

用作交联剂的甲醛可改变细胞蛋白质谱

蒋 绰^{1,3} 谢力琦² 廖日晶³ 陈付学¹ 吴嘉熙^{3*}

(¹上海大学生命科学学院, 上海 200444; ²复旦大学生物医学研究院, 上海 200032;

³中国科学院上海临床研究中心/徐汇区中心医院, 上海 200031)

摘要 甲醛能可逆性地交联蛋白质或DNA分子, 因此被广泛应用于鉴定生物大分子复合物中的蛋白质或DNA。一般认为, 甲醛分子小, 可快速进入细胞, 且作为交联剂使用时浓度较高, 因而在细胞做出应激反应之前将蛋白质和DNA交联固定在自然状态下。但高浓度甲醛是否能在交联固定生物大分子之前改变细胞蛋白质水平, 尚无相关研究。如果属实, 则可能导致相关研究产生假阳性或假阴性的结果。作者运用蛋白质组学技术分析对比了经过高浓度甲醛和未经甲醛处理过的Jurkat细胞的蛋白质谱。结果表明, 高浓度甲醛可导致多个蛋白质水平产生差异。因此, 将甲醛作为交联试剂使用时, 应考虑设置甲醛处理的样本作为对照, 或者增加实验重复次数, 以避免鉴定到假阳性或假阴性的蛋白质或DNA。

关键词 蛋白质组学; 甲醛; 交联; LC-MS/MS

Formaldehyde Used as A Technical Crosslinking Tool May Change Cell Protein Patterns

Jiang Chuo^{1,3}, Xie Liqi², Liao Rijin³, Chen Fuxue¹, Wu Jiayi^{3*}

(¹College of Life Sciences, Shanghai University, Shanghai 200444, China;

²Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China;

³Xuhui Central Hospital, Shanghai Clinical Research Center, Chinese Academy Sciences, Shanghai 200031, China)

Abstract Formaldehyde can reversibly crosslink protein-protein or protein-DNA molecules. As a result, it is widely used to study the molecular components of protein-protein or protein-DNA complexes. It is generally accepted that as a small molecule, formaldehyde, at the high concentrations that are typically used as a crosslinking tool, can rapidly permeate cross the cell membrane and instantaneously crosslink protein or DNA molecules before any cellular responses occur. However, it has not been reported whether high concentration formaldehyde may affect the proteome of a cell. If it does, a false positive or negative result may be obtained. In the present study, we have applied proteomic techniques to compare the proteomes between high concentration formaldehyde-treated and untreated control cells. The results have shown that high concentration formaldehyde can change the levels of a number of cellular proteins shortly following its permeation into the cells. Therefore, a formaldehyde-treated blank or repeated experiments are recommended for any experiment to exclude false positive or negative results when formaldehyde is technically used as a crosslinking tool.

Keywords formaldehyde; crosslink; proteome; LC-MS/MS

收稿日期: 2016-06-01 接受日期: 2016-06-17

中国科学院上海临床研究中心/徐汇区中心医院(批准号: 2014XHYYZX-05)资助的课题

*通讯作者。Tel: 13501979121, E-mail: jxwu@src.ac.cn

Received: June 1, 2016 Accepted: June 17, 2016

This work was supported by Xuhui Central Hospital, Shanghai Clinical Research Center, Chinese Academy of Sciences (Grant No.2014XHYYZX-05)

*Corresponding author. Tel: +86-13501979121, E-mail: jxwu@src.ac.cn

网络出版时间: 2016-08-31 16:17:06

URL: <http://www.cnki.net/kcms/detail/31.2035.Q.20160831.1617.014.html>

甲醛是一种具高反应活性的小分子有机物, 它可与蛋白质或DNA分子中的氨基或亚氨基等基团发生亲核加成反应, 经偶联邻近的伯胺, 使得蛋白质和蛋白质或蛋白质和DNA分子间通过共价键而交联^[1]。甲醛交联具备下述几个特性: (1)分子小, 可以瞬间通过细胞膜, 将蛋白质与蛋白质或蛋白质与DNA通过共价键而交联; (2)进入细胞后, 几乎即时阻止酶反应, 可以提供添加时瞬间蛋白质、DNA等分子的相互作用情形; (3)使用甘氨酸可以迅速终止由甲醛引发的蛋白质与蛋白质、蛋白质与DNA的交联反应; (4)一旦交联反应完成, 反应物可用于非生理条件下操作, 并保持结构完整性; (5)交联是可逆的, 只需加热样品即可逆转交联, 随后可分析复合体中的各个组分^[1]。

鉴于甲醛交联的这些特点, 它被广泛应用于蛋白质复合物中相互作用的蛋白质或DNA的鉴定分析, 包括蛋白质组学、组蛋白翻译后修饰、染色体免疫沉淀、染色体构象捕获等研究领域或研究技术。一般认为, 作为交联剂使用的高浓度甲醛可以在细胞尚未做出应激反应之前, 就通过细胞膜并瞬间将细胞中相互作用的蛋白质或DNA交联固定, 使这些生物大分子保持在自然生理状态之下^[1]。但是, 过去的研究表明, 与高浓度甲醛试剂不同, 日常环境中的低浓度甲醛可以导致DNA损伤和DNA修复等细胞反应^[2-3]。那么, 高浓度甲醛试剂在瞬间“固定”蛋白质等之前, 是否也会像低浓度甲醛那样, 引起细胞反应和蛋白质水平变化, 目前尚无相关研究报道。如果高浓度甲醛会在瞬间交联“固定”蛋白质等分子之前, 导致某些蛋白质水平发生变化, 而这些蛋白质正是被研究的对象, 就可能造成假阳性或假阴性的结果。本研究以人白血病T淋巴Jurkat细胞作为模型, 运用蛋白质组学技术对比分析了甲醛处理和未经甲醛处理的细胞蛋白质谱, 结果显示, 作为交联剂使用的高浓度甲醛会改变细胞蛋白质谱。

1 材料与方法

1.1 细胞的培养和甲醛交联

人类白血病T淋巴Jurkat细胞株从中国科学院上海生命科学研究院细胞资源中心购得; RPMI 1640培养基和胎牛血清购于Gibco公司; 甲醛购于Sigma公司; 甘氨酸购于上海朝瑞生物科技有限公司; 十二水磷酸氢二钠($\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$)、二水磷

酸二氢钠($\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$)、氯化钠(NaCl)和氯化钾(KCl)均为国药集团化学试剂有限公司产品; Eppendorf冷冻离心机(5427R型)购于Eppendorf公司; 细胞培养箱(3111型)购于ThermoFisher Scientific公司。

细胞培养: Jurkat细胞用1640培养基(含10%胎牛血清)在37 °C、5% CO_2 培养箱中培养, 处于对数生长期的细胞用于下述各实验。

甲醛固定: 向培养基中加入37%甲醛至终浓度为1%, 室温孵育10 min后, 向培养基和甲醛混合液中加入2.5 mol/L甘氨酸至终浓度为125 mmol/L, 室温孵育5 min以终止甲醛交联反应。将上述终止了甲醛交联反应的培养基离心3 min(500 $\times g$)。弃上清, 加入1 $\times$ PBS缓冲液(8 mmol/L $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$ 、2.1 mmol/L $\text{NaH}_2\text{PO}_4 \cdot 2\text{H}_2\text{O}$ 、137 mmol/L NaCl 、2.7 mmol/L KCl)重悬细胞, 离心3 min(500 $\times g$), 重复1次, 放置冰上备用。

1.2 细胞全蛋白质组的提取、酶解

裂解和蛋白质提取相关试剂: 二硫苏糖醇(DTT)、碘乙酰胺、丙酮、碳酸氢铵(NH_4HCO_3)、十二烷基磺酸钠(SDS)均购自Sigma-Aldrich公司; Bradford工作液和BSA试剂盒购自Bio-Rad公司; 混合蛋白酶抑制剂(Complete™ Tablets EDTA-free)购自Roche公司; KCl 、六水合氯化镁($\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$)、三羟甲基氨基甲烷(Tris-base)、浓盐酸(HCl)均购自国药集团化学试剂有限公司; 测序级溶液胰蛋白酶(Seq-Trypsin)购自Promega公司; 5%乙腈和0.1%甲酸购自于ThermoFisher Scientific公司。Eppendorf冷冻离心机(5427R型)购自Eppendorf公司; Beckman离心机(A1330011型)购自Beckman公司; 真空冷冻干燥机为德国Christ产品(ALPHA 1-2型); 超声破碎仪(JY92-2D型)购自宁波新芝生物科技公司。

蛋白质组的提取采用丙酮沉淀法^[4]。具体过程为: 加入800 μL 蛋白裂解液(2% SDS, 20 mmol/L HEPAS, pH7.9)将细胞充分悬浮均匀。采用热复性的方法解除甲醛对蛋白质的共价交联: 99 °C孵育60 min, 热复性解除甲醛对蛋白质的共价交联。4 °C超声(100 W, 超声3 s, 间隔5 s, 30个循环)促进蛋白质溶解, 然后4 °C离心30 min(18 000 $\times g$), 将上清转移至一个新的离心管。加入DTT至终浓度为10 mmol/L, 37 °C孵育1 h, 加入碘乙酰胺至终浓度20 mmol/L, 室温避光反应1 h。反应结束后, 在溶液中加入4倍体积的冰丙酮, 混匀, 置于-20 °C过夜沉淀。之后4 °C离心10 min(8 000 $\times g$), 保留沉淀用冰丙酮洗涤2次, 通风

干燥, 加入200 μ L的8 mol/L尿素溶液, 并用超声促进溶解沉淀。

蛋白质酶解: 用Bradford法定量提取物中蛋白质浓度^[5]。取500 μ g蛋白质, 加入7倍体积的50 mmol/L NH_4HCO_3 溶液, 然后按照质量比1:50的比例加入测序级胰蛋白酶(Seq-Trypsin), 37 $^\circ\text{C}$ 孵育过夜。酶解之后的肽段溶液用SPE C18小柱(Waters公司)脱盐, 脱盐后的肽段溶液采用低温冻干, -80 $^\circ\text{C}$ 保存, 备用。

1.3 LC-MS/MS质谱分析

液质联用分析使用的纳升级液相色谱仪为 EASY-nLC 1000(ThermoFisher Scientific公司), 质谱仪为超高分辨三合一质谱仪Orbitrap Fusion(ThermoFisher Scientific公司)。冻干的肽段重溶于含5%乙腈、0.1%甲酸的水溶液中, 上载至富集柱(75 μm i.d., 2 cm, C18, 3 μm , 100 $\text{\AA}$, Dionex公司), 短暂的在线脱盐之后, 以300 nL/min的流速, 用C18分析柱(75 μm i.d., 10.2 cm, C18, 3 μm , 120 $\text{\AA}$, New Objective公司)进行在线液质联用分析。液相的梯度为: 100 min内B相(99.9%乙腈、0.1%甲酸)线性地从5%增至50%。具体质谱条件如下: 纳升电喷雾(nano-electrospray)电压为2.1 kV; 正离子采集模式, 使用Top speed的数据依赖的自动化采集模式(data-dependent mode)实现在3 s内1次一级和若干个排名靠前的前体离子的二级质谱的自动采集, 其中一级质谱扫描范围为 m/z 350~1500, 分辨率为60 000(atm/z 400); 二级质谱分辨率为15 000(atm/z 400), 母离子窗口为2 Da, 串联质谱碰撞能量为35%归一化HCD能量, 动态排除为1次重复计数, 60 s持续排除时间。质量数据由XCalibur软件采集处理。

