

论著·基础研究

葡萄糖-6-磷酸脱氢酶基因 4 种突变及蛋白结构分析

朱之星^{1,2}, 纪伟³, 顾坚磊^{1,4}, 吕晖^{1,2,4}, 田国力³

1. 上海交通大学附属儿童医院, 上海市儿童医院生物医学信息研究中心, 上海 200040; 2. 上海交通大学附属儿童医院儿童精准医学大数据工程技术研究中心, 上海 200040; 3. 上海交通大学附属儿童医院新生儿筛查中心, 上海 200040; 4. 上海交通大学生命科学技术学院, 上海交通大学-耶鲁大学生物统计与数据科学联合中心, 上海 200240

[摘要] **目的**·研究葡萄糖-6-磷酸脱氢酶 (glucose-6-phosphate dehydrogenase, *G6PD*) 基因突变对酶蛋白结构及功能的影响, 了解 *G6PD* 突变与酶学结构、功能的关系。**方法**·回顾性分析 2014—2017 年上海市儿童医院 205 103 例新生儿 *G6PD* 缺乏症筛查记录, Wilcoxon 检验分析酶活性与突变基因类型之间的关系。采用 Psipred、SOPMA、JPred4 预测蛋白的二级结构, SWISS-MODEL 预测氨基酸链空间结构, Mupro、SDM、CUPSAT、mSCM、DUET、Dynamut 预测蛋白突变前后稳定性变化, PROVEAN 预测对蛋白功能的影响, PyMOL 和 LigPlot⁺ 修饰蛋白空间结构预测图。**结果**·230 例样本酶活性检测阳性, 对其中 121 例进行基因检测, 共检出 8 种突变, 除 c.95A>G、c.1376G>T、c.1388G>A 3 种常见突变, c.1024C>T 也表现出高发生率。4 种突变均导致蛋白结构改变, 稳定性降低, 对功能有不良影响; 此外, c.1388G>A 突变导致其远离甘油配体结合区。**结论**·4 种突变均通过改变蛋白结构进而降低蛋白的稳定性, 其中 c.1388G>A 突变还会影响蛋白与底物的结合能力。

[关键词] 葡萄糖-6-磷酸脱氢酶 (*G6PD*) 缺乏症; 基因突变; 酶活性; 酶蛋白结构; 蛋白稳定性

[DOI] 10.3969/j.issn.1674-8115.2020.12.001 **[中图分类号]** R722.1 **[文献标志码]** A

Analysis of four mutations and protein structure of glucose-6-phosphate dehydrogenase gene

ZHU Zhi-xing^{1,2}, JI Wei³, GU Jian-lei^{1,4}, LÜ Hui^{1,2,4}, TIAN Guo-li³

1. Center for Biomedical Informatics, Shanghai Children's Hospital, Shanghai Jiao Tong University, Shanghai 200040, China; 2. Shanghai Engineering Research Center for Big Data in Pediatric Precision Medicine, Shanghai 200040, China; 3. Newborn Screening Center, Shanghai Children's Hospital, Shanghai 200040, China; 4. SJTU-Yale Joint Center for Biostatistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

[Abstract] **Objective**·To analyze the effect of glucose-6-phosphate dehydrogenase (*G6PD*) gene mutations on the structure and function of enzymes. **Methods**·A retrospective analysis of *G6PD* screening records of 205 103 neonates collected in the Shanghai Children's Hospital from 2014 to 2017 was performed. Wilcoxon test was used to analyze the enzyme activity and mutant genes. The software tools Psipred, SOPMA, and JPred4 were used to predict the secondary structure of the protein, and the 3D structure of *G6PD* was predicted based on the amino acid chain by SWISS-MODEL and modified by PyMOL and LigPlot⁺. Six different analysis software programs, including Mupro, SDM, CUPSAT, mSCM, DUET, and Dynamut, were utilized to compare protein stability of the wild-type with the mutant forms. PROVEAN was used to analyze the effect of amino acid changes on the enzyme. **Results**·Two hundred and thirty samples were positive for enzyme deficiency, of which 121 positive samples were genetically tested and eight mutations were identified. Three common mutations c.95A>G, c.1376G>T, and c.1388G>A were present, and c.1024C>T mutation was identified in 14 samples. These four mutations induced a change in protein structure, reduced protein stability, and had adverse effects on function. In addition, c.1388G was adjacent to the glycerol ligand binding region in the wild type structure, whereas the c.1388G>A mutation caused this residue further away from it. **Conclusion**·Four nonsynonymous mutations (c.95A>G, c.1376G>T, c.1388G>A and c.1024C>T) reduce the stability of the *G6PD* enzyme by changing the structure of the protein, in which the c.1388G>A mutation also affects the binding ability of the protein to the substrate.

[Key words] glucose-6-phosphate dehydrogenase (*G6PD*) deficiency; gene mutation; enzyme activity; protein structure; protein stability

[基金项目] 国家重点研发计划 (2018YFC0910500); 上海市科学技术委员会专项基金 (17DZ2251200); 上海市卫生与计划生育委员会研究课题 (2018ZHYL0223); 上海交通大学医工交叉项目 (ZH2018QNA30); 上海市科学技术委员会科研项目 (18441905100); 上海市重中之重临床重点专科建设项目 (2017ZZ02019)。

[作者简介] 朱之星 (1996—), 女, 硕士生, 电子信箱: 981164372@qq.com。

[通信作者] 田国力, 电子信箱: Tiangl@shchildren.com.cn。

[Funding Information] National Key R&D Program of China (2018YFC0910500); Foundation of Science and Technology Commission of Shanghai Municipality (17DZ2251200); Foundation of Shanghai Municipal Commission of Health and Family Planning (2018ZHYL0223); Medical Conversion Cross Fund of Shanghai Jiao Tong University (ZH2018QNA30); Action Plan for Scientific and Technological Innovation in Shanghai (18441905100); Key Disciplines of Top Priority in Shanghai (2017ZZ02019)。

