联体	含错配发卡结构
示意图				
三核苷酸重复 DNA 序列	所有 10 种三核苷酸重复	GAA	CGG	CAG
		TTC	AGG	CTG
			TGG	CGG
				CCG

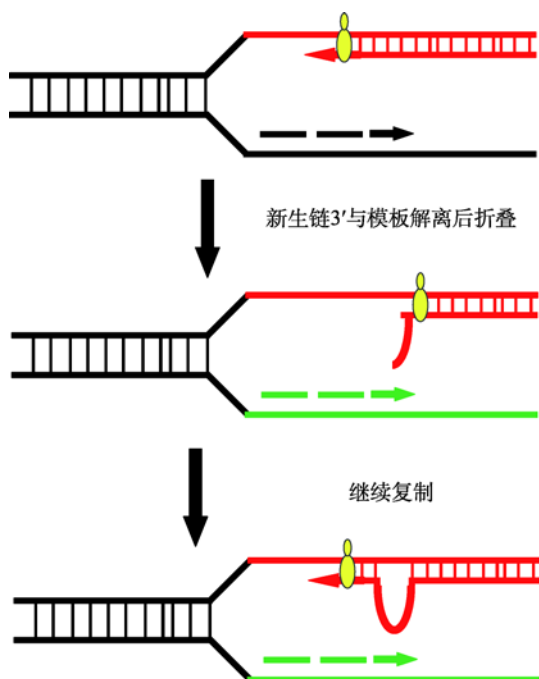


图1 三核苷酸重复 DNA 序列的滑脱复制

Fig. 1 Slippage DNA replication of the trinucleotide repeats

复制或修复过程中, 新合成链 3' 段回缩导致 DNA 局部错位; 复制或转录过程继续进行, 导致重复序列异常增加或减少。

## 2.1.2 DNA 解链元件

因为 DNA 滑脱模型无法完美解释观察到的现象, 研究者转而试图探究在包括 DNA 复制、转录、修复和重组等代谢过程中双链 DNA 被解链呈 DNA 单链后形成可替代 DNA 结构(H-DNA、G4 偶联体、含错配的发卡等(结构见表 2)对三核苷酸重复 DNA 序列扩增和缺失不稳定性可能贡献。

### 2.1.2.1 H-DNA 与三核苷酸重复 DNA 序列不稳定

1987 年首次报道了嘌呤-嘧啶镜面重复序列可以形成分子内三螺旋 DNA 结构, 被称为 H-DNA (表 2)。它们通过分子内的 Hoogsteen 或反向 Hoogsteen 碱基对堆叠以及  $\pi$ - $\pi$  相互作用而稳定<sup>[58]</sup>, 一半的 DNA 链通过折回与对应镜像链结合形成三重螺旋。H-DNA 的形成在负超螺旋 DNA 中是有利的, 因为它在拓扑学上等同于 DNA 的展开。H-DNA 结构在真核生物中普遍存在, 这暗示它们在基因组功能上有进化保守作用。它的存在可能涉及 DNA 复制及转录的停止, 促进双链断裂, 引发相应的修复<sup>[59]</sup>, 因此 H-DNA 的形成可能是部分三核苷酸重复 DNA 序

列不稳定性扩增的关键。

### 2.1.2.2 G4 偶联体与三核苷酸重复 DNA 序列不稳定

G4 偶联体首次发现是在免疫球蛋白转换区, 它是四链 DNA 结构, 由多个  $\pi$ - $\pi$  键相互作用构成。在单价阳离子钠和钾的存在下, G4 偶联体可以稳定存在(表 2)。G4 偶联体具有多种异构体形式, 以很高的丰度存在于诸多生物的基因组中, 在人等真核生物染色体端粒及一些区域最为常见。能够形成 G4 偶联体的三核苷酸重复 DNA 序列是 CGG、AGG 和 TGG, 其中 CGG 与包括脆性 X 染色体综合征在内的人类多种退行性疾病相关联<sup>[60]</sup>, 而 TGG 则与人群中复发性 14q32.2 染色体微缺失(recurrent 14q32.2 microdeletion)有关<sup>[61,62]</sup>。G4 偶联体可以干扰转录、影响基因组的完整性, 甚至会促进 DNA 的双链断裂, 这预示着 G4 偶联体在三核苷酸重复 DNA 序列不稳定性扩增中起着重要作用。

### 2.1.2.3 含错配的发卡结构与三核苷酸重复 DNA 序列不稳定

最早由美国 McMurray 实验室证明的含错配发卡结构(也称作 DNA 十字形或 S-DNA)<sup>[63]</sup>, 由其上游序列通过正常的 Watson-Crick 碱基配对与互补的下游区域退火结合形成(表 2)。实验证明三核苷酸重复 DNA 序列 CAG、CTG、CGG 和 CCG 等会形成含错配的发卡结构。DNA 十字形和 S-DNA 在形成动力学上存在差异<sup>[64]</sup>。DNA 十字形初期的形成非常缓慢, 它需要形成中心气泡, 然后 DNA 链再退火形成十字形; 而 S-DNA 则需要先解链一段较长的 DNA, 然后通过单链 DNA 进行退火配对, 因其构象在热力学上是不利的, 因而 S-DNA 无法轻易转化为双链 DNA。涉及 DNA 双链展开的含错配的发卡结构在拓扑异构学上等同于完全解链的 DNA, 因而在 DNA 复制、转录以及修复中起重要作用<sup>[65]</sup>。

### 2.1.2.4 基于非 B 型 DNA 二级结构的其他机制

三核苷酸重复 DNA 序列形成的非 B 型 DNA 二级结构可以帮助疾病相关的三核苷酸重复 DNA 序列逃避细胞内忠实的 DNA 复制模式, 从而造成包括重复序列扩增不稳定在内的多种影响, 这些影响有可能转而加剧三核苷酸序列的扩增不稳定。与此有关的研究包括 CAG 重复扩增产生的 siRNA 可以通过 RNAi 在体外有效的杀死癌症细胞, 这至少表明癌细胞对三核苷酸不稳定性扩增衍生的 siRNA 特别

敏感<sup>[66]</sup>；而利用锌指蛋白转录因子(ZFP-TFs)或者 CRISPR/Cas9 靶向 CAG 致病重复序列选择性降低突变蛋白表达的治疗策略在小鼠脑中表现较好的耐受性，并使小鼠脑在分子生物学、组织病理学、电生理学等多方面得到改进<sup>[67-70]</sup>。除此之外，重复相关的非 AUG 翻译中发现 RNA 二级结构和启动子保真度是调节 FMR1 RNA 翻译的关键因素<sup>[27]</sup>。此类实验结果暗示有必要进一步探究基于非 B 型 DNA 二级结构影响的更为复杂的三核苷酸重复 DNA 序列扩增机制。

