基于全基因转录组的结核诊断标识研究

陈献雄^{1†}, 杨倩婷^{2†}, 邱振纲³, 朱家楼¹, 张明霞², 邓国防², 邓群益², 陈心春^{1*}

1. 深圳大学医学部基础医学院病原生物学系, 深圳 518060;

2. 深圳市第三人民医院深圳市感染免疫重点实验室, 深圳 518112;

3. 赣南医学院第一附属医院, 赣州 341400

† 同等贡献

* 联系人, E-mail: chenxinchun@szu.edu.cn

收稿日期: 2017-10-15; 接受日期: 2017-11-01; 网络版发表日期: 2018-01-17

国家杰出青年科学基金(批准号: 81525016)、国家自然科学基金(批准号: 81671984)、深圳市科技计划项目基础研究(学科布局)项目(批准号: JCYJ20160427184123851)、深圳市科技计划项目基础研究项目(批准号: JCYJ20150402111430656, JCYJ20150402145015986, JCYJ20160427151540695)和深圳市科技计划项目基础研究(技术攻关)项目(批准号: JSGG20160427104724699)资助

摘要 早期诊断是结核患者得到及时治疗的关键, 目前结核病诊断手段不足, 迫切需要更加快速、简便的诊断方法. 本研究分析9例治疗前结核患者、6例结核潜伏感染者、6例健康对照的全基因转录组表达谱, 并对结核患者治疗前后表达谱进行随访. 结果发现, 与健康对照和结核潜伏感染者比较, 结核患者有77个高表达、50个低表达的差异表达基因, 其中有12个差异表达基因表达量经抗结核治疗后下调; 差异表达基因GO分析筛选出9个与炎症反应、免疫应答、趋药性相关的生物学功能通路, KEGG pathway分析发现, 1个细胞因子和细胞因子受体相关的信号转导通路, 各通路均有趋化因子或者细胞因子参与. 应用RT-qPCR方法进行临床扩大样本验证, 发现大部分基因RT-qPCR验证结果与基因芯片结果趋势一致, 其中*HBB*等8个基因一致性最好; 绘制ROC曲线评价*HBB*和*TNFRSF10C*两个基因诊断结核病的效能, AUC均在0.9以上, 敏感性和特异性均在95%以上; 本研究发现的差异表达基因为结核病研究提供了新的生物学标识, 为结核病诊断方法的建立、疗效的监测、发病机制的揭示提供了新依据.

关键词 结核, 全基因转录组, 诊断, 生物标识

结核是一种古老的由结核分枝杆菌感染引起的慢性呼吸道传染性疾病. 由于缺乏有效的保护性疫苗和理想的诊断技术, 耐多药结核高发, 结核病仍然是当前严重危害人类健康的重大公共卫生问题. 早期发现和早期治疗是当前结核病防控的主要策略, 但现有的结核病诊断主要依据病史、临床症状、体征、影像学检查、痰结核分支杆菌检查、结核菌素试验(purified protein derivative, PPD)等, 存在特异性或敏感性不足

的问题, 难以满足临床需求. 采用显微镜检查、结核菌培养、结核菌核酸PCR等技术检测痰标本中的结核菌是诊断活动性结核的金标准, 但是即便综合不同的最先进的检测技术, 能获得病原学诊断证据的结核患者也不足50%. 换句话说, 超过50%的患者仍然无法通过实验室技术获得病原学确诊, 这部分患者也被称为菌阴肺结核, 是当前结核病临床诊断的最大挑战.

