

BGC823 中 ND5 基因突变对其蛋白影响的生物信息预测

李 恋 沈晓玲 张明昱, 包丽丽 纳仁高娃

摘要 目的 分析总结胃癌细胞 BGC823 中 ND5 基因突变, 利用生物信息学方法分析比较正常和突变的 ND5 蛋白的结构和功能。方法 从胃癌细胞 BGC823 中扩增 ND5 基因, 分析蛋白错义突变的位点, 通过 ProtParam、TMHMM、Predict-Protien、Swiss-Model、Swiss-Pdb Viewer 等生物软件分析比较正常和突变的 ND5 蛋白理化性质、跨膜区、二级及三级结构等。结果 胃癌细胞 BGC823 的 ND5 蛋白有两处错义突变位点。突变后的 ND5 蛋白不稳定系数升高, 螺旋区分布和蛋白结合位点与正常 ND5 蛋白不同, 三级结构分析显示 167 位点突变后形成了新的氢键。结论 BGC823 中的突变位点对 ND5 蛋白的理化性质、二级结构、三级结构有影响。

关键词 BGC823; ND5 蛋白; 错义突变; 生物信息预测

中图分类号 R 735.2

文献标志码 A **文章编号** 1000-1492(2019)01-0074-05

doi: 10.19405/j.cnki.issn1000-1492.2019.01.015

呼吸链由 4 个功能相对独立的蛋白复合物组成, 复合物 I 是电子转运链上最大、最复杂的酶。真核生物中, 复合物 I 由 45 个不同的蛋白亚基构成, 其中仅 7 个蛋白由线粒体基因编码, 包括重要亚基 ND5, 其余由核基因编码^[1]。线粒体基因无组蛋白保护, 易受到环境中高浓度氧化物损伤而突变, 有研究^[2-3]证明线粒体突变会降低 NADH 脱氢酶活性, 使细胞供能失衡引起疾病。线粒体突变与 MELAS 综合征^[4]、帕金森病^[5]、肿瘤^[6]等有关, 甚至与肿瘤发生发展及预后相关^[7]。细菌^[8]、哺乳动物^[9-10]复合物 I 的结构已解析, 但人细胞中复合物 I 的重要

亚基 ND5 结构及功能没有详细报道。该研究采用生物信息学方法预测、比较分析正常细胞和 BGC823 中 ND5 差异, 为深入探究 ND5 基因突变与胃癌关系奠定理论基础。

1 材料与方法

1.1 材料 通过基因组提取试剂盒(北京天根生化科技有限公司)提取得到的 BGC823 总 DNA。

1.2 方法

1.2.1 PCR 反应并测序 设计两对引物扩增 ND5 (1-2)、ND5(3-4) 两个片段, 扩增体系(50 μl): DNA 模板 100 ng, dNTP(2.5 mmol/L) 3 μl, 上、下游引物(10 μmol/L) 各 2 μl, 10 × Buffer 5 μl, Taq 酶 0.5 μl, 超纯水补足体积至 50 μl。PCR 反应条件: 94 °C 预变性 5 min, 进入热循环: 94 °C 变性 30 s, 退火 30 s (退火温度见表 1), 72 °C 延伸 90 s, 30 个循环。

表 1 扩增 ND5 基因引物表

引物名称	引物序列 (5'-3')	产物 (bp)	退火温度 (°C)
ND5-1F	ACTTTTAAAGGATAACAGCTATCC	ND5(1-2)	58
ND5-2R	TAATGAGAAATCCTGCGAATA	1 476	
ND5-3F	TCATAATAGTTACAATCGGCAT	ND5(3-4)	56
ND5-4R	ATATTGTAATTGAGATTGCTCG	999	