## 2.2 三核苷酸重复 DNA 不稳定性扩增的其他机制

### 2.2.1 “ori-shift/switch”模型

已有的研究数据表明，疾病相关的三核苷酸重复 DNA 序列不稳定性可以表现出对 DNA 复制方向的依赖性。有人据此提出了“ori-shift/switch”模型(图 2)。该模型假定重复序列复制时三核苷酸重复 DNA 序列与复制起始位点 ori 之间发生了 DNA 序列插入(图 2A)，或位于三核苷酸重复 DNA 序列一侧的复制起始位点失活，重复序列转而利用其相反一侧的 DNA 复制起点进行复制(图 2B)。这种复制方向的改变会使得特定三核苷酸重复 DNA 序列的稳定性发生变化，从而表现出模型实验或临床实践中常见的

三核苷酸重复 DNA 序列的扩增和缺失<sup>[71]</sup>。这个模型强调了三核苷酸重复 DNA 序列的稳定性可能很大程度上取决于它与最近复制起始位点的距离(图 2A)或复制起始位点的使用(图 2B)。具有扩增潜力的三核苷酸重复 DNA 序列本身有可能会影响邻近 DNA 复制起始位点的使用，因此促发 DNA 复制起始位点的转换。这种模型可以帮助理解疾病相关的三核苷酸重复 DNA 序列不稳定性的改变可表现出“阈值”的现象，即当重复序列拷贝数达到一定数目之后，相应的三核苷酸重复 DNA 序列会愈发不稳定。

### 2.2.2 R-环介导的扩增机制

以上所有机制模型都无法完美的解释三核苷酸重复 DNA 序列的不稳定性扩增，基于多年的研究积累本课题组在 2006 年首次提出了 R-环介导的扩增机制<sup>[72,73]</sup>。已有的实验数据证明，重复序列越长给细胞生长造成的压力越大，而当重复数增加达到相应阈值后会造成细胞无法继续生长。有学者认为这可能是达到特定阈值的三核苷酸重复 DNA 序列会首先影响细胞内的 DNA 复制，而本课题组一步的研究表明，造成这一现象的真正原因是在转录过程中出现了障碍。通过抑制转录过程，可以使原本无法正常生长的细胞恢复生长，这表明复制障碍在其中的作用不大(待发表)。通过对这些三核苷酸重复

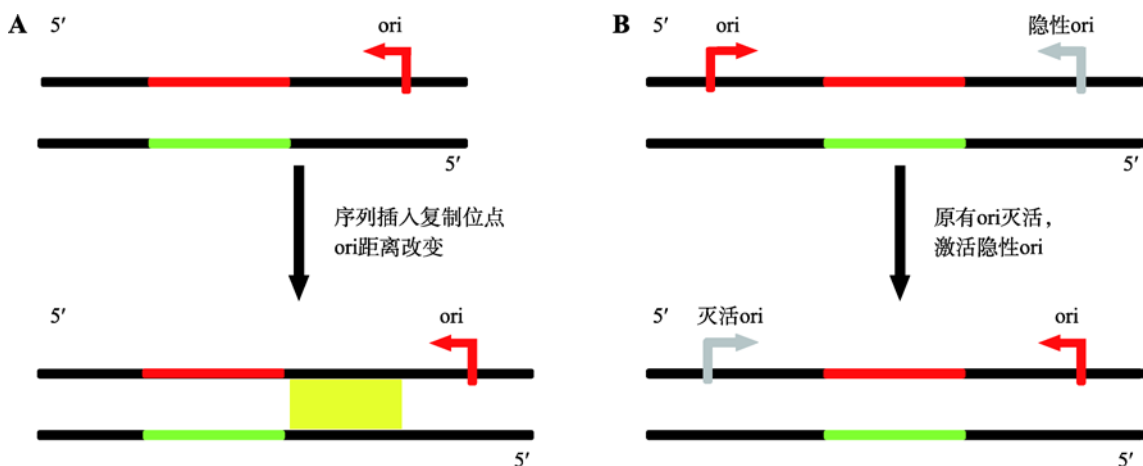


图 2 ori-shift/switch 机制

Fig. 2 ori-shift/switch mechanism

A: 插入因改变三核苷酸重复 DNA 与 ori 的距离从而引发三核苷酸重复 DNA 序列扩增或缺失；B: 重复序列一侧的复制起点失活，激活位于相反一侧的隐性复制起始位点，从而使重复序列发生扩增或缺失。

DNA 序列片段的进一步分析发现, 带有特定碱基突变或序列缺失的长三核苷酸重复 DNA 序列可逃逸基因转录压力, 从而避免对细胞生长繁殖的影响, 而未能获得碱基突变或序列缺失的长三核苷酸重复 DNA 序列依然能够在基因转录过程中出现问题, 从而影响细胞的正常繁殖。这一现象暗示了, 在转录过程中, 达到阈值的重复序列会产生特定的二级结构, 这一构象会阻碍转录的正常进行, 进而引发一系列的修复反应, 但精确修复过后的细胞仍然面临相同的问题, 无法越过转录障碍陷入修复循环, 导致细胞无法生长繁殖最终死亡。更直接的证据是在我们选用了不忠实的 DNA 复制酶后, 原本无法正常生长繁殖的带有长重复序列的细胞可以恢复生长。通过分析 DNA 提取物, 我发现长重复序列形成了特殊的二级结构, 结合之前的研究结果, 我们进一步完善了 R-环介导的扩增机制<sup>[71-76]</sup>。

基因转录需要在 RNA 聚合酶的催化下, 以双链 DNA 中的一条链为模板合成 RNA。正常情形下, 基因转录产生的 RNA 分子必须在转录泡后方与 DNA 模板分开, 以容许两条被打开呈单链状态的 DNA 分子重新恢复成双链 DNA。但包括本课题组在内的一些实验室的研究表明, 发生在 GAA•TTC、CGG•CCG 和 CAG•CTG 等疾病相关的三核苷酸重复 DNA 序列部位的基因转录所合成的 RNA 分子通常难以与其模板 DNA 分开, 或很容易重新与模板 DNA 复性形成 RNA:DNA 杂交体。这种 RNA:DNA 杂交体与非模板链一起被称为 R-环结构<sup>[45]</sup>。R-环结构的形成不仅会阻碍转录复合体的正常通过, 使转录暂停并能引发修复反应。当 R-环在模板链上形成时, 它会留下互补的部分单链, 从而诱发非 B DNA 结构的形成。因此, 一方面, R-环在调节基因表达和转录终止过程中起重要作用。另一方面, 它们还可以阻止复制和转录, 造成复制-转录冲突, 并通过引发修复造成 DNA 链断裂(DSB), 造成基因组不稳定。这一机制过程可以完美的解释目前实验中观察到的现象。在实验过程中本课题还发现, 当 CTG 重复序列位于前导链时的抑制程度要比含相同重复单位的 CAG 重复序列位于前导链更加严重。通过对转录产物进行热力学稳定性分析发现 CTG 形成的二级结构要比 CAG 形成的二级结构更加稳定(待发表)。这种情况解释了转录复合体以 CAG 重复序列为模板进行转