[Corresponding Author] TIAN Guo-li, E-mail: Tiangl@shchildren.com.cn。

葡萄糖-6-磷酸脱氢酶 (glucose-6-phosphate dehydrogenase, G6PD) 缺乏症是常见的一种先天性酶缺乏遗传病, 由定位于 Xq28 上的 *G6PD* 基因 (OMIM:305900) 突变导致 G6PD 酶活性降低或缺乏。该基因编码为一个由 515 个氨基酸组成的酶蛋白亚基, 其同源二聚体具有酶活性, 并能进一步产生酶促活性的四聚体蛋白^[1-2]。G6PD 酶活性的降低会引起红细胞不能抵抗氧化损伤而遭受破坏, 从而产生急性溶血, 部分可引起新生儿期高胆红素脑病, 导致智力低下甚至死亡^[2]。

G6PD 缺乏症作为我国新生儿疾病筛查主要疾病之一, 主要通过检测滤纸干血片血斑样品的酶活性, 并结合后续生化或基因检测来明确诊断。随着测序技术的发展, *G6PD* 基因的突变位点被广泛研究^[3-5]。1993 年杜传书团队^[5]利用 APCR 单链产物以双脱氧链终止法对该基因进行 DNA 序列测定, 首次在中国人群中发现了 6 种 *G6PD* 基因突变位点。如今, G6PD 缺乏症已知突变位点已达近 200 种^[3], 其中在中国人群中被报道的有 40 多种^[4, 6-7]。多种基因检测方法不断发展^[6, 8-10], 其中多色探针荧光 PCR 熔解曲线法^[10]因其检测速度快、简便、低成本等特点而被广泛使用。

然而, 基因突变导致 G6PD 酶活性降低或缺乏的分子机制尚未明确。有研究^[11]认为基因突变导致了 G6PD 酶蛋白稳定性的降低, 也有研究^[12-13]提示突变影响了辅酶 NADP⁺ 及底物与酶蛋白结合能力, 从而影响其正常的生物学功能。为了更好地了解不同突变从哪些方面影响了酶活性, 本研究分析了新生儿筛查 G6PD 缺乏症人群的酶活性及基因突变分布情况, 从分子、结构及功能的角度将 4 种常见突变, 即 c.95A>G、c.1376G>T、c.1388G>A、c.1024C>T, 与野生型 G6PD 酶进行比较, 并对这些突变蛋白结构及功能进行预测分析, 揭示 G6PD 突变与酶学结构、功能的关系, 以期对遗传咨询提供帮助。

1 对象与方法

1.1 研究对象

纳入上海交通大学附属儿童医院 2014—2017 年收集的 205 103 例新生儿 G6PD 缺乏症筛查记录 (为保护患者隐私, 在数据进行分析处理前, 患者的个人可识别信息均已被剔除)。研究获上海交通大学附属儿童医院伦理委员会批准 (2019R075-E01)。

1.2 方法

新生儿出生 72 h 后, 针刺足跟采血于滤纸上, 制成滤纸干血片标本。通过新生儿 G6PD 缺乏症筛查及分析并

存于医院新生儿筛查中心系统。

1.2.1 G6PD 酶活性检测 采用新生儿 G6PD 测定试剂 (PerkinElmer, 美国) 及 VICTOR™ D 荧光计 (Wallac Oy, 芬兰), 遵照实验室建立的标准操作流程检测: 在室温下使血斑样品和含有 G6P、NADP⁺ 的底物试剂反应 30 min, 可通过添加硫酸铜减缓反应速度。荧光计设定激发波长为 355 nm 和 460 nm, 进行荧光强度测量 (荧光强度与酶活性呈正比), 由机载软件处理得到 G6PD 酶活性数据。设定酶活性的最低标准, 以 U/g (每克血红蛋白中) 为酶活性单位, 活性低于 2.0 U/g 判定为初筛阳性。

1.2.2 16 个中国人群常见 *G6PD* 突变位点的 PCR 扩增 采用荧光 PCR 熔解曲线法, 取直径为 3 mm 的 3 个滤纸干血片, 参照全血 -824 核酸提取 Mini 试剂盒 (厦门致善生物科技有限公司, 中国) 说明书, 通过 Lab-aid 824 全自动核酸提取仪 (厦门致善生物科技有限公司, 中国) 提取人基因组 DNA, -20 °C 保存备用。使用 SLAN-96S 荧光定量 PCR 仪 (厦门致善生物科技有限公司, 中国) 对靶标基因进行 PCR 扩增。

反应的体系均为: DNA 模板 2.6 μL, 2×PCR 混合体系 14.0 μL, 上、下游引物 (20 μmol/L) 各 0.7 μL, 用双蒸水补足体积至 25.0 μL。反应程序为: 50 °C UnG 酶处理 2 min; 95 °C 预变性 10 min; 95 °C 变性 15 s, 65 °C ~ 56 °C 退火 15 s (每个循环下降 1 °C), 76 °C 延伸 20 s 以上过程进行 10 个循环; 随后 95 °C 变性 15 s, 55 °C 退火 15 s, 76 °C 延伸 20 s 过程进行 50 个循环。

采用 *G6PD* 基因突变检测试剂盒, 根据靶标与探针杂交产物熔点的差异来检测突变位点, 可以定性检测出 16 种中国人群常见的 *G6PD* 基因突变, 分别为: c.95A>G、c.383T>C、c.392G>T、c.487G>A、c.493A>G、c.517T>C、c.519C>T、c.592C>T、c.871G>A、c.1004C>A、c.1024C>T、c.1360C>T、c.1376G>T、c.1381G>A、c.1387C>T、c.1388G>A。

1.3 酶活性及突变情况统计学分析

应用 R3.6.3 软件对数据结果进行统计分析, 定量数据以 $M(Q_1, Q_3)$ 表示, 定性数据以 $n(\%)$ 表示。Wilcoxon 秩和检验分析酶活性与突变基因类型之间的关系, 在显著性水平 $\alpha=0.05$ 的条件下, 对检验假设 “ $H_0: \mu_1=\mu_2$; $H_1: \mu_1 \neq \mu_2$ ” 进行检验, 其中 μ_1 、 μ_2 分别为 2 个总体的均值, 以秩和 W 作为统计量, $P<0.05$ 认为有统计学意义。所有检验均为双侧检验。随后对不同突变的酶活性采用 G6PD 缺乏症严重等级分型^[14-15]进行描述性统计分析, 以比较不同突变对患者造成的影响是否具有差异。