1.4 数据分析

质谱产生的原始文件用MaxQuant(1.5.3.8版本)^[6]进行数据库搜库、数据整合、假阳性控制以及非标记定量。搜库参数基本采用默认参数: 固定修饰为半胱氨酸的乙酰胺化; 可变修饰包括甲硫氨酸的氧化和蛋白质N-端乙酰化; 酶切特异性设置为胰蛋白酶, 允许2个缺失酶切位点; 母离子最大质量偏差4.5 ppm, 子离子最大质量偏差20 ppm。数据库搜索采用MaxQuant内置的Andromeda引擎进行^[7], 数据库为人的Uniprot蛋白质库(2013年10月版本), 其中还包含常见的污染物蛋白质序列和数据库中所有蛋白质的反向序列。数据库搜索产生的结果用MaxQuant进行筛选整合, 控制肽段和蛋白质的错

误发现率(False Discovery Rate, FDR)均小于1%, 且要求肽段的长度大于等于7个氨基酸。非标记定量采用MaxQuant内置的MaxLFQ^[8]算法进行, 采用默认参数, MaxQuant会根据得到的初步定量结果在各个原始文件之间进一步根据总离子强度进行归一化。

1.5 功能注释

在线Ingenuity Pathways Analysis(IPA)软件(www.ingenuity.com/products/ipa)用于分析和构建这些差异表达蛋白质归属的亚细胞器和细胞通路。

2 结果

2.1 分析数据的重复性

我们对甲醛处理过的细胞或者对照细胞都做了两次生物学重复实验, 每次生物学实验做了三次LC-MS/MS分析。我们对两次生物学重复实验获得的蛋白质定量强度进行相关分析(图1)。无论是甲醛处理或是对照细胞样品的两次生物学重复实验的蛋白质定量强度数据的相关系数 r 均大于0.95, 表明两次生物学实验具有很好的重复性。

2.2 甲醛固定细胞与对照细胞的蛋白质组差异

我们分别检测了甲醛处理的细胞和未经甲醛处理的对照细胞蛋白质组, 图2显示了两次生物学重复实验的结果。第一次实验(图2A), 在甲醛处理细胞中, 鉴定到3 699个蛋白质, 对照细胞鉴定到3 750个蛋白质, 其中在甲醛处理和对照细胞内皆能鉴定到的蛋白质为3 536个, 仅能在甲醛处理细胞内鉴定到的蛋白质为163个, 仅能在对照细胞内鉴定到的蛋白质为214个。第二次实验(图2B), 甲醛处理细胞鉴定到蛋白质为3 529个, 对照细胞鉴定到蛋白质为3 651个, 其中在甲醛处理过和对照细胞皆能检测到的蛋白质为3 315个, 仅能在甲醛处理细胞检测到的蛋白质为214个, 仅能在对照细胞检测到的蛋白质为336个。

综合分析两次重复实验结果, 结果如下(图2C)。第一, 38个蛋白质可以在两次甲醛处理细胞中重复检测到, 但是对照细胞两次都未检测到, 换言之, 这些蛋白质在甲醛处理细胞与对照细胞之间存在差异表达。第二, 55个蛋白质可以在两次对照处理细胞重复检测到, 但两次甲醛处理细胞都未检测到, 显示这些蛋白质在两组细胞间也存在差异表达。第三, 3 142个蛋白质在两次试验中, 不论甲醛处理与

否, 都能鉴定到。对比甲醛处理细胞和对照细胞, 在这3 142个蛋白质中, 有43个浓度发生显著差异, 表明这些蛋白质水平发生了变化(图2C括号内数字)。

归纳上述三种情况, 在甲醛处理细胞中, 共有136(38+55+43)个蛋白质水平出现变化。表1~表3具体列出了这些蛋白质。

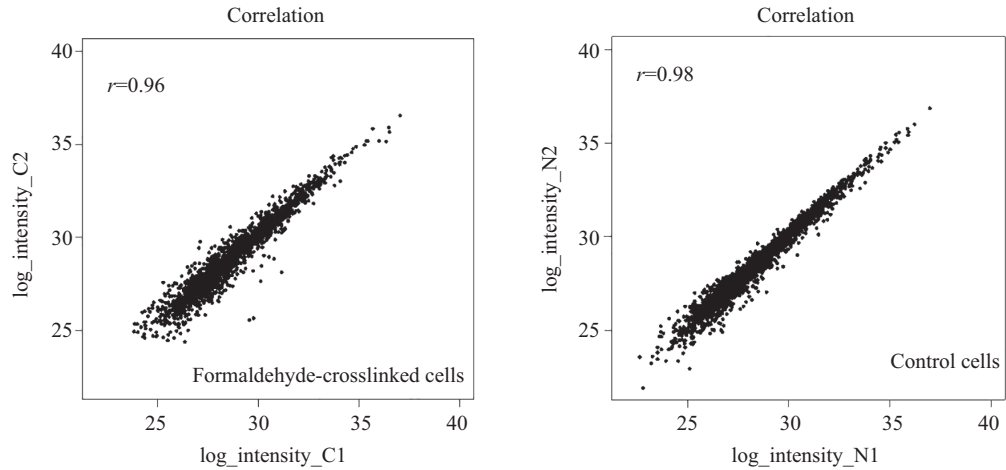


图1 两次LC-MS/MS定量结果相关分析
Fig.1 Correlation analyses of two repeated quantitative LC-MS/MS results

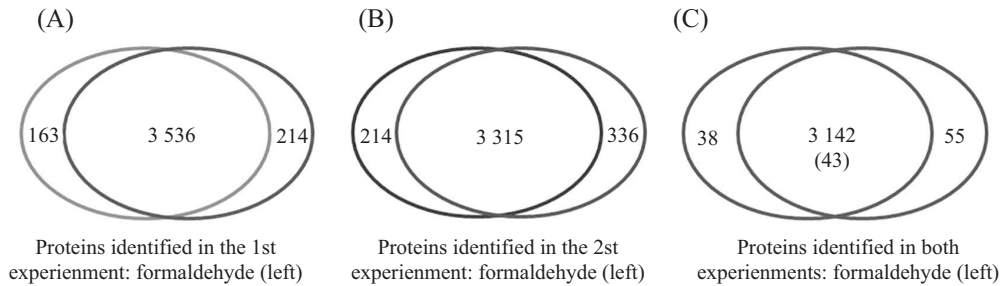


图2 两次甲醛交联和对照细胞实验中鉴定到的蛋白
Fig.2 Proteins identified in two repeated experiments of formaldehyde-treated and control cells

表1 两次甲醛出现但两次对照都不出现的蛋白

Table 1 Proteins identified in both formaldehyde treatment experiments but neither control experiments

基因名称 Gene names	蛋白名称 Protein names	基因名称 Gene names	蛋白名称 Protein names
<i>MTFP1</i>	Mitochondrial fission process protein 1	<i>RRM2B</i>	Ribonucleoside-diphosphate reductase subunit M2 B
<i>TRIM24</i>	Transcription intermediary factor 1-alpha	<i>HAUS6</i>	HAUS augmin-like complex subunit 6
<i>CTAGE5</i>	cTAGE family member 5; cTAGE family member 6; cTAGE family member 15	<i>TMED4</i>	Transmembrane emp24 domain-containing protein 4
<i>TBC1D4</i>	TBC1 domain family member 4	<i>NIT1</i>	Nitrilase homolog 1
<i>TRAPPC2P1</i>	Trafficking protein particle complex subunit 2 protein TRAPPC2P1; trafficking protein particle complex subunit 2	<i>AVL9</i>	Late secretory pathway protein AVL9 homolog
<i>ECE1</i>	Endothelin-converting enzyme 1	<i>ARID1B</i>	AT-rich interactive domain-containing protein 1B
<i>TAP1</i>	Antigen peptide transporter 1	<i>SHKBP1</i>	SH3KBP1-binding protein 1
<i>TRIP4</i>	Activating signal cointegrator 1	<i>ERGIC1</i>	Endoplasmic reticulum-Golgi intermediate compartment protein 1
<i>FOCAD</i>	Focadhesin	<i>MOB3A</i>	MOB kinase activator 3A
<i>RSBN1</i>	Round spermatid basic protein 1	<i>NGLY1</i>	Peptide-N(4)-(N-acetyl-beta-glucosaminyl) asparagine amidase
<i>LYPLAL1</i>	Lysophospholipase-like protein 1	<i>EXOC2</i>	Exocyst complex component 2
		<i>NDNL2</i>	Melanoma-associated antigen G1

(续表1)

基因名称	蛋白名称	基因名称	蛋白名称
Gene names	Protein names	Gene names	Protein names
<i>UHRF2</i>	E3 ubiquitin-protein ligase UHRF2	<i>HDAC6</i>	Histone deacetylase 6
<i>AGAP2</i>	Arf-GAP with GTPase, ANK repeat and PH domain-containing protein 2	<i>CDC23</i>	Cell division cycle protein 23 homolog
<i>REXO4</i>	RNA exonuclease 4	<i>FBXW11</i>	F-box/WD repeat-containing protein 11
<i>POLR3F</i>	DNA-directed RNA polymerase III subunit RPC6	<i>DHRS7</i>	Dehydrogenase/reductase SDR family member 7
<i>UBE2Z</i>	Ubiquitin-conjugating enzyme E2 Z	<i>SBDS</i>	Ribosome maturation protein SBDS
<i>TMEM38B</i>	Trimeric intracellular cation channel type B	<i>MTF2</i>	Metal-response element-binding transcription factor 2
<i>PAK1IP1</i>	p21-activated protein kinase-interacting protein 1	<i>TAF6L</i>	TAF6-like RNA polymerase II p300/CBP-associated factor-associated factor 65 kDa subunit 6L
<i>RRN3</i>	RNA polymerase I-specific transcription initiation factor RRN3		

表2 两次对照出现但两次甲醛都不出现的蛋白

Table 2 Proteins identified in both control but neither formaldehyde treatment experiments

