猪的主要组织相容性复合体表达谱分析

姜范波 陈 晨 邓亚军 于 军 胡松年 *

(中国科学院北京基因组研究所, 北京 101318; 中国科学院研究生院, 北京 100038.* 联系人, E-mail: husn@genomics.org.cn)

摘要 为了全面地了解 SLA 在各组织的表达情况,深入理解 SLA 的功能,利用"家猪基因组计划"中得到的大量 EST,构建了丹麦长白猪 51 种组织的 SLA 表达谱,并进行了若干组织的不同发育阶段以及丹麦长白猪和中国二花脸猪 SLA 表达水平的比较.在确定该方法可靠的同时,发现经典的抗原呈递基因高表达于免疫组织和消化道的中段;类似于 HLA-C, SLA a类基因 SLA-3 的表达丰度和组织表达模式明显不同于 SLA-1 和 SLA-2. 在进行品系间比较时发现,除空肠外,二花脸猪的其他组织抗原呈递基因的表达水平均高于丹麦长白猪.前者的 MHC 基因在多种组织(包括免疫系统、肾上腺等)的高表达可能是二花脸猪高抗逆性的基础之一,而后者的 MHC 基因在空肠(上皮细胞)的高表达可能是丹麦长白猪高生长速度、高饲料回报率的一个基础.

关键词 猪白细胞抗原 主要组织相容性复合体 表达序列标签 表达谱 猪

主要组织相容性复合体(major histocompatibility complex, MHC)见于所有的脊椎动物基因组,在进化的过程中变得越来越复杂. 哺乳动物中, MHC分为 , 3 部分, 其中 类区在进化上变异最大, 和 类区相对保守. 猪的MHC又称猪白细胞抗原(swine leukocyte antigen, SLA, 以下简称SLA), 位于7号染色体, 跨着丝粒, 长约2 Mb^[1]. 目前已经获得 类区的部分基因组序列(总长约0.9 Mb)^[2-5] 和 类区通过用猪或/和人以及小鼠的已知基因作为探针与猪的基因组杂交等方法也已基本清楚其基因组结构^[5,6].

类区的基因可以分为 4 个亚类: a. b. c和组织相容不相关(non-MHC)基因. 其中 a即经 典的 类基因含 3 个有功能的基因SLA-1~3. 该亚类 基因以不同丰度表达于所有有核细胞的表面、结合 并向CD8+ T细胞呈递内源性抗原. b类基因包括 SLA-6~8, 它们的表达具有组织特异性, 功能还不是 很清楚, 只能根据对它们的HLA 同源基因的研究推 测它们可能介导机体的免疫耐受(如母胎耐受). 类基因包括MICI 和MIC2、前者可能是假基因、这类 基因的组织表达和功能都有待研究、对人的该类基 因研究表明、其表达产物是γδT细胞和NK细胞表面 NKG2D受体的配体, 二者的结合导致表达MIC分子 的细胞溶解, 可能是机体的一种免疫监督机制. SLA 类区还包含 30 多个non-MHC基因、它们的功能各 异、其中有一些成为进化上高度保守的"框架基因",

框架基因之间成为 a, b, c类基因复制、变异、基因间重组的"容纳区"^[2]. SLA 类区基因包括 3 个亚类:() 经典的 类基因,指 SLA-DQA,-DQB,-DRA,-DRB及-DPA基因,它们编码的MHC 类分子表达于抗原呈递细胞的表面,结合外源性抗原,呈递给CD4+ T细胞.但SLA-DPA没有发现转录本^[8],可能是个假基因;() 类样基因,包括 SLA-DM和 SLA-DO基因,它们在序列上与经典的 类基因相似,但不表达于细胞表面,而是作为一种分子伴侣,在外源性抗原的处理过程中,协助经典的 类分子形成正确的折叠;() non-MHC基因,部分该类基因在内源性抗原的处理和转运中有重要作用(如TAP1, TAP2,TAPASIN, LMP等),其他基因功能各异. SLA 类区基因主要编码补体成分、细胞因子等先天免疫成分.