1.4 酶蛋白空间结构预测

G6PD 基因具有 13 个外显子, 除去 1 号外显子不编码蛋白, 根据 cDNA 翻译的氨基酸序列, 包含有 515 个氨基酸, 此外 G6PD 的活性蛋白在人体中以同源二聚体和同源四聚体形式出现^[2]。

G6PD 的 cDNA 参考序列及氨基酸序列来源于 Ensembl 基因组数据库 (ENST00000393564.6), Psipred^[16-17]、SOPMA^[18]、JPred4^[19] 被用来预测蛋白的二级结构。使用 SWISS-MODEL^[20] 预测氨基酸链的空间结构, 并参比 PDB^[21-22] 蛋白结构数据库 6e08 和 6e07.1.A 分别构建蛋白单体及二聚体。

为保证预测可靠性, 利用 Mupro^[23]、SDM^[24-25]、CUPSAT^[26]、mSCM^[27]、DUET^[28]、Dynamut^[29] 6 种基于结构、图形特征、统计势能等蛋白稳定性分析工具同时对突变前后蛋白的稳定性进行预测, 以突变后蛋白折叠自由能的变化量 $\Delta \Delta G$ 作为统计量。若 $\Delta \Delta G > 0$ ^[23-29], 则表示突变后蛋白结构的稳定性增加, 反之表示蛋白结构稳定性降低。蛋白结构及预测图使用 PyMOL 完成, LigPlot⁺^[30] 用来构建配体与蛋白相关作用预测图。

采用 PROVEAN^[31] 工具预测氨基酸的改变对蛋白功能的影响。为保证更高的有害突变灵敏度, 设置阈值为 -1.3 ^[31], PROVEAN 得分 ≤ -1.3 的变量被认为是“有害的”, 得分 > -3 的变量被认为是“中性的”。

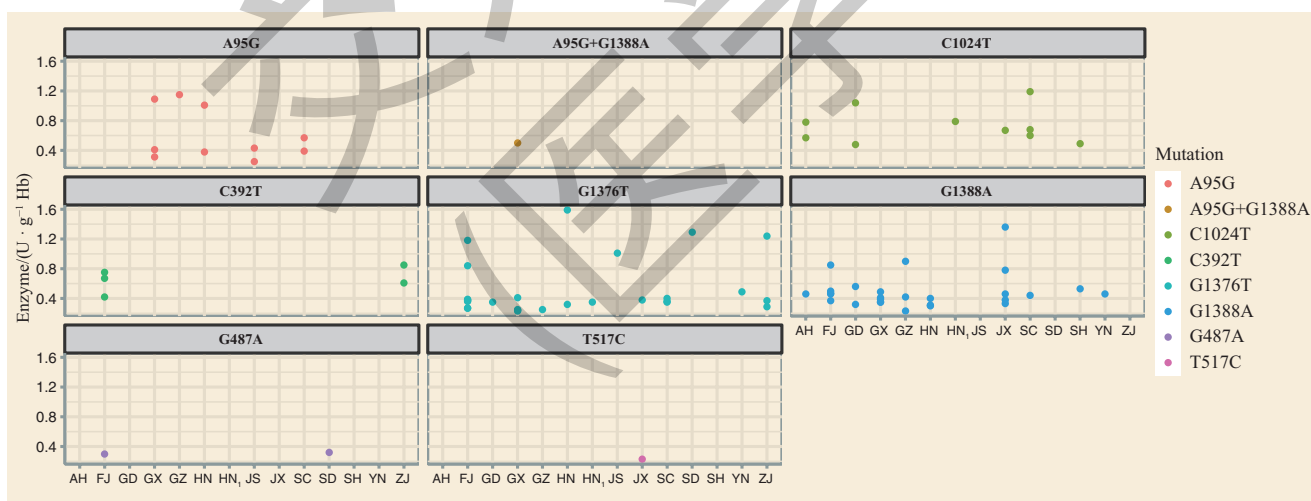
2 结果

2.1 G6PD 酶活性检测情况

2014—2017 年新生儿筛查的 205 103 例 G6PD 活性检测血样, 男婴 106 654 例 (52%), 女婴 98 449 例 (48%), 中位采血日龄 3 d (3 d, 5 d), 中位胎龄 39^{+3} 周 (38^{+4} 周, 40^{+1} 周), 中位孕周体质量 3 350 g (3 080 g, 3 620 g)。通过荧光分析法测定 G6PD 酶活性, 确诊 G6PD 缺乏症 230 例, 其中男 211 例 (占 92%), 女 19 例 (占 8%)。

2.2 基因突变检测分析

采用多色探针荧光 PCR 溶解曲线法对 121 例 G6PD 酶缺陷样本进行检测分析, 检测出 8 种点突变, 其中 7 种单碱基突变, 分别是 c.95A>G、c.1024C>T、c.392C>T、c.1376G>T、c.1388G>A、c.487G>A、c.517T>C, 1 种复合点突变为 c.95A>G+c.1388G>A。来自全国 14 个行政区不同突变的酶活性结果如图 1 所示。本研究人群除 c.95A>G、c.1376G>T、c.1388G>A 3 种常见突变, c.1024C>T 也表现出高发生率, 4 种突变占检测总人数的 91%, 且前 3 种突变酶活性主要分布在 0.2 ~ 0.6 U/g, c.1024C>T 突变则主要分布在 0.4 ~ 0.8 U/g。



Note: AH—Anhui; FJ—Fujian; GD—Guangdong; GX—Guangxi; GZ—Guizhou; HN—Hunan; HN₁—Henan; JS—Jiangsu; JX—Jiangxi; SC—Sichuan; SD—Shandong; SH—Shanghai; YN—Yunnan; ZJ—Zhejiang.