基因名称	蛋白名称	基因名称	蛋白名称
Gene names	Protein names	Gene names	Protein names
<i>CASC3</i>	Protein CASC3	<i>DPP9</i>	Dipeptidyl peptidase 9
<i>EIF1B</i>	Eukaryotic translation initiation factor 1b	<i>FOXP4</i>	Forkhead box protein P4
<i>NDUFA7</i>	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 7	<i>CCDC117</i>	Coiled-coil domain-containing protein 117
<i>MT-ND6</i>	NADH-ubiquinone oxidoreductase chain 6	<i>C9orf40</i>	Uncharacterized protein C9orf40
<i>DBT</i>	Lipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex, mitochondrial	<i>RNF169</i>	E3 ubiquitin-protein ligase RNF169
<i>CD53</i>	Leukocyte surface antigen CD53	<i>SLC30A7</i>	Zinc transporter 7
<i>ERCC3</i>	TFIIH basal transcription factor complex helicase XPB subunit	<i>MYEOV2</i>	Myeloma-overexpressed gene 2 protein
<i>USF1</i>	Upstream stimulatory factor 1	<i>GTF3A</i>	Transcription factor IIIA
<i>CD1C</i>	T-cell surface glycoprotein CD1c	<i>UBE2F</i>	NEDD8-conjugating enzyme UBE2F
<i>BLVRB</i>	Flavin reductase (NADPH)	<i>DTNBP1</i>	Dysbindin
<i>SLC6A6</i>	Sodium- and chloride-dependent taurine transporter	<i>OTUD5</i>	OTU domain-containing protein 5
<i>ZFP36L2</i>	Zinc finger protein 36, C3H1 type-like 2	<i>FAHD2A</i>	Fumarylacetoacetate hydrolase domain-containing protein 2A
<i>PPP3R1</i>	Calcineurin subunit B type 1	<i>CBWD1</i>	COBW domain-containing protein 1; COBW domain-containing protein 5; COBW domain-containing protein 3; COBW domain-containing protein 6; Putative COBW domain-containing protein 7; COBW domain-containing protein 2
<i>CENPC1</i>	Centromere protein C 1	<i>ANAPC13</i>	Anaphase-promoting complex subunit 13
<i>NFATC3</i>	Nuclear factor of activated T-cells, cytoplasmic 3	<i>MECR</i>	Trans-2-enoyl-CoA reductase, mitochondrial
<i>TRIM32</i>	E3 ubiquitin-protein ligase TRIM32	<i>MND1</i>	Meiotic nuclear division protein 1 homolog
<i>ARFRP1</i>	ADP-ribosylation factor-related protein 1	<i>EPT1</i>	Ethanolaminephosphotransferase 1
<i>ZMYM3</i>	Zinc finger MYM-type protein 3	<i>LHPP</i>	Phospholysine phosphohistidine inorganic pyrophosphate phosphatase
<i>ACAP2</i>	Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 2	<i>VIPAS39</i>	Spermatogenesis-defective protein 39 homolog
<i>CYTH1</i>	Cytohesin-1	<i>PATZ1</i>	POZ-, AT hook-, and zinc finger-containing protein 1
<i>BOLA3</i>	Bola-like protein 3	<i>BLOC1S4</i>	Biogenesis of lysosome-related organelles complex 1 subunit 4
<i>ALG10B</i>	Putative Dol-P-Glc:Glc(2)Man(9)GlcNAc(2)-PP-Dol alpha-1,2-glucosyltransferase; Dol-P-Glc:Glc(2)Man(9)GlcNAc(2)-PP-Dol alpha-1,2-glucosyltransferase	<i>BCAP29</i>	B-cell receptor-associated protein 29
<i>THOC7</i>	THO complex subunit 7 homolog	<i>SYNRG</i>	Synergism gamma
<i>C3orf17</i>	Uncharacterized protein C3orf17	<i>ANKRD26</i>	Ankyrin repeat domain-containing protein 26
<i>KDELC1</i>	KDEL motif-containing protein 1	<i>DICER1</i>	Endoribonuclease dicer
<i>C11orf57</i>	Uncharacterized protein C11orf57		

(续表2)

基因名称	蛋白名称	基因名称	蛋白名称
Gene names	Protein names	Gene names	Protein names
<i>TMED3</i>	Transmembrane emp24 domain-containing protein 3	<i>CABIN1</i>	Calcineurin-binding protein cabin-1
<i>COL4A3BP</i>	Collagen type IV alpha-3-binding protein	<i>CAPN7</i>	Calpain-7

表3 两次甲醛以及两次对照均出现但含量有显著变化的蛋白

Table 3 Significantly quantitatively changed proteins identified in both formaldehyde treated and both control experiments

基因名称	蛋白名称	基因名称	蛋白名称
Gene names	Protein names	Gene names	Protein names
<i>KRBA2</i>	60S ribosomal protein L26	<i>UTP3</i>	Something about silencing protein 10
<i>UNC45A</i>	Protein unc-45 homolog A	<i>RFK</i>	Riboflavin kinase
<i>SCD</i>	Acyl-CoA desaturase	<i>GABPA</i>	GA-binding protein alpha chain
<i>H2AFV</i>	Histone H2A.V; Histone H2A.Z	<i>ZNF622</i>	Zinc finger protein 622
<i>ICAM2</i>	Intercellular adhesion molecule 2	<i>ILVBL</i>	Acetolactate synthase-like protein
<i>LAMP2</i>	Lysosome-associated membrane glycoprotein 2	<i>ANKFY1</i>	Ankyrin repeat and FYVE domain-containing protein 1
<i>MAGED1</i>	Melanoma-associated antigen D1	<i>TARBP1</i>	Probable methyltransferase TARBP1
<i>H3F3A</i>	Histone H3.3	<i>MRPL28</i>	39S ribosomal protein L28, mitochondrial
<i>HEATR3</i>	HEAT repeat-containing protein 3	<i>PHF6</i>	PHD finger protein 6
<i>PPP4C</i>	Serine/threonine-protein phosphatase 4 catalytic subunit	<i>TUBB6</i>	Tubulin beta-6 chain
<i>DPY30</i>	Protein dpy-30 homolog	<i>CALM1</i>	Calmodulin
<i>ARL1</i>	ADP-ribosylation factor-like protein 1	<i>XRCC1</i>	DNA repair protein XRCC1
<i>TTN</i>	Titin	<i>UFM1</i>	Ubiquitin-fold modifier 1
	DNA-directed RNA polymerase II subunit RPB11-b2;	<i>UBE2G1</i>	Ubiquitin-conjugating enzyme E2 G1
<i>POLR2J3</i>	DNA-directed RNA polymerase II subunit RPB11-b1;	<i>ATP5I</i>	ATP synthase subunit e, mitochondrial
	DNA-directed RNA polymerase II subunit RPB11-a	<i>RPL19</i>	60S ribosomal protein L19
<i>PRKAG1</i>	5-AMP-activated protein kinase subunit gamma-1	<i>LSM1</i>	U6 snRNA-associated Sm-like protein LSm1
<i>AGL</i>	Glycogen debranching enzyme; 4-alpha-glucanotransferase; Amylo-alpha-1,6-glucosidase	<i>RPL26L1</i>	60S ribosomal protein L26-like 1
<i>ARHGAP4</i>	Rho GTPase-activating protein 4	<i>ADRM1</i>	Proteasomal ubiquitin receptor ADRM1
<i>AQR</i>	Intron-binding protein aquarius	<i>PTMA</i>	Prothymosin alpha; Prothymosin alpha, N-terminally processed; Thymosin alpha-1
<i>PEX14</i>	Peroxisomal membrane protein PEX14	<i>ZFHX3</i>	Zinc finger homeobox protein 3
<i>GTF3C4</i>	General transcription factor 3C polypeptide 4	<i>NAGK</i>	N-acetyl-D-glucosamine kinase
<i>SCCPDH</i>	Saccharopine dehydrogenase-like oxidoreductase		

为了观察甲醛处理引起的蛋白质水平差异是否具有亚细胞器特异性, 我们使用IPA在线软件对这136个差异蛋白质进行了亚细胞器定位分析。定位分析表明: 差异蛋白质中有53个胞质蛋白、51个核蛋白、13个膜蛋白、1个细胞外的分泌蛋白和14个其他蛋白质, 另有4个蛋白质未找到匹配(图3)。该定位结果表明: 甲醛交联所引起的细胞蛋白质组变化并不具有明显的亚细胞器特异性。有趣的是, 我们发现了1个定位于细胞外的分泌蛋白, 该蛋白质仅能在甲醛处理的细胞样品中鉴定到。这可能是其他蛋白质, 另有4个蛋白质未找到匹配(图3)。该定位结果表明: 甲醛交联所引起的细胞蛋白质组变化并不具有明显的亚细胞器特异性。有趣的是, 我们发现

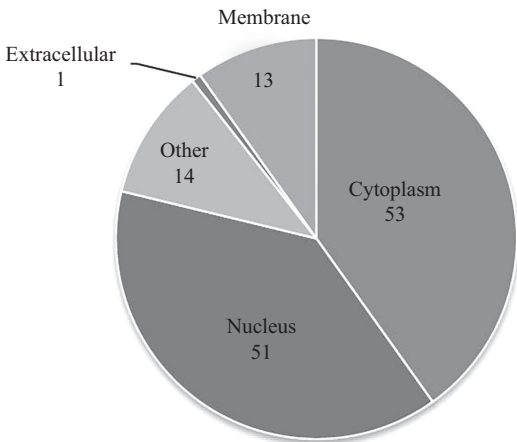


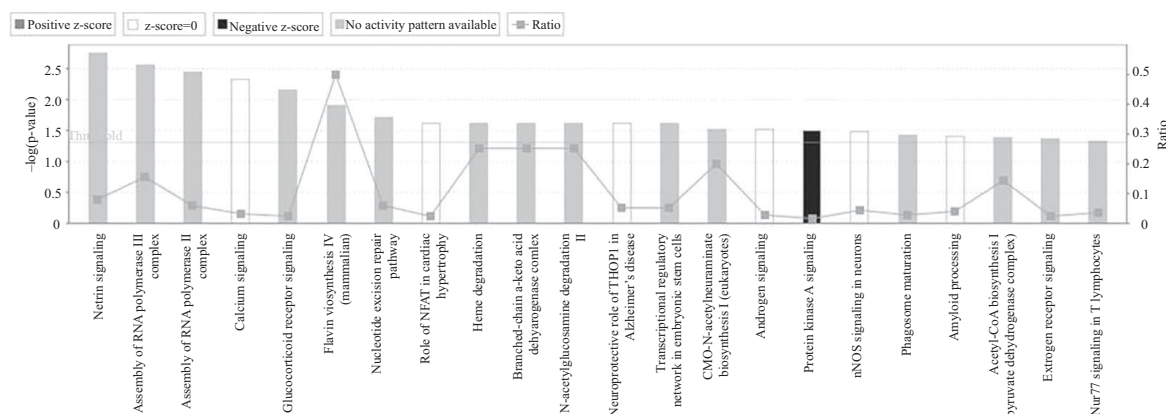
图3 全细胞差异蛋白的亚细胞器定位分析
Fig.3 Analysis of cellular locations of whole-cell differentially expressed proteins

了1个定位于细胞外的分泌蛋白, 该蛋白质仅能在甲醛处理的细胞样品中鉴定到。这可能是由于甲醛将处于细胞膜附近的分泌蛋白共价交联到细胞膜上, 使得分泌蛋白不能在后续的细胞洗涤过程中被洗脱掉。