PCR 产物 ND5(1-2)、ND5(3-4) 两个片段在 1% 的琼脂糖凝胶上进行电泳, 检测条带特异性并切胶纯化后由奥科生物技术公司和华大基因测序。

1.2.2 生物信息学比较分析 ProtParam 用于分析 ND5 蛋白质理化性质; 用 EASE-MM 分析稳定性; 采用 ProtScale、TMHMM 预测蛋白质的亲/疏水性和跨膜区; 借助 PredictProtien 和 Swiss-Model 在线预测软件对蛋白质的二级和三级结构进行预测分析, 用 Swiss-Pdb Viewer Manual v3.7 软件分析点突变对 ND5 三级结构的影响。

2 结果

2.1 胃癌细胞 BGC823 中 ND5 基因突变情况 用引物 ND5(1-2)、ND5(3-4) PCR 扩增、测序后拼接可

2018-08-24 接收

基金项目: 内蒙古自治区高等学校科学研究项目(编号: NJZC16109); 内蒙古医科大学本科教学质量工程项目(编号: NYJTXX201712); 内蒙古医科大学本科教学改革项目(编号: NYJXGG2018038); 内蒙古医科大学实验室开放基金项目(编号: 2015ZN17); 内蒙古自然科学基金项目(编号: 2015MS0888); 内蒙古医科大学博士启动基金(编号: YKD2015BQ03)

作者单位: 内蒙古医科大学基础医学院, 呼和浩特 010059

作者简介: 李 恋, 女, 讲师;

纳仁高娃, 女, 副教授, 责任作者, E-mail: narensan128@hotmail.com;

包丽丽, 女, 副教授, 责任作者, E-mail: baolili1203@126.com

以获得 ND5 基因全长。以修正剑桥参考序列 (rCRS: NC. 012920) 为对照, rCRS 作为正常细胞线粒体 ND5 基因, 分析比较获得 BGC823 细胞的 ND5 基因, 将该细胞中与 rCRS 序列有差异的位点见表 2。经分析氨基酸改变情况, BGC823 细胞中在 12835、13105 位点处发生了错义突变。

表 2 BGC823 细胞中 ND5 基因突变位点

BGC823	位点			
	12705	12835	13105	13914
碱基变化	C-T	G-G/A	A-G	C-A
氨基酸变化	I-I	A-T	I-V	L-L

注: G-A 代表 rCRS 该位点碱基 G, 但 BGC823 细胞中为 A; G-G/A 代表 rCRS 该位点碱基 G, 但 BGC823 细胞中同时存在 G 和 A

2.2 人 ND5 蛋白的理化性质预测和分析 借助 ProtParam (<http://web.expasy.org/protparam/>) 蛋白质理化性质预测网站对正常人 ND5 蛋白进行理化性质预测, 并与发生异质性点突变的 BGC823 细胞中 ND5 蛋白进行比对。人线粒体 ND5 基因编码 603 个氨基酸; 相对分子质量为 67 026. 51; 人线粒体 ND5 蛋白等电点为 9. 14, 说明 ND5 是碱性蛋白质; 哺乳动物 ND5 蛋白的半衰期达到了 30 h, ND5 蛋白的不稳定系数 33. 95, 小于 40, 属于稳定蛋白质。同时, 平均疏水性预测结果显示: 0. 589, 表明是疏水蛋白质。以上特点符合已报道的 ND5 蛋白的特性。对 BGC823 的 ND5 蛋白序列进行预测, 显示相对分子质量为 67 042. 51; 等电点相同; 平均疏水性预测结果为 0. 584, 说明点突变并未改变蛋白疏水性; ND5 蛋白的不稳定系数为 34. 9, 虽然也低于 40, 但比正常 ND5 蛋白值稍高, 说明突变降低蛋白稳定性。进一步通过 EASE-MM (<http://sparks-lab.org/server/ease>) 在线预测 12835 和 13105 两处位点发生的突变对 ND5 蛋白稳定性的影响, 结果显示两位的氨基酸的改变均很可能降低了该蛋白的稳定性。