由于SLA 区域涵盖了先天免疫与后天免疫的成分,与疾病有密切的关联,表现在不同的单体型对疾病有不同的易感性^[9]. 此外,SLA 的单体型还与猪的生产性状,如繁殖力^[10]、猪肉品质等有关,因此,SLA 越来越受到育种学家的重视. MHC在同种和异种移植中都有重要的意义^[11,12],猪作为对人异种器官移植最好的供体,深入了解SLA 将为人类攻克异种器官移植这一难关提供帮助. 虽然对SLA的研究越来越多,但系统地研究SLA的表达尚未见文献报道. 为了系统地了解SLA的表达情况,本研究通过分析中丹家猪基因组计划中得到的大量表达序列标签(expressed sequence tag, EST),构建了丹麦长白猪 51 种组织的

SLA表达谱、并进行了部分组织的不同品系和不同发 育阶段的SLA表达谱比较、以期深入地了解SLA的功 能.

材料与方法 1

结肠

() EST 数据. 所有 EST 数据均来自中丹家猪 基因组合作计划。其组织来源包括成年丹麦长白猪 的 51 种组织、成年中国二花脸猪的 14 种组织和来自 丹麦长白猪不同发育阶段的若干种组织, 所有的 cDNA 文库均为非均一化文库, 因此随机挑选克隆进 行大量测序能够得到基因的表达丰度信息. 单向一 次测序得到的 EST 去除载体序列后, 模糊碱基数少 于 5%, 长度在 100 bp 以上者为高质量 EST, 纳入下 一步的分析. 各组织的高质量 EST 数从 2208 条(前列 腺)到 10675条(大肠), 平均值为 7000条. 组织的取样 部位及各组织的高质量 EST 数见表 1. 其中的部分 EST 序列, 经本研究的分析可以明确比对到 10 个抗 原呈递基因(表 2)、在 GenBank 的登录号为: DN593918~DN596190.

() SLA 基因组及mRNA数据. 对SLA基因组 的研究[2~6,8]已经明确了SLA 各部分的基因名称和数 目. 其中只有 10 个基因的猪源性的全长mRNA序列 可以从公共数据库中找到,它们是: a类的SLA-1~3 和 b类的SLA-6~8 以及经典的 类基因SLA-DQA, -DOB、-DRA、-DRB. 这已经包括了除MIC2 之外的所 有SLA抗原呈递分子的编码基因、需要说明的是、上 述 10 个基因的序列在基因组上是惟一的. 另外 74 个 基因可以找到人源性的全长mRNA.

这 2 部分全长 mRNA 分别作为 EST 比对的数据集. MHC 基因的一个特点是, 在与病原物共同进化

		表	₹1 组	织取样	部位、	所属系							
系统	组织	LF50	LF100	LN115	LA	EA	系统	组织	LF50	LF100	LN115	LA	EA
	胸腺		5771		7247		消化系统	直肠				8386	
免疫系统	脾				7525	6297	7月16次36	肝	6678	5008		6967	6637
	支气管淋巴结				7107			前列腺				2208	
	半腱肌	8051		5251	7997			睾丸				5140	6137
	半膜肌				6808			卵巢				7574	7731
	肱三头肌				6595		生殖系统	乳腺(产初乳)				7406	3642
	冈上肌				7010			乳腺(哺乳)				5785	
肌肉	冈下肌				6141			乳腺(断奶后7d)				6045	
13/61-9	股中间肌				6993			子宫				7626	5737
	股二头肌				6312		泌尿系统	肾				8667	5967
	最长肌				10421		750 757 75	膀胱				8090	
	咬肌				4836			甲状腺				7589	
	左心室				7229		内分泌系统	肾上腺				8287	6680
	大脑皮层	8863	6245		6324	7853	13737678.26	腺垂体				9068	
	小脑	7198	7444			6357		胰腺				6808	
	脑干	8669	5896				循环系统	心脏				4883	7055
神经系统	下丘脑				7201		110 1 27 20	大动脉				5470	
111-57 37 30	海马	6019					造血系统	血液					8993
	脊髓				7695		~===	骨髓			10071		
	视网膜				8386			胎盘				7522	
	眼				6535			羊膜				2617	
呼吸系统	肺	7549		5241	6762			脐带	8331				
	气管				9031			表皮	8195		8451	6940	6970
	唾液腺				5279			脂肪				7336	
	食道				5685		其他	黏膜		8229			
	胃				5153	6870		鼻腔黏膜					4896
消化系统	十二指肠				5835			舌尖黏膜				5749	
	空肠	10210		9503	10276	5834		软骨			7373		
	回肠				5693			关节囊				8809	

a) LF50 示长白猪 50 d 胚胎; LF100 示长白猪 100 d 胚胎; LN115 示 115 d 长白猪幼猪; LA 示成年丹麦长白猪; EA 示成年中国二花脸猪