为了观察甲醛处理引起的差异蛋白质是否具有细胞通路特异性, 我们同样采用IPA分析了这些差

异蛋白质, 结果表明, 差异蛋白质涉及22条细胞通路(图4), 其中蛋白激酶A信息通路受到了抑制。

考虑到使用甲醛作为交联工具的染色体免疫沉淀、染色质构象捕捉等技术广泛应用于DNA与蛋白质相互作用方面的研究, 我们用IPA软件分析了细胞核内的51个差异蛋白质(图5), 这些蛋白质共涉及16个细胞通路, 包括DNA转录、细胞周期和DNA修

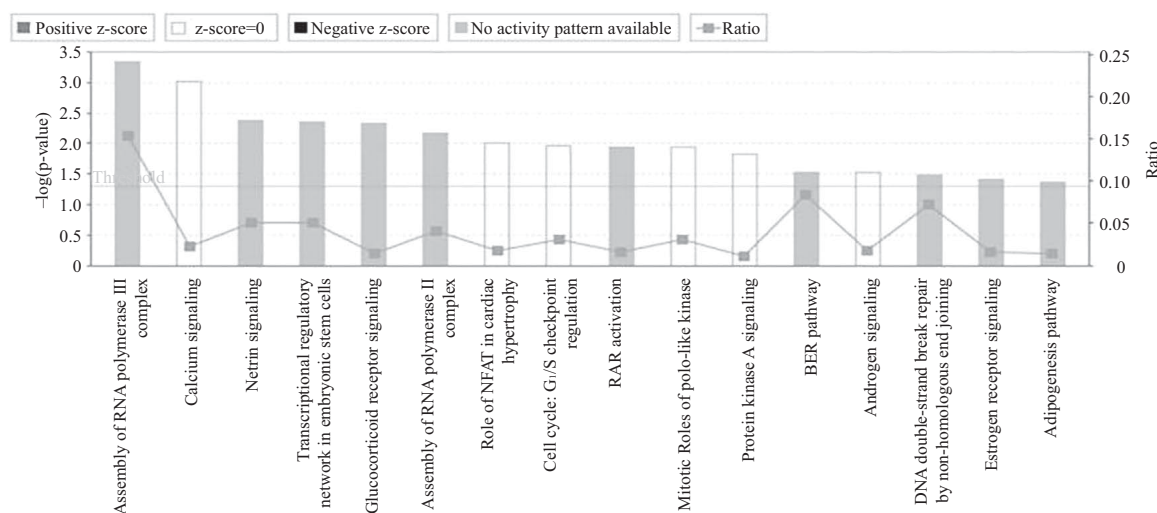


X轴是差异蛋白所涉及到的通路的名称, 柱的高度代表涉及通路的可信度; 柱子颜色: 深灰色代表通路被抑制, 白色代表既不被激活也不被抑制, 浅灰色代表不知道该通路是否被激活或抑制; 线条为ratio, 代表差异蛋白中含有的某个通路蛋白数占该通路全部蛋白数的比例。

X axis represents the names of the cellular pathways associated with whole-cell differentially expressed proteins. The heights of the columns represent the reliability of the associated pathways. The dark columns represent the suppressed pathways. The white ones, neither activated nor suppressed pathways. The light ones, status unknown. The line represents the ratio between the number of the differentially expressed proteins and the number of all proteins in a pathway.

图4 全细胞差异蛋白的相关通路分析

Fig.4 Analysis of cellular pathways associated with whole-cell differentially expressed proteins



X轴是差异蛋白所涉及到的通路的名称, 柱的高度代表涉及通路的可信度; 柱子颜色: 白色代表通路既不被激活也不被抑制, 浅灰色代表不知道该通路是否被激活或抑制; 线条为ratio, 代表差异蛋白中含有的某个通路蛋白数占该通路全部蛋白数的比例。

X axis represents the names of the cellular pathways associated with nuclear differentially expressed proteins. The heights of the columns represent the reliability of the associated pathways. The white columns neither activated nor suppressed pathways. The light ones, status unknown. The line represents the ratio between the number of the differentially expressed proteins and the number of all proteins in a pathway.

图5 核内蛋白的相关通路分析

Fig.5 Analysis of cellular pathways associated with nuclear differentially expressed proteins

复等通路。

除了上述两次实验重复性高的蛋白质,我们也分析了重复性差的蛋白质。这些蛋白质更可能与实验条件、细胞状态、细胞培养环境等因素相关。在

甲醛处理细胞的两次重复实验中仅出现一次的蛋白质,即无重复性的蛋白质共有614个(表4),所涉及的分子功能和细胞通路范围广泛,比如调节基因表达、DNA损伤修复、组蛋白翻译后修饰等。

表4 仅在一次甲醛处理细胞内发现的蛋白(614个)

Table 4 Proteins identified only once in either formaldehyde treatment experiment

基因名称 Gene names	蛋白名称 Protein names	基因名称 Gene names	蛋白名称 Protein names
<i>GOLGA3</i>	Golgin subfamily A member 3	<i>KLF16</i>	Krueppel-like factor 16
<i>KNOP1</i>	Lysine-rich nucleolar protein 1	<i>ITPA</i>	Inosine triphosphate pyrophosphatase
<i>TAF4</i>	Transcription initiation factor TFIID subunit 4	<i>SRXN1</i>	Sulfiredoxin-1
<i>FRYL</i>	Protein furry homolog-like	<i>WDR55</i>	WD repeat-containing protein 55
<i>DLGAP5</i>	Disks large-associated protein 5	<i>CNOT10</i>	CCR4-NOT transcription complex subunit 10
<i>ATF7</i>	Cyclic AMP-dependent transcription factor ATF-7	<i>BSDC1</i>	BSD domain-containing protein 1
<i>CDC7</i>	Cell division cycle 7-related protein kinase	<i>SLAIN2</i>	SLAIN motif-containing protein 2
<i>WIPF1</i>	WAS/WASL-interacting protein family member 1	<i>C19orf25</i>	UPF0449 protein C19orf25
<i>LAT</i>	Linker for activation of T-cells family member 1	<i>MKLN1</i>	Muskelin
<i>DUSP11</i>	RNA/RNP complex-1-interacting phosphatase	<i>FAM21C; FAM21A</i>	WASH complex subunit FAM21C; WASH complex subunit FAM21A
<i>GMNN</i>	Geminin	<i>ATP6V1D</i>	V-type proton ATPase subunit D
<i>HAUS5</i>	HAUS augmin-like complex subunit 5	<i>WASF2</i>	Wiskott-Aldrich syndrome protein family member 2
<i>ZWINT</i>	ZW10 interactor	<i>MTHFS</i>	5-formyltetrahydrofolate cyclo-ligase
<i>TAL1</i>	T-cell acute lymphocytic leukemia protein 1	<i>KIAA1467</i>	Uncharacterized protein KIAA1467
<i>EPS15</i>	Epidermal growth factor receptor substrate 15	<i>ZNF316</i>	Zinc finger protein 316
<i>NUMB</i>	Protein numb homolog	<i>CCDC85C</i>	Coiled-coil domain-containing protein 85C
<i>PEX5</i>	Peroxisomal targeting signal 1 receptor	<i>PTGR2</i>	Prostaglandin reductase 2
<i>MRPS11</i>	28S ribosomal protein S11, mitochondrial	<i>DNASE2</i>	Deoxyribonuclease-2-alpha
<i>RING1</i>	E3 ubiquitin-protein ligase RING1	<i>SLC33A1</i>	Acetyl-coenzyme A transporter 1
<i>MYO9B</i>	Unconventional myosin-IXb	<i>ZNF609</i>	Zinc finger protein 609
<i>VEZF1</i>	Vascular endothelial zinc finger 1	<i>FYB</i>	FYN-binding protein
<i>PEA15</i>	Astrocytic phosphoprotein PEA-15	<i>BIRC5</i>	Baculoviral IAP repeat-containing protein 5
<i>MAP1S</i>	Microtubule-associated protein 1S; MAP1S heavy chain; MAP1S light chain	<i>ABCC4</i>	Multidrug resistance-associated protein 4
<i>MPRIIP</i>	Myosin phosphatase Rho-interacting protein	<i>TAPBP</i>	Tapasin
<i>RPS19BP1</i>	Active regulator of SIRT1	<i>PJA2</i>	E3 ubiquitin-protein ligase Praja-2
<i>FNBP4</i>	Formin-binding protein 4	<i>EMC8</i>	ER membrane protein complex subunit 8
<i>PDDC1</i>	Parkinson disease 7 domain-containing protein 1	<i>HTRA2</i>	Serine protease HTRA2, mitochondrial
<i>BOD1L1</i>	Biorientation of chromosomes in cell division protein 1-like 1	<i>SORBS3</i>	Vinexin
<i>HNRNPLL</i>	Heterogeneous nuclear ribonucleoprotein L-like	<i>PFKFB2</i>	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 2; 6-phosphofructo-2-kinase; fructose-2,6-bisphosphatase
<i>C18orf25</i>	Uncharacterized protein C18orf25	<i>RNASEH1</i>	Ribonuclease H1
<i>REPS1</i>	RalBP1-associated Eps domain-containing protein 1	<i>COQ9</i>	Ubiquinone biosynthesis protein COQ9, mitochondrial
<i>PDLIM5</i>	PDZ and LIM domain protein 5	<i>HMMR</i>	Hyaluronan mediated motility receptor
<i>PPP1R9B</i>	Neurabin-2	<i>CPD</i>	Carboxypeptidase D
<i>PAFAH2</i>	Platelet-activating factor acetylhydrolase 2, cytoplasmic	<i>UR11</i>	Unconventional prefoldin RPB5 interactor 1
<i>DPH2</i>	Diphthamide biosynthesis protein 2	<i>CCNB2</i>	G2/mitotic-specific cyclin-B2
<i>NUDT16L1</i>	Protein syndesmos	<i>ETHE1</i>	Persulfide dioxygenase ETHE1, mitochondrial

(续表4)