录比以 CTG 重复序列为模板进行转录更具优势的原因。

### 2.2.3 不忠实同源重组促发三核苷酸重复 DNA 序列不稳定

人类基因组中以纯合和非纯合形式分布着 10 种类型的三核苷酸重复 DNA 序列(图 3)<sup>[77]</sup>, 其中, 尚未明确与任何人类疾病有关的 TAA、CAA 三核苷酸重复 DNA 序列多以含有 8 个重复单位及以上的纯合形式存在, ATG、GGA 和 TGG 则表现为单位重复数多于或等于 100 的非纯合形式明显多于纯合形式(图 3)。而与疾病密切相关的三核苷酸重复 DNA 序列中的 CAG•CTG、CGG•CCG 则多以重复单位数小于或等于 8 的纯合形式存在, 鲜见重复单位数大于 100 非纯合形式, 而 GAA•TTC 则表现为重复单位数多于或等于 100 的非纯合形式多于含重复单位数为 8 及以上的纯合形式(图 3)。

本课题组在研究中发现, 利用克隆手段构建与疾病相关的长三核苷酸重复 DNA 序列过程中存在着一个 DNA 突变阈值, 即当所构建的重复序列的重复单位数达到一定值后, 所得到的重复序列多为非纯合重复序列, 很少或几乎得不到不含碱基突变和缺失的纯合形式的三核苷酸重复 DNA 序列。显然这种 DNA 突变阈值有别于三核苷酸重复 DNA 序列体内所能表现出的扩增阈值。当前这种情况已被多个实验室确证<sup>[78,79]</sup>。据此, 我们推测 DNA 碱基突变阈值的出现可能与三核苷酸重复以如图 4 所示的同源

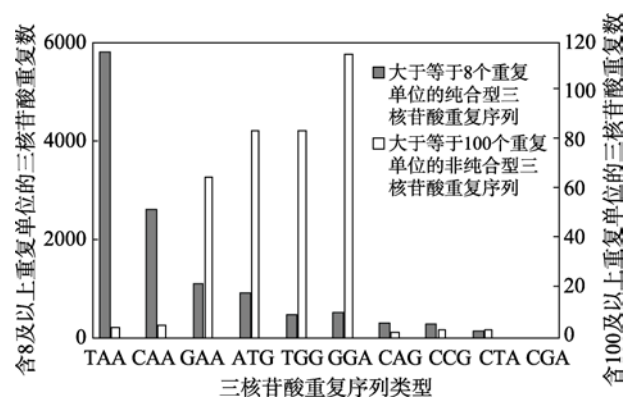


图3 人类基因组中三核苷酸重复 DNA 序列的分布

Fig. 3 Distribution of trinucleotide repeats in human genome

参照文献<sup>[77]</sup>等呈现的数据信息绘制。

复制机制有关(图 4)。当三核苷酸序列重复数目达到一定值后,序列本身会对复制或转录造成阻碍。可预见的情形包括 DNA 聚合酶转换(DNA polymerase switching),即由原本使用具有校对活性的 DNA 聚合酶转而使用不具校对活性的 DNA 聚合酶复制处于 DNA 突变阈值的三核苷酸重复 DNA 序列。在这个过程中,包括碱基错配修复蛋白 MutS (原核)和 MSH2-MSH6(真核)等会参与其中,这种情况反而带来了更大困难和更高的出错率。为了验证这种可能性,本课题组曾把 DNA 聚合酶的校对活性灭活或者高表达缺乏校对活性的不忠实 DNA 聚合酶 DinB (又称 DNA 聚合酶 IV,与人类细胞内的 DNA 聚合酶 Pol  $\kappa$  相当,是三界生物共有的一种不具 3'→5'校对活性的非忠实性 DNA 复制酶)反而会大大降低重

复序列出错的几率。之所以会出现这种 DNA 复制忠实性的悖论,我们认为这是由于已处于 DNA 突变阈值的三核苷酸重复 DNA 序列本身不容易被复制和转录,特别当碱基错配修复蛋白试图修复 DNA 复制或转录过程中的问题模板时反而会进一步增加相关三核苷酸重复 DNA 序列区的 DNA 复制和基因转录的难度,此时,有可能进一步转化出包括 DNA 链断裂(如 Mre11 等结构特异性核酸酶的内切)在内的 DNA 损伤,迫使细胞调用包括同源重组依赖的 DNA 复制机制参与三核苷酸重复 DNA 序列的复制。此时一旦发生不忠实性同源重组(如图 5 所示)就有可能使处于 DNA 突变阈值的三核苷酸重复 DNA 序列进入重复单位数扩增阈值。这种策略是生物体为了降低复制和转录时的出错率进化出来的;当只有不忠