图 1 不同突变在各地的分布情况以及酶活性缺乏情况

Fig 1 Distribution of different mutations in different administrative regions and enzyme activities

采用 Wilcoxon 秩和检验估计 4 种主要突变与酶活性之间的关系。表 1 显示, c.1376G>T、c.1388G>A、c.95A>G 突变的酶活性两两之间差异均无统计学意义, 3

种突变与 c.1024C>T 突变的酶活性比较, 差异均有统计学意义, 结果表明部分基因突变位点与酶活性有关。

对不同突变进行分组, 根据 G6PD 缺乏症严重等级

表 1 4 种主要突变位点两两间的 Wilcoxon 检验结果
Tab 1 Wilcoxon test results between four major mutation positions

Mutation position	<i>W</i>	<i>P</i> value
c.95A>G c.1024C>T	151.5	0.045
c.95A>G c.1388G>A	305.5	0.978
c.95A>G c.1376G>T	242.5	0.232
c.1376G>T c.1388G>A	1 037.5	0.068
c.1376G>T c.1024C>T	109.5	0.001
c.1024C>T c.1388G>A	114.5	0.001

表 2 突变位点与酶活性缺乏严重程度情况
Tab 2 Severity of mutation positions and enzyme activities

Mutation position	Exon	Amino acid change	Class I/n	Class II/n	Class III/n	Class IV/n	Class V/n	Total/n
c.95A>G	exon2	32His → Arg	0	0	15	0	0	15
c.392G>T	exon5	131Gly → Val	0	0	5	0	0	5
c.487G>A	exon6	163Gly → Ser	0	0	2	1	0	3
c.517T>C	exon6	173Phe → Leu	0	0	1	0	0	1
c.1024C>T	exon9	342Leu → Phe	0	0	13	1	0	14
c.1376G>T	exon12	459Arg → Leu	0	0	38	3	0	41
c.1388G>A	exon12	463Arg → His	0	0	40	1	0	41
c.95A>G ⁺ c.1388G>A	exon2 exon12	32His → Arg 463Arg → His	0	0	1	0	0	1
Total			0	0	115	6	0	121

表 3 G6PD 野生型及 4 种突变蛋白的二级结构预测情况
Tab 3 Prediction of secondary structure of G6PD wild type and four mutant proteins

Mutation	Total amino acid/n	α -helix/n (%)	β -fold/n (%)	β -turn/n (%)	Random coil/n (%)
Wild-type	515	228 (44.27)	66 (12.82)	37 (7.18)	184 (35.73)
c.95A>G	515	224 (43.49)	69 (13.40)	35 (6.80)	187 (36.31)
c.1388G>A	515	220 (42.72)	73 (14.17)	35 (6.80)	187 (36.31)
c.1376G>T	515	223 (43.30)	73 (14.17)	38 (7.38)	181 (35.15)
c.1024C>T	515	222 (43.11)	71 (13.79)	37 (7.18)	185 (35.92)

2.3.2 G6PD 蛋白野生型空间构型 G6PD 蛋白单体由多个 α -螺旋 (图 2A 蓝紫) 及 β -折叠 (图 2A 粉) 构成了 2 个结构域, 即底物结构域和 $\alpha+\beta$ 结构域。前者较小的位于肽链 N 端, 参与底物的结合; 后者主要由 9 股 β -折叠构成, 与亚基的聚合有关; 在折叠区还具有 NADP⁺ 配体的结合位点。G6PD 活性蛋白由 2 个相同的亚基构成 (图 2B

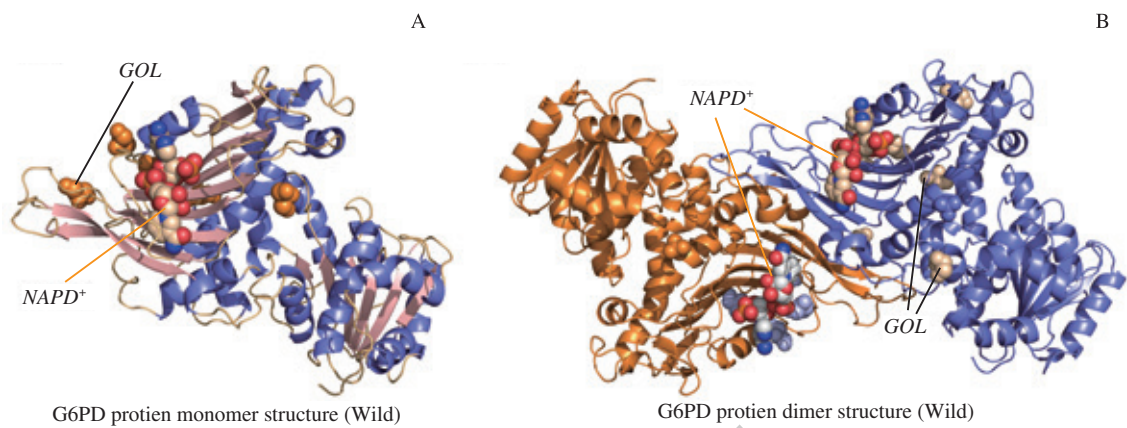
分型^[14-15]进行统计分析, 发现 121 例阳性记录仅表现出了 Class III 和 Class IV 2 种等级, 且各突变两两之间比较差异均无统计学意义, 见表 2。

2.3 G6PD 酶蛋白的空间结构

2.3.1 G6PD 酶蛋白的二级结构 G6PD 野生型蛋白主要包含 α -螺旋、 β -折叠、 β -转角以及无规则卷曲, 其中 α -螺旋有 228 个氨基酸, β -折叠占 66 个氨基酸, β -转角有 37 个氨基酸, 无规则卷曲有 184 个氨基酸, 4 种二级结构呈现于整个氨基酸链。与野生型蛋白相比, 4 种突变蛋白的二级结构均有不同程度的改变, 如表 3 所示: 4 种突变构成 α -螺旋的氨基酸均减少, β -折叠的氨基酸均增加, c.95A>G 与 c.1388G>A 2 种突变的 β -转角氨基酸均减少了 2 个, 无规卷曲增加了 3 个。