由于结核病是一个免疫介导的慢性炎症疾病, 因

引用格式: 陈献雄, 杨倩婷, 邱振纲, 等. 基于全基因转录组的结核诊断标识研究. 中国科学: 生命科学, 2018, 48: 56-64

Chen X X, Yang Q T, Qiu Z G, et al. Tuberculosis diagnostic markers based on the whole gene transcriptome (in Chinese). Sci Sin Vitae, 2018, 48: 56-64, doi: 10.1360/N052017-00225

此,理论上通过检测宿主对结核菌的免疫应答反应是结核病诊断的重要依据.基于这一理论,早期临床上采用PPD皮试和结核菌抗体的检测来辅助结核诊断,但是大量的实践表明,由于存在交叉反应和抗原特异性等问题,两者均难以满足临床要求.通过检测外周血结核菌抗原特异性T细胞免疫反应的 γ -干扰素释放试验(interferon- γ release assays, IGRAs)^[1]是近二十多年来结核病诊断领域最重大的进展之一,是目前诊断结核菌感染的金标准^[2,3].然而,结核病是一个复杂疾病^[4],虽然感染结核菌是发生结核病的必要条件,但是发生结核菌感染的人群是否发病,更多取决于宿主的免疫反应.最新的流行病学资料表明,我国人群中结核菌感染率超过1/5,但感染人群仅有不到10%发展为活动性结核病.IGRAs因无法有效区分无症状的结核菌潜伏感染者(latent tuberculosis infection, LTBI)和活动性结核患者,在临床活动性结核诊断价值有限.因此,筛选新型IGRAs的结核菌抗原,或者新型宿主应答标识,一直是近年来结核病免疫诊断研究的热点^[5].

本研究对健康对照、潜伏感染者和活动性结核患者外周血单个核细胞中全基因转录组进行系统分析,结合对结核患者抗结核治疗前后的前瞻性随访,筛选出与结核病发生发展相关宿主应答基因,一方面为结核免疫诊断标识奠定基础,另一方面为结核病免疫发病机制提供新的思路.

1 材料与方法

1.1 研究对象和样本

全基因转录组分析样本.研究对象均来自广东省深圳市第三人民医院. HD: 6例,常规体检无异常,既往无结核病史、无任何临床症状,经胸部放射学确定无异常,IGRAs检测结果为阴性. LTBI: 6例,常规体检无异常,既往无结核病史、无任何临床症状,经胸部放射学确定无异常,IGRAs检测结果为阳性. PTB: 9例,根据中华医学会结核分会制定的《肺结核诊断和治疗指南》标准纳入标本及接受正规的抗结核治疗.结核患者诊断依据为痰培养阳性,同时结合其临床表现、放射学检查结果.各组人群的年龄、性别分布: $\bar{x} \pm SD$, (男/女); HC: 28.5 ± 4.72 , (3/3); LTBI: 30.3 ± 4.27 , (3/3); TB: 30.0 ± 11.4 , (6/3).

验证人群.根据上述标准选取20个HC, 20个LTBI, 20个TB进行验证,年龄($\bar{x} \pm SD$, 岁)、性别(男/女)分布; HC: 31.2 ± 7.04 , (14/6); LTBI: 32.5 ± 11.0 , (13/7); TB: 32.6 ± 11.3 , (11/9).

以上所有研究对象均无妊娠、HIV感染、使用免疫抑制剂或增强剂、严重肝、肾功能衰竭、无陈旧性肺结核病灶.临床结核患者标本的收集是经过医院机构审查委员会批准的协议启动的,所有的研究对象都签署了参与知情同意书.

1.2 研究方法

(1) 分离外周血单个核细胞(PBMCs). 抗凝管采静脉血2 mL, 5 mL分离液(LymphoprepTM 1.077, Axis-Shield, 挪威)加入15 mL试管中.充分混匀静脉血与2 mL PBS,缓慢用滴管沿管壁叠加于分离液之液面上,以3774 $\times g$ 离心20 min.到白色云雾层吸取PBMC放入另一试管中,以2830 $\times g$ 离心5 min,再洗涤细胞两次.离心,弃上清,留取PBMCs.

(2) 人全基因组mRNA表达谱芯片分析.使用RNasey mini kit试剂盒(Qiagen, 美国)提纯定量总RNA,按照Affymetrix GeneChip[®] Human Genome U133 Plus 2.0 Array (Affymetrix, 美国)说明书进行操作,由上海博豪生物技术有限公司完成.