2.3 预测与分析突变对 ND5 蛋白亲水性/疏水性和跨膜区的影响 使用 ProtScale (<http://web.expasy.org/protscale/>) 在线预测的工具, 对人 ND5 蛋白进行亲水性/疏水性分析, 结果如图 1 所示。ND5 蛋白亲水性的最强位点是 29 位的赖氨酸 (Lysine, K), 分值为 -2. 622; 疏水性最强的位点是 257 位的异亮氨酸 (Isoleucine, I), 分值为 3. 244。BGC823 细胞中的 ND5 蛋白亲水性的最强位点未改变, 但疏水性最强的位点变为 126 位的异亮氨酸

(Isoleucine, I)、127 位的苏氨酸 (Threonine, T)、257 位的异亮氨酸 (Isoleucine, I), 分值均为 3. 211。最强疏水位点的增多可能改变蛋白表面的结合位点及疏水肽段分布。ProtScale 预测结果见图 1, 提示人 ND5 蛋白疏水肽链与亲水肽链相间分布在氨基酸序列中, 结合 ProtParam 的预测结果, 显示 ND5 蛋白是疏水蛋白质。

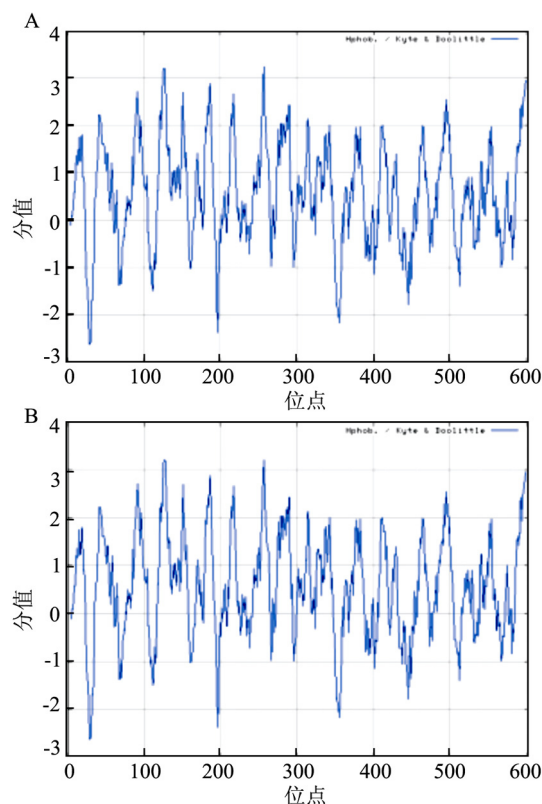


图 1 ND5 蛋白的亲水性/疏水性分析

A: 正常 ND5 蛋白结果; B: BGC823 中 ND5 蛋白结果

利用 TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>) 软件对人 ND5 蛋白进行跨膜区分分析, 结果提示该蛋白为多次跨膜蛋白质, 该结果与 ProtScale 软件预测结果一致。TMHMM 预测 ND5 蛋白螺旋分别在 35 ~ 57、85 ~ 107、120 ~ 137、141 ~ 160、172 ~ 191、211 ~ 233、245 ~ 267、272 ~ 294、301 ~ 320、325 ~ 347、368 ~ 387、407 ~ 429、450 ~ 472、482 ~ 504 和 583 ~ 602 氨基酸之间, N 端在胞内、C 端在细胞外, 见图 2。BGC823 中 ND5 蛋白跨膜区分分析结果与正常人 ND5 蛋白分析结果一致, 说明该软件分析显示突变可能并未影响跨膜区的肽段分布, 进一步在二级结构中使用其他软件进行分析, 以相互验证。

2.4 人 ND5 蛋白二级结构的预测和分析 通过

PredictProtein (<https://www.predictprotein.org/>) 在线预测 ND5 蛋白的二级结构,结果见图 3A。该蛋白含有 17 个跨膜螺旋区和 2 个 β -折叠区,螺旋区和 β -折叠区的氨基酸比率分别为 73.47% 和 2.65%,LOOP 区氨基酸比率为 23.88%,说明人 ND5 蛋白近 3/4 的氨基酸为螺旋结构。结果显示 91.71% 的氨基酸残基是疏水的,6.3% 的氨基酸残基是亲水的,1.99% 的氨基酸残基介于两者之间。