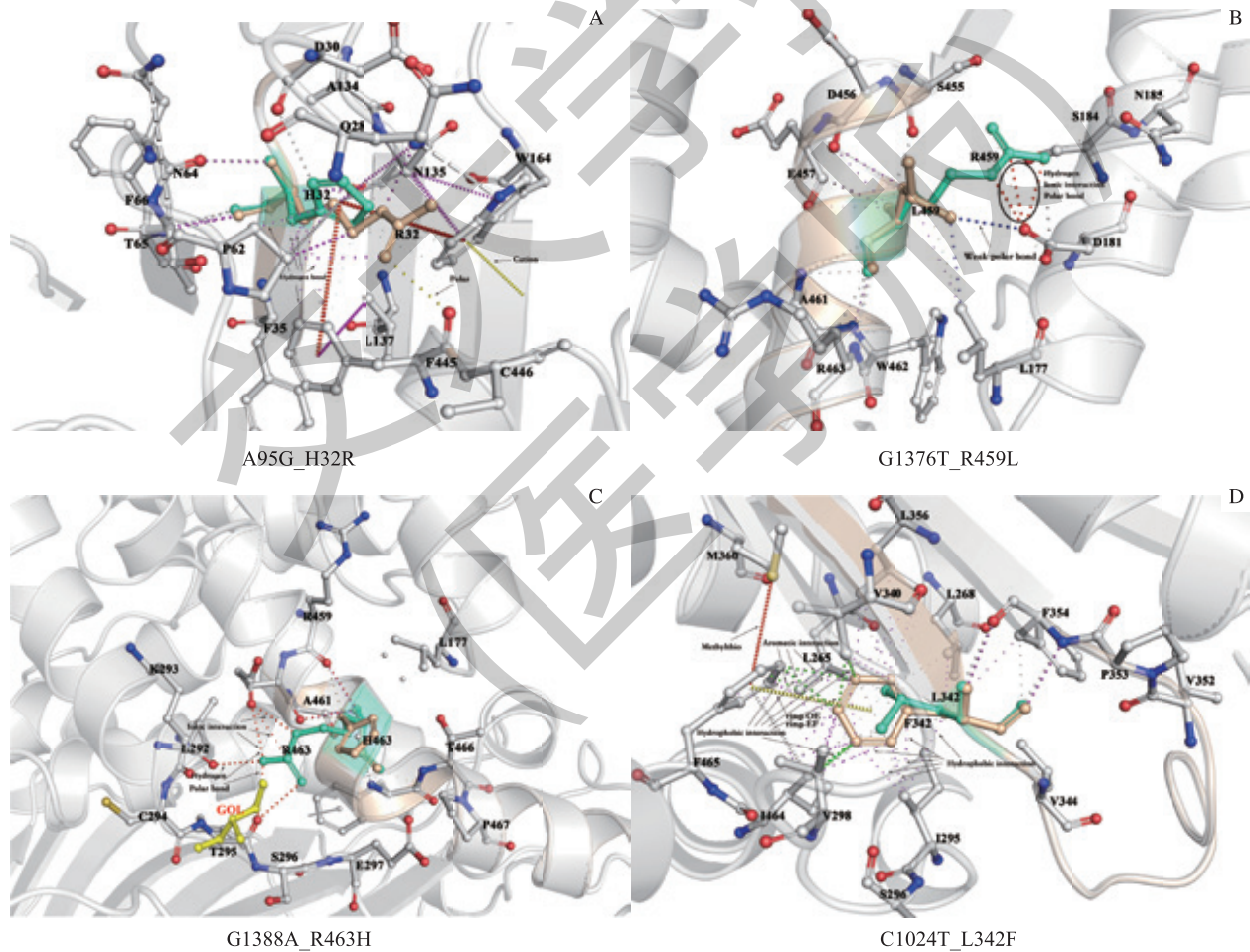
橘黄和蓝紫), 具有 2 个 NADP⁺ 配体的结合位点以及多个 GOL 位点 (图 2)。

2.3.3 G6PD 活性蛋白 4 种高频突变的空间构型 图 3 中显示了 4 种错义突变氨基酸情况, 以及突变位点与其他氨基酸共价键形成情况及相互作用的情况。



Note: A. G6PD protein monomer structure (constructed according to 6E08 in the PDB protein database). B. G6PD protein dimer structure (G6PD protein in mammals needs to be polymerized into dimer or tetramer to be active).