(3) 荧光定量RT-qPCR反应.使用Trizol[®]试剂盒(Invitrogen, 美国)抽提细胞总RNA.测定RNA纯度及浓度,反转成为cDNA,进行RT-qPCR实验,使用Primer 3软件自行设计引物,由Qiagen公司合成.实验步骤按照说明书进行.

(4) GO分析和KEGG pathway分析.使用Affymetrix公司的DMT (Gene Chip data mini tool)软件对表达谱芯片数据进行配对统计分析、聚类、功能注释等.

(5) 统计分析.应用GraphPad Prism 5软件进行分析,计量资料以 $\bar{x} \pm SD$ 表示,两组间比较采用独立样本 t 检验,3组间差异用单因素方差分析one-way ANOVA多重实验进行比较, $P < 0.05$ 被认为差异有统计学意义.

2 结果

2.1 利用基因芯片检测技术筛选结核差异表达的基因

利用基因芯片检测技术分别检测6个HC, 6个LTBI

和9个PTB标本6882个基因的表达. 比较3组人群基因的表达差异, 发现PTB有173个探针(127个基因)显著差异表达.

其中108个探针(77个基因)表达量显著高于潜伏感染者和健康对照, 差异高表达最显著的基因依次为: *SLC24A4*, *HBA1*, *CLEC12B*等. 65个探针(50个基因)表达量显著低于潜伏感染者和健康对照差异, 低表达最显著的基因依次为: *GPRC5A*, *LOC440934*, *DUSP4*等, 将所有研究对象的127个差异表达基因芯片结果制成热图(图1), 差异高表达基因芯片结果详见网络版附表1, 差异低表达基因芯片结果详见网络版附表2.

2.2 利用基因芯片检测技术筛选抗结核治疗前后差异表达的基因

从上述差异表达173个探针(127基因), 筛选出抗结核治疗前后表达量有显著差异的探针/基因, 发现20个探针(12个基因)在抗结核治疗前后表达量下调($P<0.05$), 表达量下降最显著基因依次为: *HBA1*///*HBA2*, *HBB*, *MOSCI*等, 无表达量显著上调的探针(基因). 将所有结核患者治疗前后20个探针(12个基因)差异表达基因制成热图(图2A), 随机选择4个基因(探针): *HBB* (211696_x_at), *RNASE2* (206111_at), *TNFRSF10C* (211163_s_at), *SIGLEC* (161561880_a_at), 将结核患者经抗结核治疗前后(治疗0/3/6月)差异表达基因表达量用折线图展示(图2B), 具体基因芯片结果见表1.

2.3 差异表达基因的功能分析

GO分析发现差异高表达基因富集在9个和炎症反应、免疫应答、趋药性功能相关的信号通路, 显著性最高依次为: inflammatory response ($P=6.36\times10^{-4}$), positive regulation to interleukin6 production ($P=4.88\times10^{-3}$), cellular response to interferon-gamma ($P=6.43\times10^{-3}$), 显著性最低的是 cellular response to lipopolysaccharide ($P=2.15\times10^{-2}$); 无差异低表达基因参与的功能通路. KEGG pathway分析发现差异高表达基因富集的1个信号转导通路: Cytokine-cytokine receptor interaction ($P=1.12\times10^{-2}$); 无差异低表达基因参与的信号转导途径(表2) ($P<0.05$ 为有统计学差异, 按 P 值升高为顺序).

2.4 扩大临床样本量对差异表达基因进行验证

使用RT-qPCR方法从基因芯片结果中的173个探针(127基因)进行扩大临床样本量验证, 发现部分基因RT-qPCR验证结果和基因芯片结果并不一致; 大部分基因验证结果与基因芯片结果趋势一致, 其中以4个显著高表达基因(*LRRK2*, *HBA1*, *PRAMI*, *TNFRSF10C*)和4个显著低表达基因(*MMP19*, *MET*, *SLC7A11*, *OLRI*)结果一致性最好(图3, 同一基因左侧为基因芯片结果, 右侧为对应的验证结果).