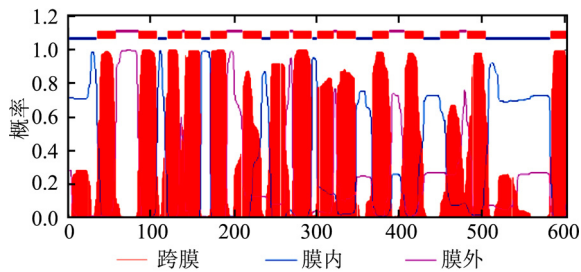


图 2 ND5 蛋白的跨膜区域预测

BGC823 中 ND5 蛋白的螺旋区与正常 ND5 蛋白有差别,见图 3B 在 243 ~ 349 位点区正常 ND5 蛋白形成三个螺旋区(243 ~ 268、270 ~ 319、320 ~ 349),BGC823 的 ND5 蛋白形成两个螺旋区(243 ~ 319、321 ~ 348);在 381 ~ 389 位点区 BGC823 的 ND5 蛋白预测结果中多出一个螺旋(385 ~ 386);在 389 ~ 431 位点区 BGC823 的 ND5 蛋白预测分为两段螺旋(389 ~ 402、404 ~ 431);在 529 ~ 580 位点区 BGC823 的 ND5 蛋白预测分为三段螺旋(529 ~ 540、542 ~ 561、563 ~ 580)。两者在跨膜螺旋区预测结果也有差异,BGC823 中 ND5 蛋白在 536 ~ 573 位点区

有两段跨膜螺旋(536 ~ 553、557 ~ 573),而正常 ND5 蛋白仅有一段跨膜螺旋(548 ~ 563)。同时,预测结果显示,蛋白结合位点两者有差别,BGC823 的 ND5 蛋白比正常 ND5 蛋白预测结果中在 438 位点和 575 位点均多了一个蛋白结合位点,而在 535 位点少了一个蛋白结合位点。BGC823 中 ND5 的螺旋区、 β -折叠区和 LOOP 区的氨基酸比率没有变化,但疏水氨基酸残基比率降为 91.04%,亲水氨基酸残基比率升高为 6.47%,介于疏水和亲水性之间的氨基酸残基的比率为 2.4%。可见 BGC823 中两处氨基酸位点的改变影响了疏水肽段和蛋白结合位点的分布,对 ND5 蛋白二级结构有一定影响。

2.5 人 ND5 蛋白三级结构的预测和分析 Swiss-Model (<https://www.swissmodel.expasy.org/>) 可利用同源建模的方法对一段未知的蛋白三级结构作预测。将人 ND5 蛋白序列输入该在线预测平台后,会自动选择与输入序列同源性最高的已知序列结构为建模主要参照结构,见图 4A,其中一个已知结构序列(SMTL ID: 5xtc.1) 同源性高达 99.5%。此序列是由 Guo et al^[11] 完成结构分析,他们研究组对人类电子传递链上的复合物 I 结构进行了分析。人 ND5 蛋白构建模型的 GMQE(全球模型质量评估)分数为 0.98,接近 1,说明预测可靠性高。进一步分析人 ND5 蛋白的氨基酸序列与模型蛋白质相似性波形图,见图 4B,预测结构的波形较稳定且绝大部分的氨基酸残基分值不低于 0.6,说明模型较趋近于真实情况。