图 2 G6PD 蛋白空间构型
Fig 2 G6PD protein spatial configuration

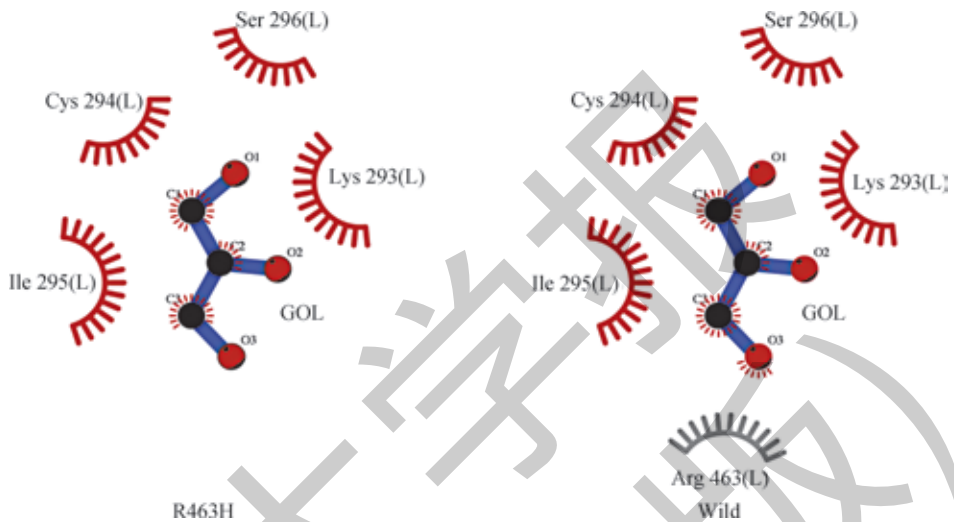


Note: Green amino acids indicate wild-type, and golden amino acids indicate mutants. Gray-white linkages indicate wild-type interactions and wild-type and mutant common linkages, and plum-colored linkages show mutant-type interactions. A. After the mutation at position 32, the aromatic ring (red) disappears, and the cation and polar bond interactions increase (yellow). B. The hydrogen bond of D181 (red) and N185 adjacent to the sacral helix is broken after the mutation at position 459. C. The mutation at position 463 causes the link broken with the adjacent zygomatc helix (red), and the substrate (yellow) is close to position 463 to interact with the wild type. D. Mutation of position 342 from leucine to phenylalanine, a forming ring (yellow), multiple aromatic groups (green), and more hydrophobic interlinkages (plum).

图 3 4 种突变位点与相邻氨基酸的相互作用变化情况
Fig 3 Interaction between four mutation sites and adjacent amino acids

从图中可以看到, H32 位于蛋白 Rossman 结构区, H32R 突变前后部分残基之间的共价键断裂, 成型环消失, 此外也形成了新的阳离子链接区及极性键作用区 (图 3A)。R459 位于蛋白表面 α 螺旋远离辅酶和底物结合区, 用于稳定自身 α 螺旋与相邻 α 螺旋。与野生型相比, R459L 突变后与相邻 α 螺旋间氢键、离子作用及极性作用均消失, 与 N181 号位氨基酸形成新的弱极性键 (图 3B)。

R463 同样处于蛋白表面 α 螺旋处, 但与底物葡萄糖 -6- 磷酸结合位点相邻 (图 4), R463H 突变后部分氢键断裂、离子相互作用减弱, 并且远离了配体甘油 (GOL) 的结合位点 (图 3C、图 4), L342 位氨基酸处于 α - β 结构域中的 β - 折叠区, L342F 突变导致该 β 结构角度发生改变, 另外形成了更多的疏水间连接, 产生芳香基, 出现成型环结构 (图 3D)。



Note: The amino acid at this position deviates from the substrate and the interaction disappears in the R463 mutation position (left).

图 4 配体 GOL 与氨基酸相互作用情况
Fig 4 Interaction between ligand GOL and amino acids

2.3.4 稳定性分析 采用 6 种常用的蛋白稳定性预测方法对 4 个突变后蛋白进行分析, 综合结果显示 4 种突变均可

能造成不同程度蛋白稳定性的降低, 尤其 c.1388G>A 突变最为显著 (表 4)。

表 4 6 种预测方法对 4 种突变蛋白稳定性预测结果
Tab 4 Six prediction methods for predicting the stability of four mutant proteins

Mutation	$\Delta\Delta G/(\text{kcal} \cdot \text{mol}^{-1})$						Stability
	Mupro	SDM	mSCM	DUET	Dynamut	CUPSAT	
c.95A>G	-0.813	-1.61	-1.207	-1.213	0.169	2.34	Destabilizing
c.1388G>A	-1.019	-1.1	-1.722	-1.806	-0.588	-0.86	Destabilizing
c.1376G>T	0.324	0.39	-0.533	-0.088	-0.605	-1.29	Destabilizing
c.1024C>T	-0.958	-0.78	-1.520	-1.490	0.690	-3.2	Destabilizing

Note: 1 kcal=4 184 J.

2.3.5 氨基酸改变对蛋白功能的影响 采用 PROVEAN 分析软件对 4 种氨基酸改变导致的蛋白功能改变情况进行预测, 结果显示 4 种突变的 PROVEAN 得分均低于阈值 -1.3, 被判定为有害影响 (表 5)。

此现象表明, 在蛋白质三维结构上, 不同部分的 4 个错义突变对 G6PD 的表达具有负面影响。

表 5 4 种突变位点对蛋白功能影响的预测结果
Tab 5 Prediction of the effect of four mutation positions on protein function

Mutation positions	PROVEAN score/score	Result
c.95A>G	-4.326	Deleterious
c.1376G>T	-2.788	Deleterious
c.1388G>A	-4.460	Deleterious
c. 1024C>T	-2.101	Deleterious

3 讨论

本研究回顾性分析了205 103例新生儿G6PD酶活性筛查记录,报道了其中的230例酶活性检测G6PD阳性样本,并对其中121例进行了基因突变位点的检测;共检出8种常见突变,其中3种中国人群最常见突变被报道,发生率与其他研究相似,并有1种突变c.1024C>T在本研究中也表现出了高发生率。G6PD基因的突变导致氨基酸的替换进而引起酶活性水平的降低。通过血液学研究,可根据红细胞中残留酶活性水平来判断G6PD缺乏症症状严重程度。本研究的测定样本均来自新生儿滤纸干血片,以2.0 U/g作为筛查临界值,所有突变体均显示出了比野生型酶蛋白低的活性(图1)。然而突变导致酶活性水平的降低可能是由于影响到G6PD二聚体的形成、蛋白稳定性、NADP⁺结合区或者底物结合区等,比如F501号位氨基酸突变成C501改变了NADP⁺结合位点的大小,L205号位氨基酸与底物G6P的结合及催化作用有关,不同的影响因素可能引起不同的临床表征。