2.5 *HBB*和*TNFRSF10C*基因在结核病诊断中的效能评价

上述8个基因中, *HBB*和*TNFRSF10C*两个基因在结核患者比对照组高表达, 而且经抗结核治疗后表达

图1 差异表达基因

红色到绿色代表表达量由高到低

图 2 探针基因的表达

A: 治疗前后差异表达的探针/基因; B: 治疗前后同一探针/基因表达水平差异

量下降, 进一步评价两个基因在诊断结核的效能, 绘制ROC曲线(图4)。HBB基因在诊断健康对照和结核患者, AUC为0.98, 敏感性为100%, 特异性为95%; 在诊断潜伏感染和结核患者, AUC为0.985, 敏感性为100%, 特异性为95%。TNFRSF10C基因在诊断健康对照和结核患者, AUC为1.00, 敏感性为100%, 特异性为100%; 在诊断潜伏感染和结核患者, AUC为0.9775, 敏感性为95%, 特异性为100%。

3 结论

从分子水平诊断结核病成为结核研究的一个重要方向, 高通量的基因组学技术极大地丰富了人们对结核分枝杆菌感染免疫的认识^[6,7]。近期在结核病的分子诊断标识研究也有一些的进展。例如, 最近发表在Lancet

的两篇文章, 一篇文章是利用meta-分析研究了3个与结核相关的基因数据库, 发现3个基因(*GBP5*, *DUSP3*和*KLF2*)组合用于活动性结核的诊断, 然后在11个数据库验证该组合在诊断结核病的效能, 结果发现该基因组在鉴别诊断结核患者和健康对照或潜伏感染的效果较好, AUC分别为0.9 (敏感性85%, 特异性93%)和0.88 (敏感性80%, 特异性86%), 但在区分其他疾病时效果稍差, AUC只为0.84 (敏感性81%, 特异性74%)^[8]。另一篇文章则利用全基因组学(whole-genome-sequencing, WGS)检测结核菌标本序列, 应用于耐药结核的诊断。与常规的检测结果相比, WGS的结核菌预测准确率达93%, 药物敏感性为93%, 但是该研究标本来自于痰, 标本质量易受影响, 且平均需要长达9天的时间^[9], 由此可见, 目前有关结核基因组学的研究, 无论在机制研究还是临床应用, 仍存在很多不足。

表 1 治疗前后差异表达的探针/基因

探针编号	基因名称	代表性的公共名称	UniGene ID	P 值	LOG (TB1/TB2,2)	LOG (TB1/TB2,2)
229373_at	—	AW139719	Hs.457403	5.05E-05	0.489488453	0.921019661
204018_x_at	<i>HBA1///HBA2</i>	NM_000558	Hs.449630	5.47E-05	1.659830481	1.726348157
209458_x_at	<i>HBA1///HBA2</i>	AF105974	Hs.449630	5.89E-05	1.614870817	1.813869697
217414_x_at	<i>HBA1///HBA2</i>	V00489	Hs.449630	9.11E-05	1.60824889	1.929411001
211699_x_at	<i>HBA1///HBA2</i>	AF349571	Hs.449630	9.70E-05	1.653670109	1.859975855
211745_x_at	<i>HBA1///HBA2</i>	BC005931	Hs.449630	0.000154	1.567474905	1.85233257
214414_x_at	<i>HBA1///HBA2</i>	T50399	Hs.449630	0.000451	1.266112622	1.481772628
209116_x_at	<i>HBB</i>	M25079	Hs.523443	0.000562	1.453075665	1.858043327
217232_x_at	<i>HBB</i>	AF059180	Hs.523443	0.000876	1.432654008	1.727574486
211696_x_at	<i>HBB</i>	AF349114	Hs.523443	0.0035	1.327617798	1.641866428
218865_at	<i>MOSC1</i>	NM_022746	Hs.497816	0.00357	0.757173297	0.986927982
211163_s_at	<i>TNFRSF10C</i>	AF012536	Hs.655801	0.00434	0.944062591	1.969903899
1561880_a_at	<i>SIGLEC16</i>	BC039008	Hs.568076	0.00461	0.780114952	2.434484199
206111_at	<i>RNASE2</i>	NM_002934	Hs.728	0.00581	0.717100209	1.573258024
209930_s_at	<i>NFE2</i>	L13974	Hs.75643	0.00767	0.931998713	2.056230841
218454_at	<i>PLBD1</i>	NM_024829	Hs.131933	0.012	0.435420085	0.638550074
221666_s_at	<i>PYCARD</i>	BC004470	Hs.499094	0.0155	0.379676715	1.24380321
218773_s_at	<i>MSRB2</i>	NM_012228	Hs.461420	0.0158	0.604751807	1.513240207
228253_at	<i>LOXL3</i>	AI917716	Hs.469045	0.0206	0.393595071	1.294531529
225059_at	<i>AGTRAP</i>	BE875567	Hs.464438	0.0261	0.313215296	1.032887469