基因名称 Gene names	蛋白名称 Protein names	基因名称 Gene names	蛋白名称 Protein names
<i>NRAS</i>	GTPase NRas	<i>NCOA6</i>	Nuclear receptor coactivator 6
<i>MBP</i>	Myelin basic protein	<i>TAF7</i>	Transcription initiation factor TFIID subunit 7
<i>EPHX1</i>	Epoxide hydrolase 1	<i>NEDD8</i>	NEDD8
<i>CD7</i>	T-cell antigen CD7	<i>HLA-B</i>	HLA class I histocompatibility antigen, B-73 alpha chain
<i>FDX1</i>	Adrenodoxin, mitochondrial	<i>EML3</i>	Echinoderm microtubule-associated protein-like 3
<i>MYB</i>	Transcriptional activator Myb	<i>BCL7A</i>	B-cell CLL/lymphoma 7 protein family member A
<i>PYGM</i>	Glycogen phosphorylase, muscle form	<i>PDCD4</i>	Programmed cell death protein 4
<i>UBL4A</i>	Ubiquitin-like protein 4A	<i>PSMG4</i>	Proteasome assembly chaperone 4
<i>TRIM27</i>	Zinc finger protein RFP	<i>ZDHHC20</i>	Probable palmitoyltransferase ZDHHC20
<i>PGAM2</i>	Phosphoglycerate mutase 2	<i>DIEXF</i>	Digestive organ expansion factor homolog
<i>CHN1</i>	N-chimaerin	<i>ZFAND6</i>	AN1-type zinc finger protein 6
<i>PPP3CB</i>	Serine/threonine-protein phosphatase 2B catalytic subunit beta isoform	<i>FAM117B</i>	Protein FAM117B
<i>ZNF24</i>	Zinc finger protein 24	<i>CEP85</i>	Centrosomal protein of 85 kDa
<i>PTMS</i>	Parathymosin	<i>CCDC174</i>	Coiled-coil domain-containing protein 174
<i>CALML3</i>	Calmodulin-like protein 3	<i>TRIM65</i>	Tripartite motif-containing protein 65
<i>SHC1</i>	SHC-transforming protein 1	<i>SAMD1</i>	Atherin
<i>RBMS1; RBMS3</i>	RNA-binding motif, single-stranded-interacting protein 1; RNA-binding motif, single-stranded-interacting protein 3	<i>MLF1IP</i>	Centromere protein U
<i>GBP1</i>	Interferon-induced guanylate-binding protein 1	<i>MCM10</i>	Protein MCM10 homolog
<i>KRT9</i>	Keratin, type I cytoskeletal 9	<i>RAB11FIP2; RAB11FIP1</i>	Rab11 family-interacting protein 2; Rab11 family-interacting protein 1
<i>CETN2</i>	Centrin-2	<i>MOB1B</i>	MOB kinase activator 1B
<i>GRK6; GRK5</i>	G protein-coupled receptor kinase 6; G protein-coupled receptor kinase 5	<i>TRAPPC11</i>	Trafficking protein particle complex subunit 11
<i>GCLM</i>	Glutamate—cysteine ligase regulatory subunit	<i>KANSL1</i>	KAT8 regulatory NSL complex subunit 1
<i>FNTB</i>	Protein farnesyltransferase subunit beta	<i>SAPCD2</i>	Suppressor APC domain-containing protein 2
<i>TSC2</i>	Tuberin	<i>ZNF598</i>	Zinc finger protein 598
<i>APLP1</i>	Amyloid-like protein 1; C30	<i>GLCCI1</i>	Glucocorticoid-induced transcript 1 protein
<i>GTF2A1</i>	Transcription initiation factor IIA subunit 1; transcription initiation factor IIA alpha chain; transcription initiation factor IIA beta chain	<i>HOOK3</i>	Protein Hook homolog 3
<i>UBE2G2</i>	Ubiquitin-conjugating enzyme E2 G2	<i>PPFIBP1</i>	Liprin-beta-1
<i>RAB11A</i>	Ras-related protein Rab-11A	<i>MFN1</i>	Mitofusin-1
<i>BLOC1S1</i>	Biogenesis of lysosome-related organelles complex 1 subunit 1	<i>RHOT1</i>	Mitochondrial Rho GTPase 1
<i>IFI35</i>	Interferon-induced 35 kDa protein	<i>TMEM192</i>	Transmembrane protein 192
<i>MLLT1</i>	Protein ENL	<i>CYP4X1</i>	Cytochrome P450 4X1
<i>CKAP4</i>	Cytoskeleton-associated protein 4	<i>PNKD</i>	Probable hydrolase PNKD
<i>TFCP2; UBP1</i>	Alpha-globin transcription factor CP2; upstream-binding protein 1	<i>KRTCAP2</i>	Keratinocyte-associated protein 2
<i>CNTN1</i>	Contactin-1	<i>ENAH</i>	Protein enabled homolog
<i>MED21</i>	Mediator of RNA polymerase II transcription subunit 21	<i>C1orf131</i>	Uncharacterized protein C1orf131
<i>EIF4EBP2</i>	Eukaryotic translation initiation factor 4E-binding protein 2	<i>PAPD5</i>	PAP-associated domain-containing protein 5
<i>TUBB2A; TUBB2B</i>	Tubulin beta-2A chain; Tubulin beta-2B chain	<i>VPS37A</i>	Vacuolar protein sorting-associated protein 37A
<i>CACNA1C</i>	Voltage-dependent L-type calcium channel subunit alpha-1C	<i>ATL2; ATL1</i>	Atlastin-2; atlastin-1
		<i>MCFD2</i>	Multiple coagulation factor deficiency protein 2
		<i>CD99L2</i>	CD99 antigen-like protein 2
		<i>PARD3</i>	Partitioning defective 3 homolog
		<i>CACTIN</i>	Cactin
		<i>NUDT8</i>	Nucleoside diphosphate-linked moiety X motif 8, mitochondrial
		<i>SKA2</i>	Spindle and kinetochore-associated protein 2
		<i>CKAP2</i>	Cytoskeleton-associated protein 2
		<i>CCNB3</i>	G2/mitotic-specific cyclin-B3

(续表4)

基因名称 Gene names	蛋白名称 Protein names	基因名称 Gene names	蛋白名称 Protein names
<i>CASKIN2</i>	Caskin-2	<i>5-Mar</i>	E3 ubiquitin-protein ligase MARCH5
<i>OVC42</i>	Ovarian cancer-associated gene 2 protein	<i>MED29</i>	Mediator of RNA polymerase II transcription subunit 29
<i>NSMAF</i>	Protein FAN	<i>SIRT5</i>	NAD-dependent protein deacylase sirtuin-5, mitochondrial
<i>PIGK</i>	GPI-anchor transamidase	<i>KCTD5</i>	BTB/POZ domain-containing protein KCTD5
<i>MCUR1</i>	Mitochondrial calcium uniporter regulator 1	<i>MIS18A</i>	Protein Mis18-alpha
<i>AKT1S1</i>	Proline-rich AKT1 substrate 1	<i>GTSE1</i>	G ₂ and S phase-expressed protein 1
<i>CHCHD1</i>	Coiled-coil-helix-coiled-coil-helix domain-containing protein 1	<i>SMARCAL1</i>	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A-like protein 1
<i>DGCR14</i>	Protein DGCR14	<i>MED11</i>	Mediator of RNA polymerase II transcription subunit 11
<i>SIRT1</i>	NAD-dependent protein deacetylase sirtuin-1; SirtT1 75 kDa fragment	<i>ABI3</i>	ABI gene family member 3
<i>CCDC101</i>	SAGA-associated factor 29 homolog	<i>KANSL3</i>	KAT8 regulatory NSL complex subunit 3
<i>NSL1</i>	Kinetochore-associated protein NSL1 homolog	<i>DPM3</i>	Dolichol-phosphate mannosyltransferase subunit 3
<i>TMX3</i>	Protein disulfide-isomerase TMX3	<i>SLC25A13</i>	Calcium-binding mitochondrial carrier protein Aralar 2
<i>ZFYVE19</i>	Zinc finger FYVE domain-containing protein 19	<i>ZNF580</i>	Zinc finger protein 580
<i>PASK</i>	PAS domain-containing serine/threonine-protein kinase	<i>ACAD8</i>	Isobutyryl-CoA dehydrogenase, mitochondrial
<i>ERGIC2</i>	Endoplasmic reticulum-Golgi intermediate compartment protein 2	<i>ACSL5; ACSL1</i>	Long-chain-fatty-acid-CoA ligase 5; long-chain-fatty-acid-CoA ligase 1
<i>RPS6KC1</i>	Ribosomal protein S6 kinase delta-1	<i>ING4</i>	Inhibitor of growth protein 4
<i>UBL7</i>	Ubiquitin-like protein 7	<i>COA3</i>	Cytochrome c oxidase assembly protein 3 homolog, mitochondrial
<i>ADO</i>	2-aminoethanethiol dioxygenase	<i>POLR1D</i>	DNA-directed RNA polymerases I and III subunit RPAC2
<i>CDC43</i>	Cell division cycle-associated protein 3	<i>MED31</i>	Mediator of RNA polymerase II transcription subunit 31
<i>CHP1</i>	Calcineurin B homologous protein 1	<i>RNF115</i>	E3 ubiquitin-protein ligase RNF115
<i>ZFYVE21</i>	Zinc finger FYVE domain-containing protein 21	<i>LRRC42</i>	Leucine-rich repeat-containing protein 42
<i>SIKE1</i>	Suppressor of IKBKE 1	<i>IER3IP1</i>	Immediate early response 3-interacting protein 1
<i>HAUS8</i>	HAUS augmin-like complex subunit 8	<i>NPTN</i>	Neuroplastin
<i>C11orf84</i>	Uncharacterized protein C11orf84	<i>DUS3L</i>	tRNA-dihydrouridine (47) synthase NADP ⁺ -like
<i>C9orf142</i>	Uncharacterized protein C9orf142	<i>IPO8</i>	Importin-8
<i>SGPP1</i>	Sphingosine-1-phosphate phosphatase 1	<i>ACSL3</i>	Long-chain-fatty-acid-CoA ligase 3
<i>MKRN2</i>	Probable E3 ubiquitin-protein ligase makorin-2	<i>FAM160B1</i>	Protein FAM160B1
<i>KLC2</i>	Kinesin light chain 2	<i>SMC6</i>	Structural maintenance of chromosomes protein 6
<i>NSRP1</i>	Nuclear speckle splicing regulatory protein 1	<i>PIK3R4</i>	Phosphoinositide 3-kinase regulatory subunit 4
<i>CSTF2T</i>	Cleavage stimulation factor subunit 2 tau variant	<i>WAC</i>	WW domain-containing adapter protein with coiled-coil
<i>TMEM222</i>	Transmembrane protein 222	<i>MED23</i>	Mediator of RNA polymerase II transcription subunit 23
<i>PCIF1</i>	Phosphorylated CTD-interacting factor 1	<i>NISCH</i>	Nischarin
<i>DHX33</i>	Putative ATP-dependent RNA helicase DHX33	<i>SNX8</i>	Sorting nexin-8
<i>PHAX</i>	Phosphorylated adapter RNA export protein	<i>UQCRCQ</i>	Cytochrome b-c1 complex subunit 8
<i>CAAP1</i>	Caspase activity and apoptosis inhibitor 1	<i>MED14</i>	Mediator of RNA polymerase II transcription subunit 14
<i>CLSPN</i>	Claspin	<i>CDC45</i>	Cell division control protein 45 homolog
<i>ZFYVE1</i>	Zinc finger FYVE domain-containing protein 1		
<i>COPRS</i>	Coordinator of PRMT5 and differentiation stimulator		
<i>EIF4ENIF1</i>	Eukaryotic translation initiation factor 4E transporter		
<i>UBQLN4</i>	Ubiquilin-4		
<i>KLC4</i>	Kinesin light chain 4		
<i>PRMT7</i>	Protein arginine N-methyltransferase 7		
<i>PIH1D1</i>	PIH1 domain-containing protein 1		

(续表4)