通过对G6PD的结构分析,发现 α -螺旋及 β -折叠氨基酸数量的改变可能会直接影响多肽链在空间上的折叠情况,造成固有结构域上原子或氨基酸之间的错位。c.95A>G突变导致32号位的组氨酸转变成精氨酸,H32位点位于蛋白Rossman结构域,是 β - α - β 组成的超二级结构单元,其序列为His-Ile-Phe-Ile-Ile-Met,由32号组氨酸开始,至37号甲硫氨酸结束,共6个氨基酸。Rossman折叠的结构域常含有功能性结合部位或活性部位,能调控整个酶的活性,因此在该位点上的突变可能会改变口袋的位置或大小,从而影响酶活性。此外,很多Rossman结构区具有磷酸化位点,磷酸化位点常出现在His或Asp位置;而在该单元中32号位恰好是His位置,因此该位点也可能与蛋白的磷酸化/去磷酸化有关,该位点的突变可能导致了该位点无法去磷酸化,从而影响酶活性。c.1376G>T突变导致精氨酸(R)突变成亮氨酸(L),亮氨酸是一个更为保守的氨基酸,氨基酸的保守突变在决定蛋白质三维结构与功能方面起重要作用。野生型R459与D181和N185之间的螺旋间相互作用对于酶的催化活性和稳定性有重要影响,而突变型L459导致螺旋间相互作

用被削弱,从而使得相邻两螺旋的位移变大,螺旋结构变得松散。另外,L459号位氨基酸的侧链与四聚体聚合有关,二聚体和四聚体的动态平衡对维持G6PD蛋白具有重要作用。R463与R459同处一个 α 螺旋,与L292、C294、I295、E460形成了稳定的氢键、离子相互作用,稳定蛋白 α 螺旋结构以及底物G6P结合区。H463突变导致该位点的侧链集团远离底物,与底物相互作用消失(图4),并且打断了与L292、C294、I295等氨基酸的各种共价键与非共价键作用(图3C),与E460号位的盐键崩溃,进而导致酶活性降低。L342位于G6PD中 α + β 结构域的9股不平衡 β -折叠区,后与c.1024C>T突变导致了亮氨酸突变成苯丙氨酸,F342增大了该位点氨基酸所需的空间(图3D),而该位点的后一位氨基酸Y343与底物GOL接触,L342突变可能在空间结构上会造成Y343与底物的错位;L342处于 β 发夹结构并靠近 β 转角(PDB 6e08),与F354号位氨基酸形成氢键稳定发夹结构,F342突变不仅会破坏稳定的发卡结构,还可能由于空间障碍阻碍 β -转角的形成,使酶的三维结构改变。另外F342突变增加了更多的疏水键,使得酶蛋白结构的刚性增加,而蛋白活性区的柔性与蛋白催化效率直接相关已经被证实,因此这一突变可能会导致蛋白催化活性的降低。4种突变上述的结构错位、空间障碍、氢键断裂、盐键崩溃、柔性降低等影响,均可能导致酶蛋白的不稳定,从而降低突变蛋白的稳定性;如表4中总结,4种突变蛋白稳定性均表现出降低,并且对G6PD蛋白均会造成不良影响,从而引起G6PD蛋白的缺乏,造成新生儿的溶血等。

本研究详细地分析了G6PD基因突变与酶活性降低的关系,根据基因突变引起的表达差异,从蛋白结构和功能上分析导致酶活性降低的可能原因。一方面对指导临床有重要作用;另一方面为我国G6PD常见基因突变酶学结构与功能的研究提供了资料和理论依据,也为研究儿童遗传代谢类疾病提供了蛋白结构的分析方向。在G6PD酶学结构及功能的研究和应用上,仍有许多未解决的问题有待探索。大多数已知的突变位点的蛋白结构仍然未被解析,酶活性降低是由于蛋白结构改变、稳定性降低、聚合能力的缺失还是配体结合能力的改变仍然不清楚,还有待未来研究人员探索。