为了进一步预测 BGC823 中具有两处氨基酸改

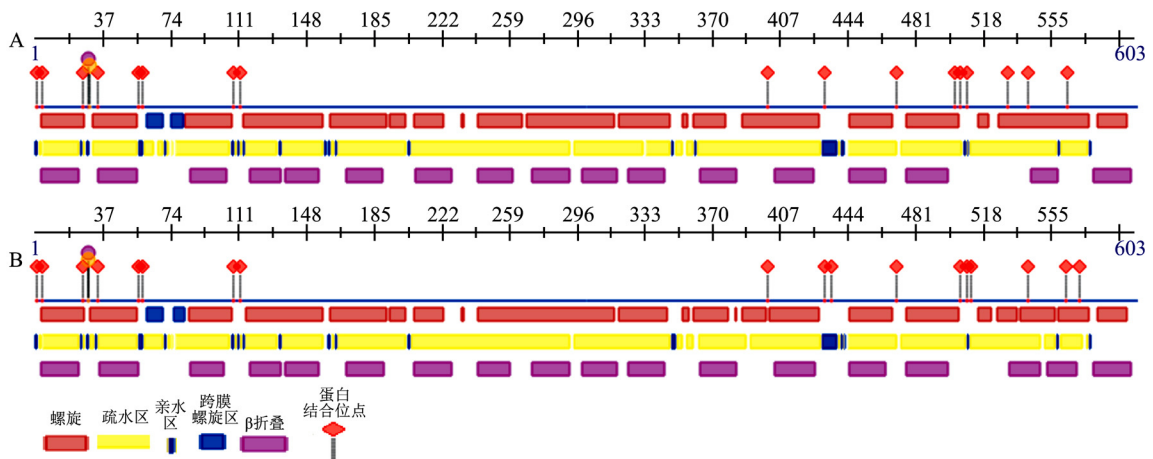


图 3 人 ND5 蛋白的二级结构分析

A: 正常 ND5 二级结构分析结果; B: BGC823 中 ND5 二级结构分析结果

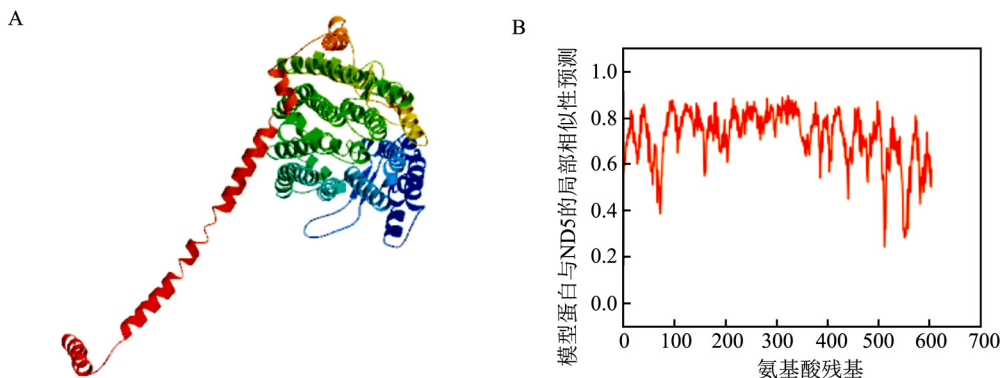


图4 ND5 蛋白三级结构及其同源蛋白质相似性波形图

A: 人 ND5 蛋白预测的三级结构; B: ND5 同源蛋白质相似性波形图

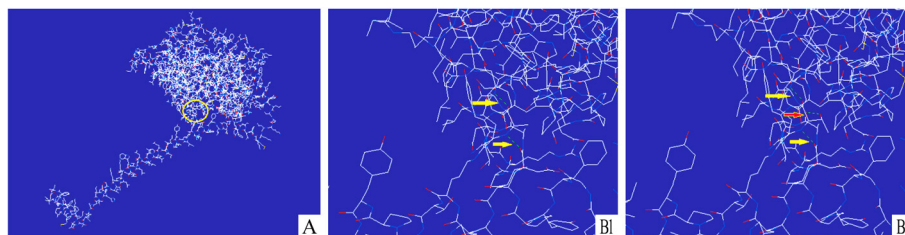


图5 人 ND5 三级结构及突变前后 167 位氢键改变

A: ND5 蛋白三级结构; B1: 正常 ND5 蛋白 167 位点处氢键情况; B2: 突变 ND5 蛋白 167 位点处氢键情况; 圆圈为突变位点在 ND5 三级结构中的位置; 黄色箭头指向突变前后一致的化学键; 红色箭头指向突变后出现的新化学键