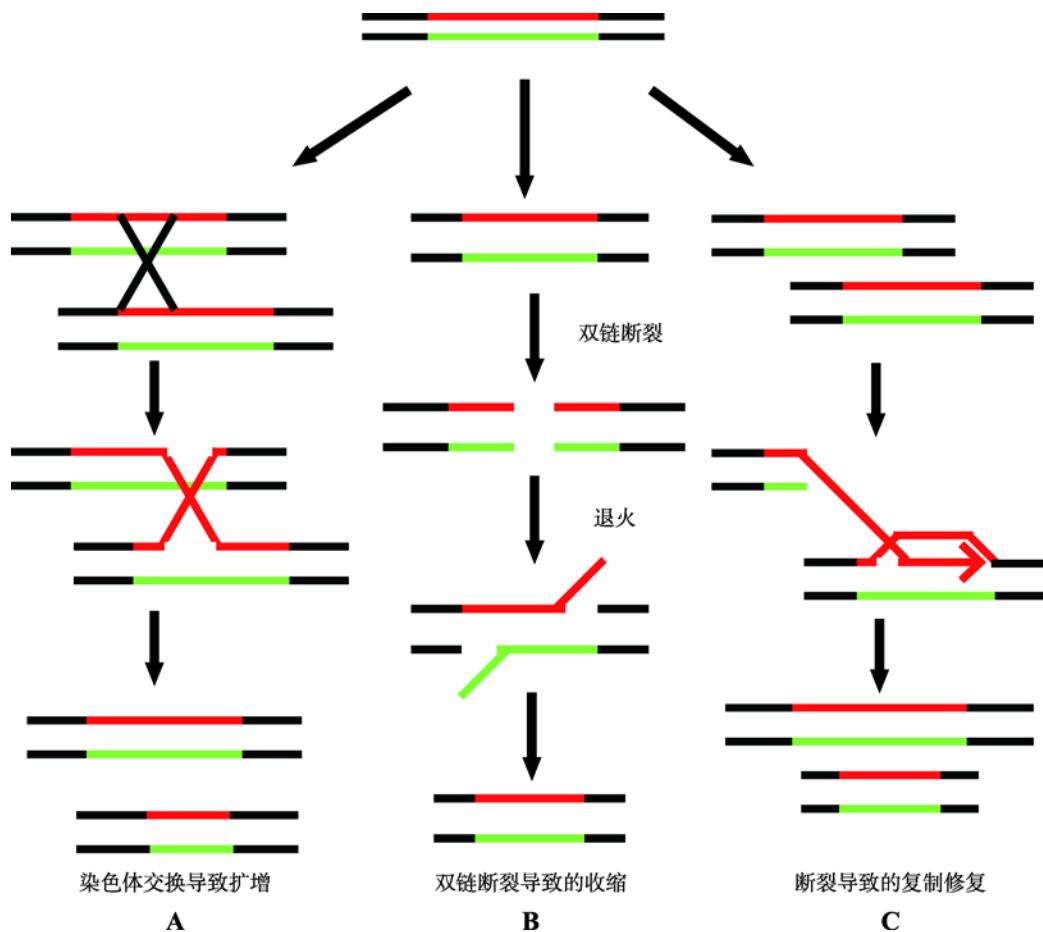


图 4 三核苷酸重复 DNA 序列同源复制机制

Fig. 4 The homologous replication mechanism of TNRs

A: 同源染色体在重复序列发生交换, 一条染色体获得收缩的重复序列, 而另一条染色体获得扩增的重复序列; B: 重复序列发生双链断裂, 单链重复末端退火连接, 多余末端切除导致重复序列收缩; C: 重复序列处双链断裂, 引发修复机制, 导致重复序列的异常扩增。

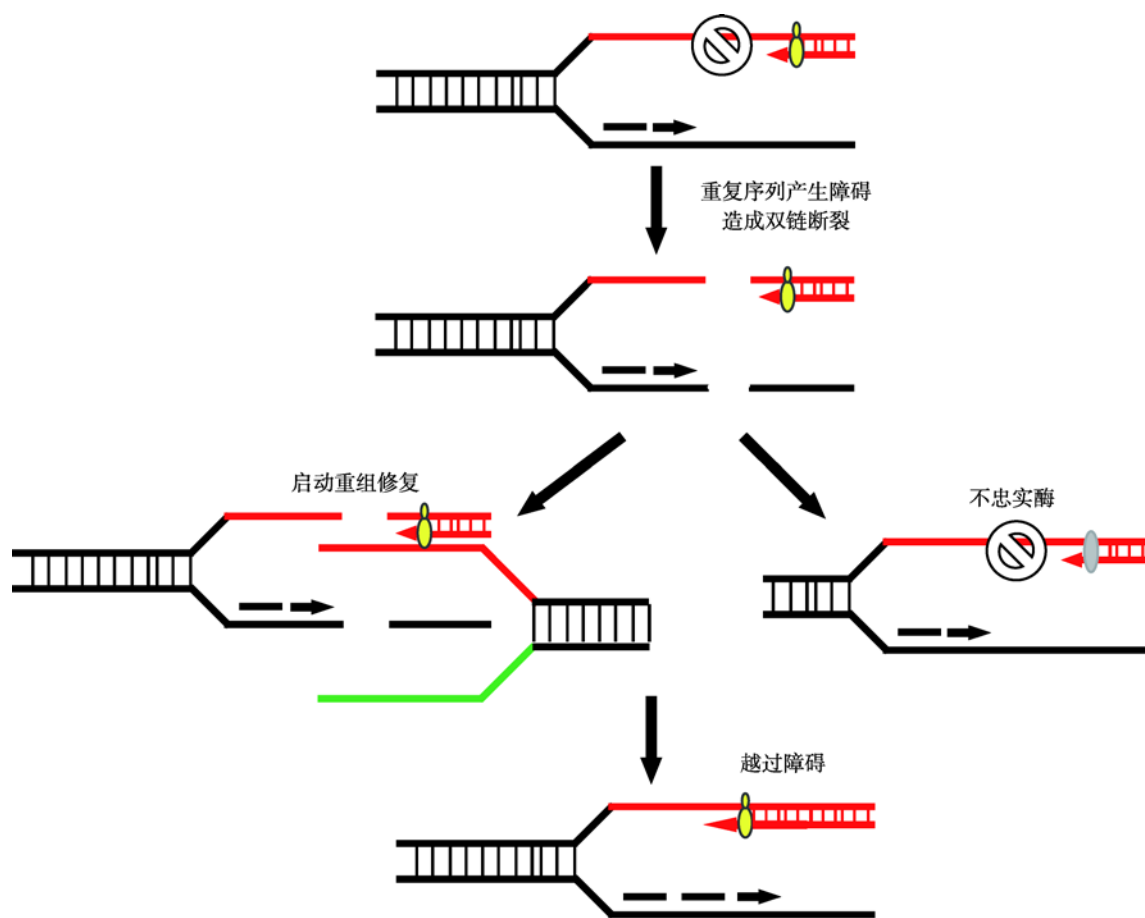


图 5 同源重组不忠实性机制

Fig. 5 Mechanism of unfaithful homologous recombination

重复序列阻碍复制正常进行, 导致双链断裂, 诱发修复机制, 在忠实修复失败后, 以出错率更高的不忠实酶进行修复越过重复序列障碍。

实酶时会更容易跨越障碍, 降低出错率。这一策略对本身有复制转录障碍的重复序列而言是可行的, 但是对生物体复制转录其他正常序列时则是一个灾难。目前本课题组已经利用大肠杆菌的 DNA 聚合酶 IV (DinB) 不忠实复制系统初步证实了这一猜测。简言之, 重复序列本身存在的复制和转录困难, 会造成双链断裂, 同源重组时为了精准修复带来了更高的出错率, 这可能是三核苷酸重复 DNA 序列碱基突变不稳定和重复单位数扩增或缺失不稳定性发生的一种综合机制。

### 3 结语与展望

三核苷酸重复不稳定相关疾病的发病年龄和疾病进展由两种不同机制共同控制, 以等位基因存在

的三核苷酸重复的拷贝数的多少并不直接表现出明确的细胞毒性。但是, 这些重复序列会在某些组织的体细胞中扩增, 当重复序列的重复单位数目超过疾病所需阈值时, 则会通过快速出现的细胞毒性引发疾病。随后, 症候进程主要决定于扩增所能产生的细胞毒性大小。这解释了杂合性和纯合性患者的发病年龄对遗传重复序列数目的依赖性。大量的研究表明, 三核苷酸重复 DNA 序列通常在复制、转录或者修复时形成非 B 型 DNA 二级结构, 然后诱发多种类型的修复。将 DNA 二级结构切除或转而衍生出包括 DNA 双链断裂在内的 DNA 损伤, 对于这种损伤的重组修复可能导致了三核苷酸重复 DNA 的扩增或缺失。显然, 上述过程受制于三核苷酸重复 DNA 序列的种类、在染色体中的位置, 以及所处细胞类型等因素。尽管目前已在包括人类细胞系在