纵隔

8818

660 www.scichina.com

5201

6242

基因	GeneBank 登录号												
	AY135592.1	AY135587.1	AY247767.1	AY459298.1	AF464036.1	AY135591.1							
SLA-1	AF464045.1	AY135600.1	AF464043.1	AY135594.1	AY459297.1	AY459299.1							
	AY135595.1	AF464016.1	AY459306.1	AF464015.1	AF464013.1								
	AY247774.1	AY247775.1	AY247773.1	AF464023.1	AF464049.2	AF464039.1							
SLA-2	AY135599.1	AF464059.2	AF464058.1	AY135598.1	AF464004.1	AF464005.1							
	AF464003.1												
SLA-3	AY247769.1	AY247770.1	AF464042.1	AF464041.1	AF464040.1	AF464012.1							
SLA-3	AY247768.1	AF464011.1	AY135602.1	AF464010.1	AF464009.1								
SLA-6	AF464019.1	AY459304.1	AF464020.2	AY247772.1	AY247771.1	AF464008.1							
SLA-0	AF464007.1	AF464006.1											
SLA-7	AY463541.1												
SLA-8	AY463542.1												
	AY243101.1	AY243105.1	AY135574.1	AF464024.1	AF464028.1	AY135569.1							
SLA-DQB	AY191778.1	AF464025.1	AF464026.1	AY102476.1	AY102477.1	AF464047.1							
SLA-DQB	AB009659.1	AF464061.1	AY135571.1	AY135572.1	AF464038.1	AF464027.1							
	AY135568.1	AF464029.1	AY102478.1	AF464030.1									
	AY243103.1	AY243107.1	AY135581.1	AY135579.1	AY135582.1	AY135583.1							
SLA-DRB	AY191776.1	AY135578.1	AY102479.1	AY102480.1	AF464050.1	AY135575.1							
SLA-DKD	AF464051.1	AY126721.1	AF464060.1	AY135584.1	AF464032.1	AF464033.2							
	AY102481.1	AF464034.1											
SLA-DRA	AY243102.1												
SLA-DQA	AY243100.1	AY243104.1	AY191777.1	AY102473.1	AY102474.1	AY102475.1							

表 2 SLA 10 个抗原呈递基因的 mRNA 序列在 GenBank 的登录号

的过程中,多个位点(主要是抗原结合位点)产生碱基突变,使得几乎每一个抗原呈递分子编码基因都有众多的等位基因. 考虑到本研究中,等位基因型并不清楚,因此以 GenBank 中能找到的所有等位基因作为该基因的一个集合(表 2). 由人源性 mRNA 构成的74 个基因中,部分基因也有多个等位基因和可变剪接,这些均放在比对集中. 通过对猪的 10 个抗原呈递基因的100 个等位基因序列进行 Clustal W 分析发现,等位基因之间形成很清晰的聚类关系. SLA 的10 个抗原呈递基因与 HLA 的对应基因的 ClustalW 分析显示,同一物种来源的基因聚在一起,因而从序列相似性难以直接判断基因间的同源对应关系(结果未显示).

() 比对参数. 由于EST比对的对象分成猪源性和人源性2部分, 因此相应的比对方法也分成2种. 对猪源性mRNA, 采用与BLAST相似, 但运行更快的BLAT程序^[13]直接在核苷酸水平上比对, 参数为 100 bp(即连续比对长度大于等于 100 bp). 一条EST若同时比对上多个基因,则以精确比对长度最长者为其对应者(即加上max参数). 以同一个基因的多个等位基因所对应的EST数相加作为该基因的EST总数. 表达丰度的计算公式为: 表达丰度 = 基因的EST总数×1000/所在文库的高质量EST总数. 对人源性的

mRNA, 由于核苷酸水平上可能存在较大的种间差异, 因此比对在蛋白水平上进行(人与猪之 间蛋白水平的相似性在 60%~80%). 比对程序为TBLASTX, 参数仍选用 100 bp, 同时也用到了max参数. 等位基因和可变剪接的处理与猪源性部分相同.