基因名称 Gene names	蛋白名称 Protein names	基因名称 Gene names	蛋白名称 Protein names
<i>PIK3R1</i>	Phosphatidylinositol 3-kinase regulatory subunit alpha	<i>ZNF22</i>	Zinc finger protein 22
<i>NEDD4</i>	E3 ubiquitin-protein ligase NEDD4	<i>PSMB9</i>	Proteasome subunit beta type-9
<i>ARPP19</i>	cAMP-regulated phosphoprotein 19	<i>ACTR1B</i>	Beta-centractin
<i>MRPS6</i>	28S ribosomal protein S6, mitochondrial	<i>CASP3</i>	Caspase-3; Caspase-3 subunit p17; Caspase-3 subunit p12
<i>FPGS</i>	Folypolyglutamate synthase, mitochondrial	<i>KNTC1</i>	Kinetochore-associated protein 1
<i>TWF1</i>	Twinfilin-1	<i>MECP2</i>	Methyl-CpG-binding protein 2
<i>TSTA3</i>	GDP-L-fucose synthase	<i>PMS2</i>	Mismatch repair endonuclease PMS2
<i>PDCD2</i>	Programmed cell death protein 2	<i>NDUFA6</i>	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 6
<i>MAPKAPK3</i>	MAP kinase-activated protein kinase 3	<i>CXCR4</i>	C-X-C chemokine receptor type 4
<i>KDM3B</i>	Lysine-specific demethylase 3B	<i>RRAS2</i>	Ras-related protein R-Ras2
<i>TBC1D1</i>	TBC1 domain family member 1	<i>TIMM10</i>	Mitochondrial import inner membrane translocase subunit Tim10
<i>PPWD1</i>	Peptidylprolyl isomerase domain and WD repeat-containing protein 1	<i>SUPT4H1</i>	Transcription elongation factor SPT4
<i>PIGS</i>	GPI transamidase component PIG-S	<i>LACTB</i>	Serine beta-lactamase-like protein LACTB, mitochondrial
<i>TSG101</i>	Tumor susceptibility gene 101 protein	<i>SRSF4</i>	Serine/arginine-rich splicing factor 4
<i>GINS4</i>	DNA replication complex GINS protein SLD5; DNA replication complex GINS protein SLD5, N-terminally processed	<i>CDC20</i>	Cell division cycle protein 20 homolog
<i>REPIN1</i>	Replication initiator 1	<i>USP4</i>	Ubiquitin carboxyl-terminal hydrolase 4
<i>TRABD</i>	TraB domain-containing protein	<i>PPFIA1</i>	Liprin-alpha-1
<i>KAT8</i>	Histone acetyltransferase KAT8	<i>DNAJC3</i>	DnaJ homolog subfamily C member 3
<i>POLR1B</i>	DNA-directed RNA polymerase I subunit RPA2	<i>PDIA5</i>	Protein disulfide-isomerase A5
<i>EIF2B3</i>	Translation initiation factor eIF-2B subunit gamma	<i>TMED2</i>	Transmembrane emp24 domain-containing protein 2
<i>NCDN</i>	Neurochondrin	<i>TARBP2</i>	RISC-loading complex subunit TARBP2
<i>ZC3H7B</i>	Zinc finger CCCH domain-containing protein 7B	<i>CHI3L2</i>	Chitinase-3-like protein 2
<i>VAV3</i>	Guanine nucleotide exchange factor VAV3	<i>BAK1</i>	Bcl-2 homologous antagonist/killer
<i>GTPBP10</i>	GTP-binding protein 10	<i>GUK1</i>	Guanylate kinase
<i>PDXK</i>	Pyridoxal kinase	<i>QRICH1</i>	Glutamine-rich protein 1
<i>CDK2AP1; CDK2AP2</i>	Cyclin-dependent kinase 2-associated protein 1; cyclin-dependent kinase 2-associated protein 2	<i>PREPL</i>	Prolyl endopeptidase-like
<i>MAP3K7</i>	Mitogen-activated protein kinase kinase kinase 7	<i>EARS2</i>	Probable glutamate-tRNA ligase, mitochondrial
<i>NDUFB5</i>	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 5, mitochondrial	<i>NIPBL</i>	Nipped-B-like protein
<i>STX6</i>	Syntaxin-6	<i>RETSAT</i>	All-trans-retinol 13,14-reductase
<i>MRPS14</i>	28S ribosomal protein S14, mitochondrial	<i>TTI2</i>	TELO2-interacting protein 2
<i>GBAS</i>	Protein NipSnap homolog 2	<i>GALNT7</i>	N-acetylgalactosaminyltransferase 7
<i>DFFB</i>	DNA fragmentation factor subunit beta	<i>METTL16</i>	Methyltransferase-like protein 16
<i>SNUPN</i>	Snurportin-1	<i>ABI1</i>	Abl interactor 1
<i>PTBP3</i>	Polypyrimidine tract-binding protein 3	<i>ATPAF2</i>	ATP synthase mitochondrial F1 complex assembly factor 2
<i>PCCB</i>	Propionyl-CoA carboxylase beta chain, mitochondrial	<i>LEMD2</i>	LEM domain-containing protein 2
<i>MRPL3</i>	39S ribosomal protein L3, mitochondrial	<i>PDPR</i>	Pyruvate dehydrogenase phosphatase regulatory subunit, mitochondrial
<i>PC</i>	Pyruvate carboxylase, mitochondrial	<i>FBXO22</i>	F-box only protein 22
<i>CDK4</i>	Cyclin-dependent kinase 4	<i>TMEM167A</i>	Protein kish-A
<i>GLUL</i>	Glutamine synthetase	<i>C12orf23</i>	UPF0444 transmembrane protein C12orf23
<i>NCK1</i>	Cytoplasmic protein NCK1	<i>IRGQ</i>	Immunity-related GTPase family Q protein
		<i>GCDH</i>	Glutaryl-CoA dehydrogenase, mitochondrial
		<i>WBP2</i>	WW domain-binding protein 2

(续表4)

基因名称 Gene names	蛋白名称 Protein names	基因名称 Gene names	蛋白名称 Protein names
<i>EXOC4</i>	Exocyst complex component 4	<i>HIGD1A</i>	HIG1 domain family member 1A, mitochondrial
<i>FAM105B</i>	Ubiquitin thioesterase otulin	<i>PI4K2A</i>	Phosphatidylinositol 4-kinase type 2- α
<i>C3orf37</i>	UPF0361 protein C3orf37	<i>AMDHD2</i>	Putative N-acetylglucosamine-6-phosphate deacetylase
<i>XPO6</i>	Exportin-6	<i>TMEM256</i>	Transmembrane protein 256
<i>MYADM</i>	Myeloid-associated differentiation marker	<i>IFRD1</i>	Interferon-related developmental regulator 1
<i>MPHOSPH8</i>	M-phase phosphoprotein 8	<i>ZNF593</i>	Zinc finger protein 593
<i>PAAF1</i>	Proteasomal ATPase-associated factor 1	<i>GAPDHS</i>	Glyceraldehyde-3-phosphate dehydrogenase, testis-specific
<i>CNPY3</i>	Protein canopy homolog 3	<i>MGST3</i>	Microsomal glutathione S-transferase 3
<i>CHID1</i>	Chitinase domain-containing protein 1	<i>GEMIN2</i>	Gem-associated protein 2
<i>NUF2</i>	Kinetochore protein Nuf2	<i>AURKA</i>	Aurora kinase A
<i>GHITM</i>	Growth hormone-inducible transmembrane protein	<i>NVL</i>	Nuclear valosin-containing protein-like
<i>PTPN23</i>	Tyrosine-protein phosphatase non-receptor type 23	<i>DDX3Y</i>	ATP-dependent RNA helicase DDX3Y
<i>DNAJC5</i>	DnaJ homolog subfamily C member 5	<i>FIBP</i>	Acidic fibroblast growth factor intracellular-binding protein
<i>PTGES2</i>	Prostaglandin E synthase 2; prostaglandin E synthase 2 truncated form	<i>RAD51C</i>	DNA repair protein RAD51 homolog 3
<i>WDR54</i>	WD repeat-containing protein 54	<i>FADS1</i>	Fatty acid desaturase 1
<i>PLEKHA1</i>	Pleckstrin homology domain-containing family A member 1	<i>KPNA6; KPNA5</i>	Importin subunit α -7; importin subunit α -6
<i>SPC25</i>	Kinetochore protein Spc25	<i>CCDC22</i>	Coiled-coil domain-containing protein 22
<i>APOBEC3C</i>	DNA dC->dU-editing enzyme APOBEC-3C	<i>KIF21B</i>	Kinesin-like protein KIF21B
<i>IMPAD1</i>	Inositol monophosphatase 3	<i>N4BP1</i>	NEDD4-binding protein 1
<i>USE1</i>	Vesicle transport protein USE1	<i>PPP6R2</i>	Serine/threonine-protein phosphatase 6 regulatory subunit 2
<i>MPP6</i>	MAGUK p55 subfamily member 6	<i>NDUFS6</i>	NADH dehydrogenase [ubiquinone] iron-sulfur protein 6, mitochondrial
<i>CWC15</i>	Spliceosome-associated protein CWC15 homolog	<i>MED24</i>	Mediator of RNA polymerase II transcription subunit 24
<i>RNF181</i>	E3 ubiquitin-protein ligase RNF181	<i>CRCP</i>	DNA-directed RNA polymerase III subunit RPC9
<i>CXXC1</i>	CpG-binding protein	<i>ERAL1</i>	GTPase Era, mitochondrial
<i>PHRF1</i>	PHD and RING finger domain-containing protein 1	<i>PIAS2</i>	E3 SUMO-protein ligase PIAS2
<i>METTL1</i>	tRNA (guanine-N(7)-)-methyltransferase	<i>VAMP5</i>	Vesicle-associated membrane protein 5
<i>SEL1L</i>	Protein sel-1 homolog 1	<i>SIX6</i>	Homeobox protein SIX6
<i>DAXX</i>	Death domain-associated protein 6	<i>MT-CO3</i>	Cytochrome c oxidase subunit 3
<i>TUBGCP4</i>	Gamma-tubulin complex component 4	<i>HLA-A</i>	HLA class I histocompatibility antigen, A-68 α chain
<i>LAMTOR3</i>	Ragulator complex protein LAMTOR3	<i>RAF1</i>	RAF proto-oncogene serine/threonine-protein kinase
<i>ANAPC4</i>	Anaphase-promoting complex subunit 4	<i>HMGN1</i>	Non-histone chromosomal protein HMG-14
<i>MINPP1</i>	Multiple inositol polyphosphate phosphatase 1	<i>HMGN2; HMGN3</i>	Non-histone chromosomal protein HMG-17; high mobility group nucleosome-binding domain-containing protein 3
<i>RPL26L1</i>	60S ribosomal protein L26-like 1	<i>SRP19</i>	Signal recognition particle 19 kDa protein
<i>AZII</i>	5-azacytidine-induced protein 1	<i>POLR2M; GCOMI</i>	DNA-directed RNA polymerase II subunit GRINL1A
<i>TRAPPC4</i>	Trafficking protein particle complex subunit 4	<i>HIST1H1E</i>	Histone H1.4
<i>RAP2C</i>	Ras-related protein Rap-2c	<i>BCR</i>	Breakpoint cluster region protein
<i>RNF114</i>	RING finger protein 114		
<i>NDUFB9</i>	NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 9		
<i>GOPC</i>	Golgi-associated PDZ and coiled-coil motif-containing protein		
<i>RGPD3; RGPD4</i>	RanBP2-like and GRIP domain-containing protein 3; RanBP2-like and GRIP domain-containing protein 4		