内的多种模式生物中试图重现疾病状态下的三核苷酸重复不稳定性扩增过程,但由于技术限制,有关三核苷酸重复是否在体内形成非 B 型 DNA 构象,以及形成何种类型的非 B 型 DNA 构象很难直接呈现。临床状态下,三核苷酸重复 DNA 序列可能形成的非 B 型 DNA 二级结构如何通过影响 DNA 的复制、转录发生扩增需深入研究。简言之,尽管研究使用的模式生物拥有完整的复制、转录、修复机制,但它们在疾病发生发展过程中的实际作用依然难以确认。特别是疾病相关的三核苷酸重复 DNA 序列所在细胞内的蛋白组差异(功能不同的细胞,其蛋白组不同)、所在染色体上的位置差异是否会直接影响非 B 型 DNA 二级结构的形成以及非 B 型 DNA 二级结构形成后细胞对它们的修复也需进一步研究。到目前为止,虽然三核苷酸重复相关的绝大多数人类疾病尚缺乏有效的治疗手段,但随着对疾病相关的三核苷酸重复 DNA 序列扩增机制的深入了解,相信终将会获得有效治疗的方法。比如,利用类似于 CRISPR/Cas9 的基因编辑技术删除扩增后的三核苷酸重复 DNA 序列,避免这些三核苷酸重复 DNA 序列处在拷贝数扩增阈值之上,通过降低非 B 型 DNA 结构形成的能力,可能会是一种很有前途的治疗的策略。

## 参考文献(References):

- [1] Verkerk AJ, Pieretti M, Sutcliffe JS, Fu YH, Kuhl DP, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang FP, Eussen BE, van Ommen GJ, Blonden LA, Riggins GJ, Chastain JL, Kunst CB, Galjaard H, Caskey CT, Nelson DL, Oostra BA, Warren ST. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell*, 1991, 65(5): 905–914. [DOI]
- [2] La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature*, 1991, 352(6330): 77–79. [DOI]
- [3] Murmann AE, Gao QQ, Putzbach WE, Patel M, Bartom ET, Law CY, Bridgeman B, Chen SQ, McMahon KM, Thaxton CS, Peter ME. Small interfering RNAs based on huntingtin trinucleotide repeats are highly toxic to cancer cells. *EMBO Rep*, 2018, 19(3): e45336. [DOI]
- [4] Swinnen B, Robberecht W, Van Den Bosch L. RNA toxicity in non-coding repeat expansion disorders. *EMBO J*, 2020, 39(1): e101112. [DOI]
- [5] Kourkouta E, Weij R, González-Barriga A, Mulder M, Verheul R, Bosgra S, Groenendaal B, Puoliväli J, Toivanen J, van Deutekom JCT, Datson NA. Suppression of mutant protein expression in SCA3 and SCA1 mice using a CAG repeat-targeting antisense oligonucleotide. *Mol Ther Nucleic Acids*, 2019, 17: 601–614. [DOI]
- [6] Langbehn DR, Stout JC, Gregory S, Mills JA, Durr A, Leavitt BR, Roos RAC, Long JD, Owen G, Johnson HJ, Borowsky B, Craufurd D, Reilmann R, Landwehrmeyer GB, Scallin RI, Tabrizi SJ, TRACK-HD and Track-On HD Groups. Association of CAG repeats with long-term progression in Huntington disease. *JAMA Neurol*, 2019, 76(11): 1375–1385. [DOI]
- [7] Lieberman AP, Shakkottai VG, Albin RL. Polyglutamine repeats in neurodegenerative diseases. *Annu Rev Pathol*, 2019, 14: 1–27. [DOI]
- [8] Pulst SM, Nechiporuk A, Nechiporuk T, Gispert S, Chen XN, Lopes-Cendes I, Pearlman S, Starkman S, Orozco-Diaz G, Lunke A, DeJong P, Rouleau GA, Auburger G, Korenberg JR, Figueroa C, Sahba S. Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat Genet*, 1996, 14(3): 269–276. [DOI]
- [9] Massey TH, Jones L. The central role of DNA damage and repair in CAG repeat diseases. *Dis Model Mech*, 2018, 11(1): dmm031930. [DOI]
- [10] Ghosh R, Tabrizi SJ. Clinical features of Huntington's disease. *Adv Exp Med Biol*, 2018, 1049: 1–28. [DOI]
- [11] Cleary JD, Pattamatta A, Ranum LPW. Repeat-associated non-ATG (RAN) translation. *J Biol Chem*, 2018, 293(42): 16127–16141. [DOI]
- [12] Wilson H, Politis M. Molecular imaging in Huntington's disease. *Int Rev Neurobiol*, 2018, 142: 289–333. [DOI]
- [13] Hu WT, Sun GF, Liu X, Liu YR, Wang S, Peng T, Yang HC, Lu H. Clinical analysis of 10 cases of myotonic muscular dystrophy with brain damage. *J Apopl Nerv Dis*, 2019, 36(9): 819–823.  
胡文涛, 孙桂芳, 刘希, 刘艳茹, 王赏, 彭涛, 杨贺成, 卢宏. 伴有脑部损害的强直性肌营养不良 10 例临床分析. *中风与神经疾病杂志*, 2019, 36(9): 819–823. [DOI]
- [14] Stoyas CA, La Spada AR. The CAG-polyglutamine repeat diseases: a clinical, molecular, genetic, and pathophysiologic nosology. *Handb Clin Neurol*, 2018, 147: 143–170. [DOI]