表 2 差异表达基因的GO分析和KEGG Pathway分析

ID	描述	P 值	基因
GO biological process			
GO:0006954	Inflammatory response	6.36×10^{-4}	<i>CCL2, OLR1, AIF1, CXCL3, CCL7, CCRL2, TNFAIP6, CCL22, TNFRSF10C, CCL20, PYCARD, IL1B, THBS1, IL1A</i>
GO:0032755	Positive regulation of interleukin-6 production	4.88×10^{-3}	<i>WNT5A, RAB7B, CD36, PYCARD, IL1B, IL1A</i>
GO:0071346	Cellular response to interferon-gamma	6.43×10^{-3}	<i>WNT5A, CCL22, CCL2, CCL20, AIF1, CCL7</i>
GO:0030593	Neutrophil chemotaxis	1.22×10^{-2}	<i>CCL22, CCL2, CCL20, CXCL13, IL1B, CCL7</i>
GO:0070098	Chemokine-mediated signaling pathway	1.34×10^{-2}	<i>CCRL2, CCL22, CCL2, CCL20, CXCL3, CCL7</i>
GO:0071356	Cellular response to tumor necrosis factor	1.36×10^{-2}	<i>CCL22, CCL2, CCL20, PYCARD, COL1A1, THBS1, CCL7</i>
GO:0002548	Monocyte chemotaxis	1.87×10^{-2}	<i>CCL22, CCL2, FLT1, CCL20, CCL7</i>
GO:0071347	Cellular response to interleukin-1	1.99×10^{-2}	<i>CCL22, CCL2, CCL20, PYCARD, TFPI, CCL7</i>
GO:0071347	Cellular response to lipopolysaccharide	2.15×10^{-2}	<i>WNT5A, CD36, CCL2, CCL20, SERPINE1, PYCARD, TFPI</i>
KEGG pathway			
hsa04060	Cytokine-cytokine receptor interaction	1.12×10^{-2}	<i>CCL2, FLT1, CCL20, CRLF2, MET, IL1B, TNFSF14, TNFSF13, CCL7, IL1A</i>

图3 差异表达基因的验证

图4 *HBB*和*TNFRSF10C*基因诊断结核病的效能评价

为此,本研究基于基因芯片技术对受试者外周血进行全转录组表达谱分析,从中筛选出能用于结核诊断的分子标识。本研究共筛选出127个差异表达基因,其中77个基因在结核患者高表达,而50个基因在结核患者低表达(图1)。筛选出的部分差异表达基因,如*CCL20*, *CCL22*, *CXCL2*, *CXCL3*, *IL1B*, *TNF*编码的趋化

因子或者细胞因子已经被证实和结核病先天性免疫应答相关^[10-13]。然而,筛选出的差异表达基因大部分在以往研究并未发现和结核病相关,如网络版附表1最显著差异高表达的*SLC24A4*基因和牙釉质发育不全的钙离子转运有关^[14],显著性最低的高表达*MEGF9*基因和神经组织内的新型转膜蛋白编码相关^[15],网络版附表2最

显著差异低表达的*GPRC5A*基因是抑癌基因, 在COPD患者体内低表达^[16], 显著性最低的低表达*C15orf48*基因主要和肝癌转移有关^[17]。这提示, 结核病诊断标识的筛选还是有很大的空间, 并不只是局限在过往的研究结果中。