(续表4)

基因名称 Gene names	蛋白名称 Protein names	基因名称 Gene names	蛋白名称 Protein names
<i>BCKDHA</i>	2-oxoisovalerate dehydrogenase subunit alpha, mitochondrial	<i>PLP2</i>	Proteolipid protein 2
<i>ALAD</i>	Delta-aminolevulinic acid dehydratase	<i>GOLGA2</i>	Golgin subfamily A member 2
<i>PLS3</i>	Plastin-3	<i>TRIM26</i>	Tripartite motif-containing protein 26
<i>ACYP2</i>	Acylphosphatase-2	<i>STK3</i>	Serine/threonine-protein kinase 3; Serine/threonine-protein kinase 3 36 kDa subunit; Serine/threonine-protein kinase 3 20 kDa subunit
<i>ERCC2</i>	TFIIH basal transcription factor complex helicase XPD subunit	<i>NMI</i>	N-myc-interactor
<i>RAB3A</i>	Ras-related protein Rab-3A	<i>XRCC4</i>	DNA repair protein XRCC4
<i>RAB4A</i>	Ras-related protein Rab-4A	<i>TOP3A</i>	DNA topoisomerase 3-alpha
<i>GSTM3</i>	Glutathione S-transferase Mu 3	<i>TUBB3</i>	Tubulin beta-3 chain
<i>RFX1</i>	MHC class II regulatory factor RFX1	<i>DGKZ</i>	Diacylglycerol kinase zeta
<i>ATP6V0C</i>	V-type proton ATPase 16 kDa proteolipid subunit	<i>MTMR1</i>	Myotubularin-related protein 1
<i>GSTM2</i>	Glutathione S-transferase Mu 2	<i>DYRK1A; DYRK1B</i>	Dual specificity tyrosine-phosphorylation-regulated kinase 1A; dual specificity tyrosine-phosphorylation-regulated kinase 1B
<i>PRKAR1B</i>	cAMP-dependent protein kinase type I-beta regulatory subunit	<i>C1D</i>	Nuclear nucleic acid-binding protein C1D
<i>IL2RG</i>	Cytokine receptor common subunit gamma	<i>COX17</i>	Cytochrome c oxidase copper chaperone
<i>KDEL2; KDEL1</i>	ER lumen protein retaining receptor 2; ER lumen protein retaining receptor 1	<i>MORC3</i>	MORC family CW-type zinc finger protein 3
<i>CA8</i>	Carbonic anhydrase-related protein	<i>ARHGEF7</i>	Rho guanine nucleotide exchange factor 7
<i>MPV17</i>	Protein Mpv17	<i>TFDP1</i>	Transcription factor Dp-1
<i>TSPAN7</i>	Tetraspanin-7	<i>GSE1</i>	Genetic suppressor element 1
<i>SLC1A4</i>	Neutral amino acid transporter A	<i>PPP2R5A</i>	Serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit alpha isoform
<i>SYK</i>	Tyrosine-protein kinase SYK	<i>RBBP5</i>	Retinoblastoma-binding protein 5
<i>ACADSB</i>	Short/branched chain specific acyl-CoA dehydrogenase, mitochondrial	<i>TOMM20</i>	Mitochondrial import receptor subunit TOM20 homolog
<i>IREB2</i>	Iron-responsive element-binding protein 2	<i>SS18</i>	Protein SSXT
<i>HSPA13</i>	Heat shock 70 kDa protein 13	<i>DECRI</i>	2,4-dienoyl-CoA reductase, mitochondrial
<i>CD97</i>	CD97 antigen; CD97 antigen subunit alpha; CD97 antigen subunit beta	<i>DGUOK</i>	Deoxyguanosine kinase, mitochondrial
<i>CENPF</i>	Centromere protein F	<i>PFKFB3</i>	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3; 6-phosphofructo-2-kinase; Fructose-2,6-bisphosphatase
<i>HARS2</i>	Probable histidine-tRNA ligase, mitochondrial	<i>HNRNPA1L2</i>	Heterogeneous nuclear ribonucleoprotein A1-like 2
<i>CSNK1E</i>	Casein kinase I isoform epsilon	<i>TWISTNB</i>	DNA-directed RNA polymerase I subunit RPA43
<i>GSK3A</i>	Glycogen synthase kinase-3 alpha	<i>SNAI3</i>	Zinc finger protein SNAI3
<i>GALK1</i>	Galactokinase	<i>TUBB8</i>	Tubulin beta-8 chain
<i>VAMP7</i>	Vesicle-associated membrane protein 7	<i>TMEM41B</i>	Transmembrane protein 41B
<i>RABGGTB</i>	Geranylgeranyl transferase type-2 subunit beta	<i>OXLD1</i>	Oxidoreductase-like domain-containing protein 1
<i>HIRA</i>	Protein HIRA	<i>DNAJC21</i>	DnaJ homolog subfamily C member 21
<i>TMBIM6</i>	Bax inhibitor 1	<i>TMEM201</i>	Transmembrane protein 201
<i>HSD17B7</i>	3-keto-steroid reductase	<i>C9orf114</i>	Uncharacterized protein C9orf114
<i>TMEM258</i>	Transmembrane protein 258	<i>EMC10</i>	ER membrane protein complex subunit 10
<i>RPS29</i>	40S ribosomal protein S29	<i>SYDE2</i>	Rho GTPase-activating protein SYDE2
<i>RAB1A</i>	Ras-related protein Rab-1A	<i>MBLAC2</i>	Metallo-beta-lactamase domain-containing protein 2
<i>PRAME</i>	Melanoma antigen preferentially expressed in tumors	<i>C12orf73</i>	Uncharacterized protein C12orf73
<i>CCZ1; CCZ1B</i>	Vacuolar fusion protein CCZ1 homolog; vacuolar fusion protein CCZ1 homolog B	<i>GRAMD4</i>	GRAM domain-containing protein 4
<i>REEP5</i>	Receptor expression-enhancing protein 5		
<i>DHODH</i>	Dihydroorotate dehydrogenase (quinone), mitochondrial		

(续表4)

基因名称 Gene names	蛋白名称 Protein names	基因名称 Gene names	蛋白名称 Protein names
<i>LAIR1</i>	Leukocyte-associated immunoglobulin-like receptor 1	<i>PIK3C3</i>	Phosphatidylinositol 3-kinase catalytic subunit type 3
<i>RPUSD3</i>	RNA pseudouridylate synthase domain-containing protein 3	<i>TOR1AIP2</i>	Torsin-1A-interacting protein 2
<i>DUS1L</i>	tRNA-dihydrouridine (16/17) synthase NADP ⁺ -like	<i>ABHD11</i>	Alpha/beta hydrolase domain-containing protein 11
<i>FAHD2B</i>	Fumarylacetoacetate hydrolase domain-containing protein 2B	<i>MRPL30</i>	39S ribosomal protein L30, mitochondrial
<i>MTFR2</i>	Mitochondrial fission regulator 2	<i>SCFD2</i>	Sec1 family domain-containing protein 2
<i>WDR73</i>	WD repeat-containing protein 73	<i>DCAKD</i>	Dephospho-CoA kinase domain-containing protein
<i>FAHD1</i>	Acylpyruvase FAHD1, mitochondrial	<i>SNRNP27</i>	U4/U6.U5 small nuclear ribonucleoprotein 27 kDa protein
<i>RPL22L1</i>	60S ribosomal protein L22-like 1	<i>SLC39A7</i>	Zinc transporter SLC39A7
<i>C6orf58</i>	UPF0762 protein C6orf58	<i>NDRG1</i>	Protein NDRG1
<i>CYP20A1</i>	Cytochrome P450 20A1	<i>ALG3</i>	Dol-P-Man:Man(5)GlcNAc(2)-PP-Dol alpha-1,3-mannosyltransferase
<i>DNMBP</i>	Dynamin-binding protein	<i>RREB1</i>	Ras-responsive element-binding protein 1
<i>QSOX2</i>	Sulphydryl oxidase 2	<i>ERCC4</i>	DNA repair endonuclease XPF
<i>FRA10AC1</i>	Protein FRA10AC1	<i>KIAA1143</i>	Uncharacterized protein KIAA1143
<i>COX15</i>	Cytochrome c oxidase assembly protein COX15 homolog	<i>FLYWCH2</i>	FLYWCH family member 2
<i>SV2A</i>	Synaptic vesicle glycoprotein 2A	<i>FOXRED1</i>	FAD-dependent oxidoreductase domain-containing protein 1
<i>PKD1L3</i>	Polycystic kidney disease protein 1-like 3	<i>C10orf35</i>	Uncharacterized protein C10orf35
<i>COMMD6</i>	COMM domain-containing protein 6	<i>HVCN1</i>	Voltage-gated hydrogen channel 1
<i>CMC1</i>	COX assembly mitochondrial protein homolog	<i>FAM120B</i>	Constitutive activator of peroxisome proliferator-activated receptor gamma
<i>ISCA2</i>	Iron-sulfur cluster assembly 2 homolog, mitochondrial	<i>RRP36</i>	Ribosomal RNA processing protein 36 homolog
<i>LDB1</i>	LIM domain-binding protein 1	<i>MMAB</i>	Cob(II)yrinic acid a,c-diamide adenosyltransferase, mitochondrial
<i>PARG</i>	Poly(ADP-ribose) glycohydrolase	<i>DYNLL2</i>	Dynein light chain 2, cytoplasmic
<i>NDUFA11</i>	NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 11	<i>CHMP6</i>	Charged multivesicular body protein 6
<i>MGA</i>	MAX gene-associated protein	<i>RNF185</i>	E3 ubiquitin-protein ligase RNF185
<i>ZC3H7A</i>	Zinc finger CCCH domain-containing protein 7A	<i>TOE1</i>	Target of EGR1 protein 1
<i>TPH2</i>	Tryptophan 5-hydroxylase 2	<i>POM121</i>	Nuclear envelope pore membrane protein POM 121
<i>CDAN1</i>	Codanin-1	<i>INTS4</i>	Integrator complex subunit 4
<i>DNAJC10</i>	DnaJ homolog subfamily C member 10	<i>SMIM8</i>	Small integral membrane protein 8
<i>SIRT2</i>	NAD-dependent protein deacetylase sirtuin-2	<i>ARAP1</i>	Arf-GAP with Rho-GAP domain, ANK repeat and PH domain-containing protein 1
<i>EXOC8</i>	Exocyst complex component 8	<i>OSBPL9</i>	Oxysterol-binding protein-related protein 9
<i>C1orf174</i>	UPF0688 protein C1orf174	<i>RNF2</i>	E3 ubiquitin-protein ligase RING2
<i>THNSL1</i>	Threonine synthase-like 1	<i>MYD88</i>	Myeloid differentiation primary response protein MyD88
<i>KIAA2013</i>	Uncharacterized protein KIAA2013	<i>ARPC5L</i>	Actin-related protein 2/3 complex subunit 5-like protein
<i>C8orf59</i>	Uncharacterized protein C8orf59	<i>MICU1</i>	Calcium uptake protein 1, mitochondrial
<i>ORMDL3; ORMDL2</i>	ORM1-like protein 3; ORM1-like protein 2	<i>FAM118B</i>	Protein FAM118B
<i>ANKRD35</i>	Ankyrin repeat domain-containing protein 35	<i>C17orf62</i>	Uncharacterized protein C17orf62
<i>ARMC10</i>	Armadillo repeat-containing protein 10	<i>TSEN34</i>	tRNA-splicing endonuclease subunit Sen34
<i>SMIM20</i>	Small integral membrane protein 20	<i>DCTN5</i>	Dynactin subunit 5
<i>JAGN1</i>	Protein jagunal homolog 1		
<i>PARP16</i>	Mono [ADP-ribose] polymerase PARP16		
<i>C19orf47</i>	Uncharacterized protein C19orf47		
<i>NHLRC2</i>	NHL repeat-containing protein 2		
<i>DAGLB</i>	Sn1-specific diacylglycerol lipase beta		