- [15] Genetic Modifiers of Huntington's Disease (GeM-HD) Consortium. CAG repeat not polyglutamine length determines timing of Huntington's disease onset. *Cell*, 2019, 178(4): 887–900.e14. [DOI]
- [16] Orr HT, Chung M-y, Banfi S, Kwiatkowski TJ, Servadio A, Beaudet AL, McCall AE, Duvick LA, Ranum LP, Zoghbi HY. Expansion of an unstable trinucleotide CAG repeat in spinocerebellar ataxia type 1. *Nat Genet*, 1993, 4(3): 221–226. [DOI]
- [17] Imbert G, Saudou F, Yvert G, Devys D, Trottier Y, Garnier JM, Weber C, Mandel JL, Cancel G, Abbas N, Dürr A, Didierjean O, Stevanin G, Agid Y, Brice A. Cloning of the gene for spinocerebellar ataxia 2 reveals a locus with high sensitivity to expanded CAG/glutamine repeats. *Nat Genet*, 1996, 14(3): 285–291. [DOI]
- [18] Kim YE, Oh KW, Noh MY, Park J, Kim HJ, Park JE, Ki CS, Kim SH. Analysis of ATXN2 trinucleotide repeats in Korean patients with amyotrophic lateral sclerosis. *Neurobiol Aging*, 2018, 67: 201.e5–201.e8. [DOI]
- [19] Kawaguchi Y, Okamoto T, Taniwaki M, Aizawa M, Inoue M, Katayama S, Kawakami H, Nakamura S, Nishimura M, Akiyuchi I, Kimura J, Narumiya S, Kakizuka A. CAG expansions in a novel gene for Machado-Joseph disease at chromosome 14q32.1. *Nat Genet*, 1994, 8(3): 221–228. [DOI]
- [20] Zhuchenko O, Bailey J, Bonnen P, Ashizawa T, Stockton DW, Amos C, Dobyns WB, Subramony SH, Zoghbi HY, Lee CC. Autosomal dominant cerebellar ataxia (SCA6) associated with small polyglutamine expansions in the  $\alpha 1A$ -voltage-dependent calcium channel. *Nat Genet*, 1997, 15(1): 62–69. [DOI]
- [21] David G, Abbas N, Stevanin G, Dürr A, Yvert G, Cancel G, Weber C, Imbert G, Saudou F, Antoniou E, Drabkin H, Gemmill R, Giunti P, Benomar A, Wood N, Ruberg M, Agid Y, Mandel JL, Brice A. Cloning of the SCA7 gene reveals a highly unstable CAG repeat expansion. *Nat Genet*, 1997, 17(1): 65–70. [DOI]
- [22] Holmes SE, O'Hearn EE, McInnis MG, Gorelick-Feldman DA, Kleiderlein JJ, Callahan C, Kwak NG, Ingersoll-Ashworth RG, Sherr M, Sumner AJ, Sharp AH, Ananth U, Seltzer WK, Boss MA, Viera-Saecker AM, Epplen JT, Riess O, Ross CA, Margolis RL. Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. *Nat Genet*, 1999, 23(4): 391–392. [DOI]
- [23] Nakamura K, Jeong SY, Uchihara T, Anno M, Nagashima K, Nagashima T, Ikeda S, Tsuji S, Kanazawa I. SCA17, a novel autosomal dominant cerebellar ataxia caused by an expanded polyglutamine in TATA-binding protein. *Hum Mol Genet*, 2001, 10(14): 1441–1448. [DOI]
- [24] Gonzalez-Alegre P. Recent advances in molecular therapies for neurological disease: triplet repeat disorders. *Hum Mol Genet*, 2019, 28(R1): R80–R87. [DOI]
- [25] La Spada AR, Wilson EM, Lubahn DB, Harding AE, Fischbeck KH. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature*, 1991, 352(6330): 77–79. [DOI]
- [26] Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, Schlessinger D, Sutherland GR, Richards RI. Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science*, 1991, 252(5013): 1711–1714. [DOI]
- [27] Linsalata AE, He F, Malik AM, Glineburg MR, Green KM, Natla S, Flores BN, Krans A, Archbold HC, Fedak SJ, Barmada SJ, Todd PK. DDX3X and specific initiation factors modulate FMR1 repeat-associated non-AUG-initiated translation. *EMBO Rep*, 2019, 20(9): e47498. [DOI]
- [28] Knight SJ, Flannery AV, Hirst MC, Campbell L, Christodoulou Z, Phelps SR, Pointon J, Middleton-Price HR, Barnicoat A, Pembrey ME, Holland J, Oostra BA, Bobrow M, Davies KE. Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. *Cell*, 1993, 74(1): 127–134. [DOI]
- [29] Campuzano V, Montermini L, Moltò MD, Pianese L, Cossée M, Cavalcanti F, Monros E, Rodius F, Duclos F, Monticelli A, Zara F, Cañizares J, Koutnikova H, Bidichandani SI, Gellera C, Brice A, Trouillas P, De Michele G, Filla A, De Frutos R, Palau F, Patel PI, Di Donato S, Mandel JL, Coccozza S, Koenig M, Pandolfo M. Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. *Science*, 1996, 271(5254): 1423–1427. [DOI]
- [30] Harley HG, Rundle SA, Reardon W, Myring J, Crow S, Brook JD, Harper PS, Shaw DJ. Unstable DNA sequence in myotonic dystrophy. *Lancet*, 1992, 339(8802): 1125–1128. [DOI]
- [31] Koob MD, Moseley ML, Schut LJ, Benzow KA, Bird TD, Day JW, Ranum LP. An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). *Nat Genet*, 1999, 21(4): 379–384. [DOI]
- [32] Lee JM, Zhang J, Su AI, Walker JR, Wiltshire T, Kang K, Dragileva E, Gillis T, Lopez ET, Boily MJ, Cyr M,