为进一步筛选出能用于指导监测抗结核疗效的基因, 将上述127基因中筛选治疗前后有表达差异的基因。发现其中有12个基因经抗结核治疗后表达量显著下降, 其中以*HBA1*和*HBA2*基因和血红蛋白- α 相关^[18]基因差异最为显著, 显著性最低的*AGTRAP*基因则和心血管疾病的病理生理相关^[19]。

本研究对差异表达基因进行了功能分析, 发现了10个差异表达基因富集的功能信号通路。和以往研究发现结核感染可导致机体复杂炎症应答的结果一致^[10,20], 表1 GO分析也发现结核病差异表达基因功能富集最显著的就是炎症应答, 也富集在和 γ 干扰素应答^[21]、IL-1应答^[22]、IL-6正调节、脂多糖应答^[23]等与细胞免疫应答相关的功能, 提示了炎症反应和细胞免疫应答均在结核抗感染免疫机制起关键作用。表1还发现, 各信号通路均有趋化因子或者细胞因子参与, 特别是*CCL20*, 已有研究发现人或者牛感染结核菌后单核巨噬细胞、胸膜腔积液的*CCL20*高表达, *CCL20*可以把IL-27⁺CD4⁺ T细胞聚集到胸腔积液, 促进胸膜间皮细胞的增殖^[24-27], 表1的*CCL2*, *CCL22*也显著富集在各通路上, 提示趋化因子和结核发病密切相关^[13]。而KEGG pathway分析结果则提示细胞因子在结核发病起了重要的作用^[28], IL-1, IL-2, IL-4, IL-10, IL-12, IL-17, IFN- γ ,

TNF- α 等细胞因子和结核密切关系已获得证实^[29,30], 上述差异基因功能分析发现了和结核病发病相关的信号通路, 一定程度揭示了结核病发病机制。

另外, 为验证所筛选的基因在临床应用具有可行性, 利用RT-qPCR方法对基因芯片结果进行了验证, 从127个差异基因中选取差异最大的88个基因, 发现部分基因RT-qPCR验证结果和基因芯片结果并不一致; 而大部分基因验证结果与基因芯片结果趋势一致, 其中8个基因*LRRK2*, *HBB*, *PRAMI*, *TNFRSF10C*, *MMP19*, *MET*, *SLC7A11*, *OLR1* RT-qPCR检测的结果与芯片结果一致性最好。另外绘制ROC曲线进一步评价所检测基因在诊断结核的效能。发现*HBB*和*TNFRSF10C*两个基因在结核患者比对照组表达高, 而且经抗结核治疗后有显著下降, 其中*HBB*基因在诊断健康对照和结核患者, AUC为0.98, 敏感性为100%, 特异性为95%; 在诊断潜伏感染和结核患者, AUC为0.985, 敏感性为100%, 特异性为95%。*TNFRSF10C*基因在诊断健康对照和结核患者, AUC为1.00, 敏感性为100%, 特异性为100%; 在诊断潜伏感染和结核患者, AUC为0.9775, 敏感性为95%, 特异性为100%, 上述结果提示, 所筛选的基因在结核诊断应用中有较好的潜力, 可以作为结核诊断的候选基因。

显然本研究样本量还不小, 也未将结核病和其他疾病进行对比研究, 但是筛选出一些有意义的和结核病相关的生物学标识, 为建立适合我国人群结核病诊断和疗效监测的分子生物学方法, 揭示结核病发病机制提供了新依据。