变位点的 ND5 蛋白三级结构是否有改变,使用 Swiss-Pdb Viewer Manual v3.7 软件进行分析。BGC823 中 ND5 蛋白 167 位的丙氨酸(alanine, A) 转变为苏氨酸(threonine, T), 疏水氨基酸转变为亲水氨基酸, 257 位的异亮氨酸(Isoleucine, I) 转变为缬氨酸(Valine, V), 将人 ND5 蛋白模型的 167、257 位点进行突变, 正常 ND5 蛋白 167 位与 171、163 位分别形成氢键, BGC823 中 ND5 蛋白 167 位与 171、163 和 164 位分别形成氢键, 结果显示 167 位点突变使新的氢键形成, 见图 5。

3 讨论

ND5 基因位于线粒体染色体上, 其编码蛋白是组成复合物 I 的重要亚单位。复合物 I 由跨膜臂和伸入线粒体基质的亲水臂两部分组成, 亲水臂和跨膜臂连接呈 L 形。ND5 位于跨膜臂上, 是构成反向运转质子泵模式的核心亚基。复合物 I 的跨膜臂与亲水臂连接, 另一端与复合物 IV 结合, 主要通过 ND5 与复合物 IV 中的 COX7C 蛋白亚基相互作用^[10]。

本研究检测胃癌细胞系中 BGC823 的 ND5 基因点突变情况, 发现在 12835 和 13105 位点的突变会引起其翻译蛋白在 167 和 254 位点的错义突变。

进一步通过生物信息学分析比较了正常 ND5 蛋白与 BGC823 中出现错义突变的 ND5 蛋白在理化性质、二级结构、三级结构的差异。经 ProtParam 和 EASE-MM 生物信息在线数据库分析发现 BGC823 中 ND5 基因的两处突变很可能引起其蛋白稳定性下降。虽然突变前后等电点均为 9.14, 但经 ProtScale 软件预测发现突变后的 ND5 蛋白最强疏水位点在蛋白序列 126、127 位点增加两处。PredictProtein 软件预测结果显示突变后的 ND5 蛋白二级结构中螺旋区分布与正常蛋白有差别且蛋白结合位点增加, 这些改变很可能影响蛋白高级结构形成及蛋白活性。经 Swiss-Pdb Viewer Manual v3.7 软件预测在蛋白 167 位点突变后形成了新的氢键。两处突变造成 ND5 蛋白的物理性状和结构的改变, 很可能影响该蛋白正常生物学功能。ND5 基因突变与肿瘤发生、发展、预后的相关性已在不断被证实。已有相关研究报告 ND5 基因突变在多种肿瘤组织中被检测到, 与癌症的发生发展有关。Chattopadhyay et al^[12] 在口腔癌患者中检测 ND5 基因, 找到的 5 处点突变和 7 处 SNP 与肿瘤生长有关。Li et al^[13] 建议将 ND5 基因 13273 位碱基缺失作为患乳腺癌高风险检测的生物标记。Kabekkodu et al^[14] 发现宫颈癌组织中 ND5 基因的突变位点数较多, 仅次于 D-

Loop 基因。Li et al^[15] 高通量测序检测乳腺癌患者血液中的线粒体基因突变,其中包括 ND5 基因在内的一些线粒体基因突变可能被用作乳腺癌诊断的潜在标志物。

根据本研究结果,从发生机制上推测,很可能是由于 ND5 基因突变引起 ND5 蛋白结构改变,从而进一步影响复合物 I 的活性,甚至可能影响与其他复合物结合活性,最终引起细胞的能量代谢失衡,而导致肿瘤发生或预后不良等问题。