(续表4)

基因名称 Gene names	蛋白名称 Protein names	基因名称 Gene names	蛋白名称 Protein names
<i>LRRC1</i>	Leucine-rich repeat-containing protein 1	<i>DECR2</i>	Peroxisomal 2,4-dienoyl-CoA reductase
<i>FSD1</i>	Fibronectin type III and SPRY domain-containing protein 1	<i>C19orf66</i>	UPF0515 protein C19orf66
<i>WRAP53</i>	Telomerase Cajal body protein 1	<i>UFSP2</i>	Ufm1-specific protease 2
<i>TARS2</i>	Threonine--tRNA ligase, mitochondrial	<i>MED17</i>	Mediator of RNA polymerase II transcription subunit 17
<i>CDC47</i>	Cell division cycle-associated protein 7	<i>TMEM161A</i>	Transmembrane protein 161A
<i>CEP41</i>	Centrosomal protein of 41 kDa	<i>BET1L</i>	BET1-like protein
<i>EGLN1</i>	Egl nine homolog 1	<i>CCDC88C</i>	Protein Daple
<i>TMEM126A</i>	Transmembrane protein 126A	<i>STIM2</i>	Stromal interaction molecule 2
<i>SAP130</i>	Histone deacetylase complex subunit SAP130	<i>SLC25A10</i>	Mitochondrial dicarboxylate carrier
<i>QRSL1</i>	Glutamyl-tRNA (Gln) amidotransferase subunit A, mitochondrial	<i>TJP2</i>	Tight junction protein ZO-2
<i>VPS33B</i>	Vacuolar protein sorting-associated protein 33B	<i>ADD3</i>	Gamma-adducin
<i>SLC25A32</i>	Mitochondrial folate transporter/carrier	<i>LIMD1</i>	LIM domain-containing protein 1
<i>TRIT1</i>	tRNA dimethylallyltransferase, mitochondrial	<i>RSL24D1</i>	Probable ribosome biogenesis protein RLP24
<i>SRMS</i>	Tyrosine-protein kinase Srms	<i>SHPK</i>	Sedoheptulokinase
<i>CUEDC2</i>	CUE domain-containing protein 2	<i>BAZ2A</i>	Bromodomain adjacent to zinc finger domain protein 2A
<i>FBXO44</i>	F-box only protein 44	<i>TSR3</i>	Ribosome biogenesis protein TSR3 homolog
<i>TUT1</i>	Speckle targeted PIP5K1A-regulated poly(A) polymerase	<i>TUBE1</i>	Tubulin epsilon chain
<i>REEP4</i>	Receptor expression-enhancing protein 4	<i>ANAPC5</i>	Anaphase-promoting complex subunit 5
<i>METTL17</i>	Methyltransferase-like protein 17, mitochondrial	<i>FAM184B</i>	Protein FAM184B
<i>NRDE2</i>	Protein NRDE2 homolog	<i>PALD1</i>	Paladin
<i>NT5DC2</i>	5-nucleotidase domain-containing protein 2	<i>COG5</i>	Conserved oligomeric Golgi complex subunit 5
<i>RMND5A</i>	Protein RMD5 homolog A	<i>FOXJ3</i>	Forkhead box protein J3
<i>ZWILCH</i>	Protein zwilch homolog	<i>SHOC2</i>	Leucine-rich repeat protein SHOC-2
<i>MED20</i>	Mediator of RNA polymerase II transcription subunit 20	<i>DNM3</i>	Dynamin-3
<i>WDR76</i>	WD repeat-containing protein 76	<i>TMA7</i>	Translation machinery-associated protein 7
<i>RPF1</i>	Ribosome production factor 1	<i>KLF13</i>	Krueppel-like factor 13
<i>UCK1</i>	Uridine-cytidine kinase 1	<i>POLR3H</i>	DNA-directed RNA polymerase III subunit RPC8
<i>SP110</i>	Sp110 nuclear body protein	<i>NUB1</i>	NEDD8 ultimate buster 1
<i>TMEM165</i>	Transmembrane protein 165	<i>NME7</i>	Nucleoside diphosphate kinase 7
<i>NOP10</i>	H/ACA ribonucleoprotein complex subunit 3	<i>MPC1</i>	Mitochondrial pyruvate carrier 1
<i>SDR39U1</i>	Epimerase family protein SDR39U1	<i>AUP1</i>	Ancient ubiquitous protein 1
<i>ZCCHC3</i>	Zinc finger CCHC domain-containing protein 3	<i>CEPT1</i>	Choline/ethanolaminephosphotransferase 1

3 讨论

作为研究工具, 高浓度甲醛被广泛应用于交联蛋白质-蛋白质、蛋白质-DNA分子, 研究这些分子的相互作用, 所涉及到的研究领域或研究工具包括蛋白质组学、组蛋白翻译后修饰、染色体免疫沉淀、染色质构象捕捉等^[1]。之前的观点认为, 甲醛可以快速通过细胞膜高效地瞬间将蛋白质等分子交联, 将它们“冻结”在自然状态中^[2]。但是, 高浓度甲醛是否在其瞬间“冻结”蛋白质等分子之前就诱导细胞反应, 导致

最终研究结果的失真, 目前尚未见具体报道。

我们通过非同位素标记的蛋白质组学定量手段, 研究对比了经过高浓度甲醛处理和未经甲醛处理的对照Jurkat细胞的蛋白质谱系的变化情况, 甲醛交联条件采用了文献中常用的1%浓度交联10 min。实验结果表明, 高浓度甲醛处理细胞的蛋白质水平发生变化。这些差异蛋白质可分为两类。第一类的差异蛋白质(136个)在两次甲醛实验中重复出现, 约占全部鉴定到蛋白质的3%~4%。这些差异蛋白质

存在于细胞膜、细胞质和细胞核(图3), 并与22条细胞通路相关(图4)。尽管无法完全排除部分被鉴定出的蛋白质, 受到了实验条件、细胞状态、细胞培养环境等随机因素的影响, 但这类蛋白质水平的差异, 与甲醛关联的可能性较大, 甚至由甲醛诱导发生。如果一个课题的研究对象不涉及这些与甲醛特异相关的蛋白质或与其相联的细胞通路, 那么, 将甲醛作为交联剂使用的实验(比如组蛋白翻译后修饰或染色体免疫沉淀), 鉴定到的蛋白质呈现假阳性或假阴性的可能性相对较小。

第二类的差异蛋白质(614个), 在两次甲醛实验中仅鉴定到一次, 重复性差。这类差异蛋白质比第一类蛋白质多, 约占全部鉴定到蛋白质的六分之一, 涉及到多个细胞反应, 比如基因转录、DNA修复蛋白、DNA结合蛋白、组蛋白翻译后修饰蛋白等。这类蛋白质可能与甲醛的关联较小, 更可能源自于随机因素, 比如实验条件(包括超净工作台的温度)、细胞状态(包括G₁、S、G₂和M期细胞或凋亡细胞比例、悬浮细胞的细胞密度或单层细胞的细胞融合程度)、细胞培养环境(包括不同批次的小牛血清)等。根据文献报道, 在蛋白组学研究中, 这种重复性差的现象比较常见。此外, 甲醛也可交联一些非特异的、随机出现的所谓“路过”蛋白质。更为复杂的是, 在一个试验中, 这些非特异性的随机因素常常以不同组合、不同程度地叠加在一起出现, 极难控制, 因此, 阐明这些因素所引起的特异蛋白质水平变化极具挑战性, 文献中也极少报道。但是, 对于将甲醛作为交联试剂使用的实验(比如蛋白质组学或染色体免疫沉淀)而言, 由于这类蛋白质被交联的随机性和不可预见性, 如果这些蛋白质恰与课题所研究的对象蛋白质重合, 则可导致假阳性或假阴性的结果。

综合上述结果, 在甲醛作为交联试剂使用时, 实验(比如组蛋白翻译后修饰或染色质免疫沉淀)中必须包括一个仅经过甲醛处理的平行样本作为其他样本的对照, 或者增加实验重复次数, 以避免甲醛本身引起的假阳性或假阴性结果。如果可以用其他研究方法或技术进一步验证甲醛交联的实验结果, 则更为稳妥, 比如免疫组化、酵母双杂交系统等可用于证明甲醛交联到的蛋白质-蛋白质间相互作用的特异性。

参考文献 (References)

- 1 Sutherland BW, Toews J, Kast J. Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. *J Mass Spectrom* 2008; 43(6): 699-715.
- 2 Swenberg JA, Moeller BC, Lu K, Rager JE, Fry RC, Starr TB. Formaldehyde carcinogenicity research: 30 years and counting for mode of action, epidemiology, and cancer risk assessment. *Toxicol Pathol* 2013; 41(2): 181-9.
- 3 Szende B, Tyihák E. Effect of formaldehyde on cell proliferation and death. *Cell Biol Int* 2010; 34(12): 1273-82.
- 4 Yin X, Zhang Y, Liu X, Chen C, Lu H, Shen H, *et al.* Systematic comparison between SDS-PAGE/RPLC and high-/low-pH RPLC coupled tandem mass spectrometry strategies in a whole proteome analysis. *Analyst* 2015; 140(4): 1314-22.
- 5 Cao J, Shen C, Wang H, Shen H, Chen Y, Nie A, *et al.* Identification of N-Glycosylation sites on secreted proteins of human hepatocellular carcinoma cells with a complementary proteomics approach. *J Proteome Res* 2009; 8: 662-72.
- 6 Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008; 26: 1367-72.
- 7 Cox J, Neuhauser N, Michalski A, Scheltema RA, Olsen JV, Mann M. Andromeda: A peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 2011; 10: 1794-805.
- 8 Cox J, Hein MY, Lubner CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* 2014; 13(9): 2513-26.