参考文献

- 1 Diel R, Nienhaus A. Aktuelle fragestellungen zum tuberkulose-screening mit interferon-gamma-release assays (IGRAs). *Pneumologie*, 2015, 69: 271-275
- 2 Chen X, Yang Q, Zhang M, et al. Diagnosis of active tuberculosis in China using an in-house gamma interferon enzyme-linked immunospot assay. *Clin Vaccine Immunol*, 2009, 16: 879-884
- 3 Zhang M, Wang H, Liao M, et al. Diagnosis of latent tuberculosis infection in bacille Calmette-Guérin vaccinated subjects in China by interferon-gamma ELISpot assay. *Int J Tuberc Lung Dis*, 2010, 14: 1556-1563
- 4 Chen Z B. Mycobacterium tuberculosis favors its survival by utilizing host ubiquitin to impair innate immunity. *Nat Sci Rev*, 2015, 2: 260-261
- 5 王睿, 施焕中. 干扰素- γ 释放试验在结核病诊断中的价值. *中国科学: 生命科学*, 2015, 45: 320-324
- 6 Mikhailovich V, Gryadunov D, Kolchinsky A, et al. DNA microarrays in the clinic: infectious diseases. *Bioessays*, 2008, 30: 673-682
- 7 Rachman H, Kaufmann S H E. Exploring functional genomics for the development of novel intervention strategies against tuberculosis. *Int J Med Microbiol*, 2007, 297: 559-567
- 8 Sweeney T E, Braviak L, Tato C M, et al. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet*

- Respir Med, 2016, 4: 213–224
- 9 Pankhurst L J, del Ojo Elias C, Votintseva A A, et al. Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study. [Lancet Respir Med](#), 2016, 4: 49–58
 - 10 Dorhoi A, Kaufmann S H E. Perspectives on host adaptation in response to *Mycobacterium tuberculosis*: modulation of inflammation. [Semin Immunol](#), 2014, 26: 533–542
 - 11 Lee J S, Lee J Y, Son J W, et al. Expression and regulation of the CC-chemokine ligand 20 during human tuberculosis. *Scand J Immunol*, 2008, 67: 77–85
 - 12 Francisco N M, Hsu N J, Keeton R, et al. TNF-dependent regulation and activation of innate immune cells are essential for host protection against cerebral tuberculosis. [J Neuroinflamm](#), 2015, 12: 125
 - 13 Slight S R, Khader S A. Chemokines shape the immune responses to tuberculosis. [Cytokine Growth Factor Rev](#), 2013, 24: 105–113
 - 14 Parry D A, Poulter J A, Logan C V, et al. Identification of mutations in SLC24A4, encoding a potassium-dependent sodium/calcium exchanger, as a cause of amelogenesis imperfecta. *Am J Hum Genet*, 2013, 92: 7–12
 - 15 Brandt-Bohne U, Keene D R, White F A, et al. MEGF9: a novel transmembrane protein with a strong and developmentally regulated expression in the nervous system. [Biochem J](#), 2007, 401: 447–457
 - 16 Zhou H, Rigoutsos I. The emerging roles of GPRC5A in diseases. [Oncoscience](#), 2014, 1: 765
 - 17 Fang T T, Sun X J, Chen J, et al. Long non-coding RNAs are differentially expressed in hepatocellular carcinoma cell lines with differing metastatic potential. [Asian Pac J Cancer Prevent](#), 2014, 15: 10513–10524
 - 18 Higgs D R, Weatherall D J. The alpha thalassaemias. [Cell Mol Life Sci](#), 2009, 66: 1154–1162
 - 19 Wakui H, Dejima T, Tamura K, et al. Activation of angiotensin II type 1 receptor-associated protein exerts an inhibitory effect on vascular hypertrophy and oxidative stress in angiotensin II-mediated hypertension. *Cardiovasc Res*, 2013, 100: 511–519
 - 20 Caccamo N, Dieli F. Inflammation and the coagulation system in tuberculosis: tissue factor leads the dance. [Eur J Immunol](#), 2016, 46: 303–306
 - 21 Berry M P R, Graham C M, McNab F W, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. [Nature](#), 2010, 466: 973–977
 - 22 Mayer-Barber K D, Andrade B B, Oland S D, et al. Host-directed therapy of tuberculosis based on interleukin-1 and type I interferon crosstalk. [Nature](#), 2014, 511: 99–103
 - 23 Mishra A K, Driessen N N, Appelmelk B J, et al. Lipoarabinomannan and related glycoconjugates: structure, biogenesis and role in *Mycobacterium tuberculosis* physiology and host-pathogen interaction. [FEMS Microbiol Rev](#), 2011, 35: 1126–1157
 - 24 Ye Z J, Xu L L, Zhou Q, et al. Recruitment of IL-27-producing CD4⁺ T cells and effect of IL-27 on pleural mesothelial cells in tuberculous pleurisy. [Lung](#), 2015, 193: 539–548
 - 25 Plessers E, Wyns H, Watteyn A, et al. Characterization of an intravenous lipopolysaccharide inflammation model in calves with respect to the acute-phase response. [Veterin Immunol Immunopathol](#), 2015, 163: 46–56
 - 26 Rivero-Lezcano O M, González-Cortés C, Reyes-Ruvalcaba D, et al. CCL20 is overexpressed in *Mycobacterium tuberculosis*-infected monocytes and inhibits the production of reactive oxygen species (ROS). [Clin Exp Immunol](#), 2010, 162: 289–297
 - 27 Lee J S, Lee J Y, Son J W, et al. Expression and regulation of the CC-chemokine ligand 20 during human tuberculosis. *Scand J Immunol*, 2008, 67: 77–85
 - 28 Mayer-Barber K D, Sher A. Cytokine and lipid mediator networks in tuberculosis. [Immunol Rev](#), 2015, 264: 264–275
 - 29 Chandrashekar S, Anupama K R, Sambarey A, et al. High IL-6 and low IL-15 levels mark the presence of TB infection: a preliminary study. [Cytokine](#), 2016, 81: 57–62
 - 30 John S H, Kenneth J, Gandhe A S. Host biomarkers of clinical relevance in tuberculosis: review of gene and protein expression studies. [Biomarkers](#), 2012, 17: 1–8