参考文献

- [1] Bridges H R, Mohammed K, Harbour M E, et al. Subunit NDUFV3 is present in two distinct isoforms in mammalian complex I[J]. *Biochim Biophys Acta*, 2017, 1858(3): 197–207.
- [2] Seong M W, Choi J, Park S S, et al. Novel MT-ND5 gene mutation identified in Leber's hereditary optic neuropathy patient using mitochondrial genome sequencing[J]. *J Neurol Sci*, 2017, 375: 301–3.
- [3] Mohamed Yusoff A A, Zulfakhar F N, Mohd Khair S Z N, et al. Mitochondrial 10398A > G NADH-dehydrogenase subunit 3 of complex I is frequently altered in intra-axial brain tumors in malaysia[J]. *Brain Tumor Res Treat*, 2018, 6(1): 31–8.
- [4] Grady J P, Pickett S J, Ng Y S, et al. mtDNA heteroplasmy level and copy number indicate disease burden in m. 3243A > G mitochondrial disease[J]. *EMBO Mol Med*, 2018, 10: e8262.
- [5] Pignataro D, Francia S, Zanetta F, et al. A missense MT-ND5 mutation in differentiated parkinson disease cytoplasmic hybrid induces ROS-dependent DNA damage response amplified by DRO-SHA[J]. *Sci Rep*, 2017, 7(1): 9528.
- [6] Ju Y S, Alexandrov L B, Gerstung M, et al. Origins and functional consequences of somatic mitochondrial DNA mutations in human cancer[J]. *eLife*, 2014, 3: e02935.
- [7] Kim H, Komiyama T, Inomoto C, et al. Mutations in the mitochondrial ND1 gene are associated with postoperative prognosis of localized renal cell carcinoma[J]. *Int J Mol Sci*, 2016, 17: 2049.
- [8] Berrisford J M, Baradaran R, Sazanov L A. Structure of bacterial respiratory complex I[J]. *Biochim Biophys Acta*, 2016, 1857(7): 892–901.
- [9] Vinothkumar K R, Zhu J, Hirst J. Architecture of mammalian respiratory complex I[J]. *Nature*, 2014, 515: 80–4.
- [10] Gu J K, Wu M, Guo R Y, et al. The architecture of the mammalian respirasome[J]. *Nature*, 2016, 537: 639–43.
- [11] Guo R, Zong S, Wu M, et al. Architecture of human mitochondrial respiratory megacomplex I₂III₂IV₂[J]. *Cell*, 2017, 170(6): 1247–57.
- [12] Chattopadhyay E, De Sarkar N, Singh R, et al. Genome-wide mitochondrial DNA sequence variations and lower expression of OXPHOS genes predict mitochondrial dysfunction in oral cancer tissue[J]. *Tumour Biol*, 2016, 37(9): 11861–71.
- [13] Li L H, Chen L D, Li J, et al. Correlational study on mitochondrial DNA mutations as potential risk factors in breast cancer[J]. *Oncotarget*, 2016, 7: 31270–83.
- [14] Kabekkodu S P, Bhat S, Mascarenhas R, et al. Mitochondrial DNA variation analysis in cervical cancer[J]. *Mitochondrion*, 2014, 16: 73–82.
- [15] Li L H, Kang T, Chen L D, et al. Detection of mitochondrial DNA mutations by high-throughput sequencing in the blood of breast cancer patients[J]. *Int J Mol Med*, 2014, 33: 77–82.

Bioinformatic analysis of the mutated ND5 protein in BGC823

Li Lian, Shen Xiaoling, Zhang Mingyu, et al

(The Basic Medical College of Inner Mongolia Medical University, Hohhot 010059)

Abstract Objective To sequence and identify the mutation of ND5 gene in BGC823, and to analyze the normal and mutant ND5 protein by bioinformatics. **Methods** To amplify ND5 gene from BGC823, and to analyze the missense mutations. ProtParam, TMHMM, PredictProtien, Swiss Model and Swiss-Pdb Viewer were used to analyze the physicochemical properties, secondary structure and tertiary structure of the normal and mutant proteins. **Results** Two missense mutation sites of ND5 protein were found in BGC823. The mutant ND5 protein with higher instability index had different helix regions and protein binding sites. The analysis of three dimensional structure showed that a new hydrogen bond form due to the mutation at 167 site. **Conclusion** Bioinformatics analysis indicates that the missense mutation changes the physicochemical properties, secondary structure and tertiary structure of ND5 protein.

Key words BGC823; ND5 protien; missense mutation; bioinformatic analysis