- Kohane I, Gusella JF, MacDonald ME, Wheeler VC. A novel approach to investigate tissue-specific trinucleotide repeat instability. *BMC Syst Biol*, 2010, 4: 29. [DOI]
- [33] Helma R, Bažantová P, Petr M, Adámik M, Renčiuk D, Tichý V, Pastuchová A, Soldánová Z, Pečinka P, Bowater RP, Fojta M, Brázdová M. p53 binds preferentially to non-B DNA structures formed by the pyrimidine-rich strands of GAA•TTC trinucleotide repeats associated with Friedreich's ataxia. *Molecules*, 2019, 24(11): 2078. [DOI]
- [34] Ooi J, Langley SR, Xu XH, Utami KH, Sim B, Huang YH, Harmston NP, Tay YL, Ziaei A, Zeng RZ, Low D, Aminkeng F, Sobota RM, Ginhoux F, Petretto E, Pouladi MA. Unbiased profiling of isogenic Huntington disease hPSC-derived CNS and peripheral cells reveals strong cell-type specificity of CAG length effects. *Cell Rep*, 2019, 26(9): 2494–2508.e7. [DOI]
- [35] Morton AJ, Skillings EA, Wood NI, Zheng Z. Antagonistic pleiotropy in mice carrying a CAG repeat expansion in the range causing Huntington's disease. *Sci Rep*, 2019, 9(1): 37. [DOI]
- [36] Jones L, Houlden H, Tabrizi SJ. DNA repair in the trinucleotide repeat disorders. *Lancet Neurol*, 2017, 16(1): 88–96. [DOI]
- [37] Nolin SL, Glicksman A, Tortora N, Allen E, Macpherson J, Mila M, Vianna-Morgante AM, Sherman SL, Dobkin C, Latham GJ, Hadd AG. Expansions and contractions of the FMR1 CGG repeat in 5,508 transmissions of normal, intermediate, and premutation alleles. *Am J Med Genet A*, 2019, 179(7): 1148–1156. [DOI]
- [38] Khan E, Biswas S, Mishra SK, Mishra R, Samanta S, Mishra A, Tawani A, Kumar A. Rationally designed small molecules targeting toxic CAG repeat RNA that causes Huntington's disease (HD) and spinocerebellar ataxia (SCAs). *Biochimie*, 2019, 163: 21–32. [DOI]
- [39] Santoro M, Fontana L, Maiorca F, Centofanti F, Massa R, Silvestri G, Novelli G, Botta A. Expanded [CCTG]<sub>n</sub> repetitions are not associated with abnormal methylation at the CNBP locus in myotonic dystrophy type 2 (DM2) patients. *Biochim Biophys Acta Mol Basis Dis*, 2018, 1864(3): 917–924. [DOI]
- [40] Zafarullah M, Tassone F. Molecular biomarkers in fragile X syndrome. *Brain Sci*, 2019, 9(5): 96. [DOI]
- [41] Chen HM. Clinical and molecular biology research of myotonic muscular dystrophy. *Zhejiang Med J*, 2019, 41(5): 407–408, 418.
- 陈慧敏. 强直性肌营养不良临床与分子生物学研究. 浙江医学, 2019, 41(5): 407–408, 418. [DOI]
- [42] Chen YY, Hao Y, Zhang J, Zhang X, Xie KM, Ding M, Gu WH. Capillary electrophoresis fragment analysis and clone sequencing in detection of dynamic mutations of spinocerebellar ataxia. *Chin J Contemp Neurol Neurosurg*, 2018, 18(3): 192–197.
- 陈园园, 郝莹, 张瑾, 张鑫, 谢坤铭, 丁铭, 顾卫红. 基于毛细管电泳的片段分析和克隆测序在脊髓小脑共济失调动态突变检测中的应用研究. 中国现代神经疾病杂志, 2018, 18(3): 192–197. [DOI]
- [43] Westenberger A, Reyes CJ, Saranza G, Dobricic V, Hanssen H, Domingo A, Laabs BH, Schaake S, Pozojevic J, Rakovic A, Grütz K, Begemann K, Walter U, Dressler D, Bauer P, Rolfs A, Münchau A, Kaiser FJ, Ozelius LJ, Jamora RD, Rosales RL, Diesta CCE, Lohmann K, König IR, Brüggemann N, Klein C. A hexanucleotide repeat modifies expressivity of X-linked dystonia parkinsonism. *Ann Neurol*, 2019, 85(6): 812–822. [DOI]
- [44] Ohshima K, Montermini L, Wells RD, Pandolfo M. Inhibitory effects of expanded GAA.TTC triplet repeats from intron I of the Friedreich ataxia gene on transcription and replication *in vivo*. *J Biol Chem*, 1998, 273(23): 14588–14595. [DOI]
- [45] Grabczyk E, Mancuso M, Sammarco MC. A persistent RNA:DNA hybrid formed by transcription of the Friedreich ataxia triplet repeat in live bacteria, and by T7 RNAP *in vitro*. *Nucleic Acids Res*, 2007, 35(16): 5351–5359. [DOI]
- [46] Jain A, Vale RD. RNA phase transitions in repeat expansion disorders. *Nature*, 2017, 546(7657): 243–247. [DOI]
- [47] Burguete AS, Almeida S, Gao FB, Kalb R, Akins MR, Bonini NM. GGGGCC microsatellite RNA is neuritically localized, induces branching defects, and perturbs transport granule function. *eLife*, 2015, 4: e08881. [DOI]
- [48] Wang XY, Goodrich KJ, Conlon EG, Gao JC, Erbse AH, Manley JL, Cech TR. C9orf72 and triplet repeat disorder RNAs: G-quadruplex formation, binding to PRC2 and implications for disease mechanisms. *RNA*, 2019, 25(8): 935–947. [DOI]
- [49] Perutz MF, Johnson T, Suzuki M, Finch JT. Glutamine repeats as polar zippers: their possible role in inherited neurodegenerative diseases. *Proc Natl Acad Sci USA*, 1994, 91(12): 5355–5358. [DOI]
- [50] Treangen TJ, Salzberg SL. Repetitive DNA and next-generation sequencing: computational challenges and