Tuberculosis diagnostic markers based on the whole gene transcriptome

CHEN XianXiong, YANG QianTing, QIU ZhenGang, ZHU JiaLou, ZHANG MingXia,
DENG GuoFang, DENG QunYi & CHEN XinChun

1 Department of Pathogenic Biology, School of Basic Medical Sciences, Health Science Center, Shenzhen University, Shenzhen 518060;

2 Key Laboratory of Infection and Immunity, the Third People's Hospital of Shenzhen, Shenzhen 518112;

3 The First Affiliated Hospital of Gannan Medical University, Gannan 341400

Early diagnosis is essential for the timely treatment of tuberculosis patients. At present, there is a lack of ideal clinical diagnostic methods for this disease. Therefore, rapid and simple diagnostic methods are urgently needed. This study analyzed expression of the whole genome transcript of nine tuberculosis patients before treatment, six cases of latent tuberculosis infection, and six healthy controls. The expression profiles of tuberculosis patients were followed before and after treatment. The results showed that, compared with healthy controls and patients infected with latent tuberculosis, 77 genes were highly expressed in tuberculosis patients and 50 genes were expressed at low levels. In response to different periods of anti-tuberculosis therapy (0, 3, and 6 months), the expression of 12 genes was reduced. Gene Ontology enrichment and Kyoto Encyclopedia of Genes and Genomes biological pathways analyses identified nine biological processes related to the inflammatory response, chemotaxis, and the cellular immune response. In addition, one signaling pathway was associated with a cytokine-cytokine receptor interaction. Chemokines and cytokines were involved in each signaling pathway. Our results were verified in additional subjects using the quantitative real-time polymerase chain reaction. Specifically, most measured mRNA levels were consistent with the microarray data. In the eight genes with the best consistency, two (*HBB* and *TNFRSF10C*) were ideal biomarkers for diagnosing tuberculosis. The receiver operating characteristic curve and the area under the curve were above 0.9, and the sensitivity and specificity were over 95% for these genes. In summary, two differentially expressed genes were identified as novel biological markers for tuberculosis. They provide a new basis for the establishment of diagnostic methods, efficacy monitoring, and revealing the pathogenesis for this disease.

tuberculosis, whole-gene transcriptome, diagnosis, biomarker

doi: [10.1360/N052017-00225](https://doi.org/10.1360/N052017-00225)