参·考·文·献

- [1] Cohen P, Rosemeyer MA. Subunit interactions of glucose-6-phosphate dehydrogenase from human erythrocytes[J]. Eur J Biochem, 1969, 8(1): 8-15.
- [2] Au SW, Gover S, Lam VM, et al. Human glucose-6-phosphate dehydrogenase: the crystal structure reveals a structural NADP⁺ molecule and provides insights into enzyme deficiency[J]. Structure, 2000, 8(3): 293-303.
- [3] Minucci A, Moradkhani K, Hwang MJ, et al. Glucose-6-phosphate dehydrogenase (G6PD) mutations database: review of the "old" and update of the new mutations[J]. Blood Cells Mol Dis, 2012, 48(3): 154-165.
- [4] Lin F, Lou ZY, Xing SY, et al. The gene spectrum of glucose-6-phosphate dehydrogenase (G6PD) deficiency in Guangdong Province, China[J]. Gene, 2018, 678: 312-317.
- [5] 杜传书, 王菁, 陈路明, 等. 中国人中发现的 6 种 *G6PD* 基因点突变 [J]. 中山大学学报 (医学科学版), 1993, 14(4): 291.
- [6] Yan JB, Xu HP, Xiong C, et al. Rapid and reliable detection of glucose-6-phosphate dehydrogenase (G6PD) gene mutations in Han Chinese using high-resolution melting analysis[J]. J Mol Diagn, 2010, 12(3): 305-311.
- [7] 林芬, 杨辉, 杨立业. 我国葡萄糖-6-磷酸脱氢酶缺乏症的分布特征和基因突变 [J]. 分子诊断与治疗杂志, 2016, 8(2): 73-77.
- [8] Hsu J, Fink D, Langer E, et al. PCR-based allelic discrimination for glucose-6-phosphate dehydrogenase (G6PD) deficiency in Ugandan umbilical cord blood[J]. Pediatr Hematol Oncol, 2014, 31(1): 68-75.
- [9] Tseng CP, Huang CL, Chong KY, et al. Rapid detection of glucose-6-phosphate dehydrogenase gene mutations by denaturing high-performance liquid chromatography[J]. Clin Biochem, 2005, 38(11): 973-980.
- [10] Xia ZM, Chen P, Tang N, et al. Rapid detection of G6PD mutations by multicolor melting curve analysis[J]. Mol Genet Metab, 2016, 119(1/2): 168-173.
- [11] Scopes DA, Bautista JM, Naylor CE, et al. Amino acid substitutions at the dimer interface of human glucose-6-phosphate dehydrogenase that increase thermostability and reduce the stabilising effect of NADP[J]. Eur J Biochem, 1998, 251(1/2): 382-388.
- [12] Hirono A, Kuhl W, Gelbart T, et al. Identification of the binding domain for NADP⁺ of human glucose-6-phosphate dehydrogenase by sequence analysis of mutants[J]. PNAS, 1989, 86(24): 10015-10017.
- [13] Jiang WY, Yu GL, Liu P, et al. Structure and function of glucose-6-phosphate dehydrogenase-deficient variants in Chinese population[J]. Hum Genet, 2006, 119(5): 463-478.
- [14] WHO Working Group. Glucose-6-phosphate dehydrogenase deficiency[J]. Bull World Health Organ, 1989, 67(6): 601-611.
- [15] Cappellini MD, Fiorelli G. Glucose-6-phosphate dehydrogenase deficiency[J]. Lancet, 2008, 371(9606): 64-74.
- [16] Buchan DWA, Jones DT. The psipred protein analysis workbench: 20 years on[J]. Nucleic Acids Res, 2019, 47(W1): W402-W407.
- [17] McGuffin LJ, Bryson K, Jones DT. The PSIPRED protein structure prediction server[J]. Bioinformatics, 2000, 16(4): 404-405.
- [18] Geourjon C, Deléage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments[J]. Bioinformatics, 1995, 11(6): 681-684.
- [19] Drozdetskiy A, Cole C, Procter J, et al. JPred4: a protein secondary structure prediction server[J]. Nucleic Acids Res, 2015, 43(W1): W389-W394.
- [20] Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: homology modelling of protein structures and complexes[J]. Nucleic Acids Res, 2018, 46(W1): W296-W303.
- [21] ww PDB. Protein Data Bank: the single global archive for 3D macromolecular structure data[J]. Nucleic Acids Res, 2019, 47(D1): D520-D528.
- [22] Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank[J]. Nucleic Acids Res, 2000, 28(1): 235-242.
- [23] Cheng JL, Randall A, Baldi P. Prediction of protein stability changes for single-site mutations using support vector machines[J]. Proteins, 2005, 62(4): 1125-1132.
- [24] Pandurangan AP, Ochoa-Montano B, Ascher DB, et al. SDM: a server for predicting effects of mutations on protein stability[J]. Nucleic Acids Res, 2017, 45(W1): W229-W235.
- [25] Worth CL, Preissner R, Blundell TL. SDM: a server for predicting effects of mutations on protein stability and malfunction[J]. Nucleic Acids Res, 2011, 39(suppl): W215-W222.
- [26] Parthiban V, Gromiha MM, Schomburg D. CUPSAT: prediction of protein stability upon point mutations[J]. Nucleic Acids Res, 2006, 34(Web Server): W239-W242.
- [27] Pires DEV, Ascher DB, Blundell TL. mCSM: predicting the effects of mutations in proteins using graph-based signatures[J]. Bioinformatics, 2014, 30(3): 335-342.
- [28] Pires DEV, Ascher DB, Blundell TL. DUET: a server for predicting effects of mutations on protein stability using an integrated computational approach[J]. Nucleic Acids Res, 2014, 42(W1): W314-W319.
- [29] Rodrigues CH, Pires DE, Ascher DB. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability[J]. Nucleic Acids Res, 2018, 46(W1): W350-W355.
- [30] Laskowski RA, Swindells MB. LigPlot⁺: multiple ligand-protein interaction diagrams for drug discovery[J]. J Chem Inf Model, 2011, 51(10): 2778-2786.
- [31] Choi Y, Chan AP. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels[J]. Bioinformatics, 2015, 31(16): 2745-2747.
- [32] Hanukoglu I. Proteopedia: Rossmann fold: a beta-alpha-beta fold at dinucleotide binding sites[J]. Biochem Mol Biol Educ, 2015, 43(3): 206-209.
- [33] 徐芸, 罗建明. 我国 G6PD 缺乏症基因突变的研究现状 [J]. 中国小儿血液与肿瘤杂志, 2009, 14(3): 143-145.
- [34] 董莹, 刘淑萍, 徐艳春, 等. 云南省 1 例伯氨喹诱发溶血患者的 *G6PD* 基因编码区突变分析及酶蛋白空间结构预测 [J]. 中国寄生虫学与寄生虫病杂志, 2019, 37(4): 399-405.
- [35] Rowland P, Basak AK, Gover S, et al. The three-dimensional structure of glucose 6-phosphate dehydrogenase from *Leuconostoc mesenteroides* refined at 2.0 Å resolution[J]. Struct Lond Engl, 1994, 2(11): 1073-1087.
- [36] Wang XT, Lam VMS, Engel PC. Marked decrease in specific activity contributes to disease phenotype in two human glucose 6-phosphate dehydrogenase mutants, G6PDUnion and G6PDAndalus[J]. Hum Mutat, 2005, 26(3): 284.
- [37] Boonyuen U, Chamchoy K, Swangsri T, et al. A trade off between catalytic activity and protein stability determines the clinical manifestations of glucose-6-phosphate dehydrogenase (G6PD) deficiency[J]. Int J Biol Macromol, 2017, 104: 145-156.
- [38] Bautista J, Mason PJ, Luzzatto L. Human glucose-6-phosphate dehydrogenase Lysine 205 is dispensable for substrate binding but essential for catalysis[J]. FEBS Lett, 1995, 366(1): 61-64.
- [39] Lesk AM. NAD-binding domains of dehydrogenases[J]. Curr Opin Struct Biol, 1995, 5(6): 775-783.
- [40] Hwang S, Mruk K, Rahighi S, et al. Correcting glucose-6-phosphate dehydrogenase deficiency with a small-molecule activator[J]. Nat Commun, 2018, 9(1): 4045.
- [41] 邹承鲁. 酶活性部位的柔性 [J]. 科学通报, 1989(5): 321-325.
- [42] 邹承鲁. 活性部位的柔性 [J]. 生理科学进展, 2001(1): 7-12.

[收稿日期] 2020-01-10

[本文编辑] 徐 敏