- solutions. *Nat Rev Genet*, 2011, 13(1): 36–46. [DOI]
- [51] Gemayel R, Vences MD, Legendre M, Verstrepen KJ. Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annu Rev Genet*, 2010, 44: 445–477. [DOI]
- [52] Sznajder LJ, Swanson MS. Short tandem repeat expansions and RNA-mediated pathogenesis in myotonic dystrophy. *Int J Mol Sci*, 2019, 20(13): 3365. [DOI]
- [53] Gellon L, Kaushal S, Cebrián J, Lahiri M, Mirkin SM, Freudenreich CH. Mrc1 and Tof1 prevent fragility and instability at long CAG repeats by their fork stabilizing function. *Nucleic Acids Res*, 2019, 47(2): 794–805. [DOI]
- [54] Mouro Pinto R, Arning L, Giordano JV, Razghandi P, Andrew MA, Gillis T, Correia K, Mysore JS, Grote Urtubey DM, Parwez CR, von Hein SM, Clark HB, Nguyen HP, Förster E, Beller A, Jayadaev S, Keene CD, Bird TD, Lucente D, Vonsattel JP, Orr H, Saft C, Petrasch-Parwez E, Wheeler VC. Patterns of CAG repeat instability in the central nervous system and periphery in Huntington's disease and in spinocerebellar ataxia type 1. *Hum Mol Genet*, 2020, 29(15): 2551–2567. [DOI]
- [55] Pan XF, Xiao P, Li HQ, Zhao DX, Duan F. The Gratuitous repair on undamaged DNA misfold. In: DNA Repair. Kruman I, Ed. Tech Press, 2011, DOI: 10.5772/22441. [DOI]
- [56] Kunkel TA. Nucleotide repeats. Slippery DNA and diseases. *Nature*, 1993, 365(6443): 207–208. [DOI]
- [57] Mosbach V, Poggi L, Richard GF. Trinucleotide repeat instability during double-strand break repair: from mechanisms to gene therapy. *Curr Genet*, 2019, 65(1): 17–28. [DOI]
- [58] Frank-Kamenetskii MD, Mirkin SM. TRIPLEX DNA STRUCTURES. *Annu Rev Biochem*, 1995, 64(1): 65–95. [DOI]
- [59] Faruqi AF, Datta HJ, Carroll D, Seidman MM, Glazer PM. Triple-helix formation induces recombination in mammalian cells via a nucleotide excision repair-dependent pathway. *Mol Cell Biol*, 2000, 20(3): 990–1000. [DOI]
- [60] Tian Y, Wang JL, Huang W, Zeng S, Jiao B, Liu Z, Chen Z, Li YJ, Wang Y, Min HX, Wang XJ, You Y, Zhang RX, Chen XY, Yi F, Zhou YF, Long HY, Zhou CJ, Hou X, Wang JP, Xie B, Liang F, Yang ZY, Sun QY, Allen EG, Shafik AM, Kong HE, Guo JF, Yan XX, Hu ZM, Xia K, Jiang H, Xu HW, Duan RH, Jin P, Tang BS, Shen L. Expansion of human-specific GGC repeat in neuronal intranuclear inclusion disease-related disorders. *Am J Hum Genet*, 2019, 105(1): 166–176. [DOI]
- [61] Béna F, Gimelli S, Miglia vacca E, Sharp AJ. Recurrent 14q32.2 microdeletion mediated by expanded TGG repeats. *Hum Mol Genet*, 2010, 19 (10): 1967–1973. [DOI]
- [62] Pan XF, Leach D. The roles of *mutS*, *sbcCD* and *recA* in the propagation of TGG repeat in *Escherichia coli*. *Nucleic Acids Res*, 2000, 28: 3178–3184. [DOI]
- [63] Gacy AM, Goellner G, Juranić N, Macura S, McMurray CT. Trinucleotide repeats that expand in human disease form hairpin structures *in vitro*. *Cell*, 1995, 81(4): 533–540. [DOI]
- [64] Rajan-Babu IS, Chong SS. Triplet-repeat primed PCR and capillary electrophoresis for characterizing the fragile X mental retardation 1 CGG repeat hyperexpansions. *Methods Mol Biol*, 2019, 1972: 199–210. [DOI]
- [65] Murchie AI, Lilley DM. The mechanism of cruciform formation in supercoiled DNA: initial opening of central basepairs in salt-dependent extrusion. *Nucleic Acids Res*, 1987, 15(23): 9641–9654. [DOI]
- [66] Murmann AE, Yu JD, Opal P, Peter ME. Trinucleotide repeat expansion diseases, RNAi, and cancer. *Trends Cancer*, 2018, 4(10): 684–700. [DOI]
- [67] Zeitler B, Froelich S, Marlen K, Shivak DA, Yu Q, Li D, Pearl JR, Miller JC, Zhang L, Paschon DE, Hinkley SJ, Ankoudinova I, Lam S, Guschin D, Kopan L, Cherone JM, Nguyen HOB, Qiao GJ, Ataei Y, Mendel MC, Amora R, Surosky R, Laganieri J, Vu BJ, Narayanan A, Sedaghat Y, Tillack K, Thiede C, Gärtner A, Kwak S, Bard J, Mrzljak L, Park L, Heikkinen T, Lehtimäki KK, Svedberg MM, Häggkvist J, Tari L, Tóth M, Varrone A, Halldin C, Kudwa AE, Ramboz S, Day M, Kondapalli J, Surmeier DJ, Urnov FD, Gregory PD, Rebar EJ, Muñoz-Sanjuán I, Zhang HS. Allele-selective transcriptional repression of mutant HTT for the treatment of Huntington's disease. *Nat Med*, 2019, 25(7): 1131–1142. [DOI]
- [68] Raaijmakers RHL, Ripken L, Ausems CRM, Wansink DG. CRISPR/Cas applications in myotonic dystrophy: expanding opportunities. *Int J Mol Sci*, 2019, 20(15): 3689. [DOI]
- [69] Dabrowska M, Juzwa W, Krzyzosiak WJ, Olejniczak M. Precise excision of the CAG tract from the Huntingtin gene by Cas9 nickases. *Front Neurosci*, 2018, 12: 75. [DOI]
- [70] Dastidar S, Ardui S, Singh K, Majumdar D, Nair N, Fu YF, Reyon D, Samara E, Gerli MFM, Klein AF, De Schrijver W, Tipanee J, Seneca S, Tulalamba W, Wang H, Chai YC, In't Veld P, Furling D, Tedesco FS, Vermeesch JR, Joung JK, Chuah MK, VandenDriessche T. Efficient CRISPR/

- Cas9-mediated editing of trinucleotide repeat expansion in myotonic dystrophy patient-derived iPS and myogenic cells. *Nucleic Acids Res*, 2018, 46(16): 8275–8298. [DOI]
- [71] Mirkin SM, Smirnova EV. Positioned to expand. *Nat Genet*, 2002, 31(1): 5–6. [DOI]
- [72] Pan XF. Mechanism of trinucleotide repeats instabilities: the necessities of repeat non-B secondary structure formation and the roles of cellular trans-acting factors. *Acta Genet Sin*, 2006, 33(1): 1–11. [DOI]
- [73] Wen YL, Lv KN, Xun XK, Zhang X, Ding L, Pan XF. The role and mechanism of annealing helicase SMARCAL1 in maintaining genome stability. *Hereditas(Beijing)*, 2019, 41(12): 1084–1098.  
文雅蕾, 吕柯舜, 徐小康, 张欣, 丁良, 潘学峰. 退火解旋酶 SMARCAL1 在维持基因组稳定中的作用与机制. *遗传*, 2019, 41(12): 1084–1098. [DOI]
- [74] Pan XF, Ding YF, Shi LF. The roles of SbcCD and RNaseE in the transcription of GAA × TTC repeats in *Escherichia coli*. *DNA Repair*, 2009, 8(11): 1321–1327. [DOI]
- [75] Pan XF, Liao YH, Liu YM, Chang P, Liao LN, Yang L, Li HQ. Transcription of AAT•ATT triplet repeats in *Escherichia coli* is silenced by H-NS and IS1E transposition. *PLoS One*, 2010, 5(12): e14271. [DOI]
- [76] Wei JP, Pan XF, Li HQ, Duan F. Distribution and evolution of simple repeats in the mtDNA D-loop in mammalian. *Hereditas(Beijing)*, 2011, 33(1): 67–74.  
危金普, 潘学峰, 李红权, 段斐. 简单重复 dna 序列在哺乳动物 mtDNA d-loop 区的分布及进化特征. *遗传*, 2011, 33(1): 67–74. [DOI]
- [77] Kozłowski P, de Mezer M, Krzyżosiak WJ. Trinucleotide repeats in human genome and exome. *Nucleic Acids Res*, 2010, 38(12): 4027–4039. [DOI]
- [78] Wright GEB, Collins JA, Kay C, McDonald C, Dolzhenko E, Xia QW, Bečanović K, Drögemöller BI, Semaka A, Nguyen CM, Trost B, Richards F, Bijlsma EK, Squitieri F, Ross CJD, Scherer SW, Eberle MA, Yuen RKC, Hayden MR. Length of uninterrupted CAG, independent of polyglutamine size, results in increased somatic instability, hastening onset of Huntington disease. *Am J Hum Genet*, 2019, 104(6): 1116–1126. [DOI]
- [79] Tojima M, Murakami G, Hikawa R, Yamakado H, Yamashita H, Takahashi R, Matsui M. Homozygous 31 trinucleotide repeats in the SCA2 allele are pathogenic for cerebellar ataxia. *Neurol Genet*, 2018, 4(6): e283. [DOI]

(责任编辑: 何淑君)