() 统计检验方法和基因的表达聚类. 在进行品系间及不同发育阶段表达水平差异的比较时, 2 个组织之间的比较用Audic和Claverie [14]提供的方法进行显著性检验; 涉及到 3 个或更多组织时, 采用总体 χ^2 检验 [15,16]. Romualdi和Bortoluzzi [17]的网站上提供了很方便的统计检验软件IDEG6(Identification Differentially Expressed Genes 6 test statistic, http:// telethon.bio.unipd.it/bioinfo/IDGE6_form/index.html).

在得到各基因在不同组织的丰度信息后,借鉴了基因芯片研究不同环境下某种组织若干基因的表达谱差异的分析方法,即以不同环境为横坐标,所研究的基因为纵坐标,得到一个方阵,以此方阵为基础,根据表达谱的变化对基因进行聚类,上调/下调表达越一致的基因距离越近. 本研究中,以成年丹麦长白猪的 51 种组织为横坐标,以 SLA 区域的基因为纵坐标,利用加州大学伯克利分校和伯克利国家实验室的免费软件 Cluster (http://rana.lbl.gov/downloads/Cluster.zip)得到表达丰度的方阵. 之后,利用该软件,

对 10 个抗原呈递基因进行了基于表达水平的聚类分析, 在 51 种组织中表达水平越一致的基因, 其上游调控机制应越相似, 基因的功能也应较接近.

2 结果

2.1 由猪源性 mRNA 代表的 10 个基因

()组织表达谱. 图1显示的是这10个基因的组织表达概况. 从图中可以看出: (1) 10 个基因在51种组织中的表达水平是迥异的. 少数组织, 如支气管淋巴结、关节囊、十二指肠、回肠、空肠、结肠等表达水平较高, 而其他大部分组织的表达水平很低; (2)基因的表达水平依类别而不同. a类的表达水平最高, 经典的 类基因其次, b类最低; (3) 10 个基因在51 个组织的表达模式在大体相似的基础上有细微的差别. 它们的高表达组织均主要集中免疫系统和消化道, 但即使在同一个亚类内, 各基因的表达模式也有一些细微的差别. 最典型的是 a类基因 SLA-3, 它的表达丰度和表达组织均不同于另外2个 a类基因.

SLA a 的 3 个基因中, SLA-1 和 SLA-2 在 51 种组织中的表达水平几乎总是相似的,它们在呼吸、循环、神经、肌肉、泌尿、生殖、内分泌系统均呈较低的表达,各组织丰度不超过 1‰(内分泌系统的腺垂体 SLA-1 略超,为1.21‰);在免疫系统的支气管淋巴结、消化系统的十二指肠、结肠以及纵隔、关节囊中呈较高丰度,在 2‰~3‰左右;而在消化系统的回肠和空肠显示了突出的高表达: SLA-1 在回肠和空肠中分别为 7.90‰和 9.73‰, SLA-2 则分别为 8.78‰和10.51‰.

SLA-3 虽然与 SLA-1, SLA-2 同属 a 类, 但是它的表达水平明显低于二者, 除了在十二指肠、回肠和肝中分别为 1.54‰, 1.58‰和 1.44‰之外, 在其他组织表达水平均低于 1‰. 但是, 在某些 SLA-1, SLA-2 没有检测到 EST 的组织, 如肺、气管、膀胱、肾、肾上腺、前列腺、表皮和某些肌肉组织, SLA-3 反而有表达, 此外, 在肝中, SLA-3 的表达也明显高于 SLA-1, SLA-2.

SLA b 类基因在所有组织的表达水平均低于 0.25%, 且分布在有限的几个组织. SLA-6 见于纵隔、 关节囊、脾、空肠、食道和左心室, SLA-7 仅见于脾和纵隔, SLA-8 见于 9 种组织: 支气管淋巴结、结肠、 关节囊、心脏、胎盘、唾液腺、卵巢、子宫、和肾. 这些表达 SLA b 的组织中, 有的也同时高表达经典的 SLA 基因(指 SLA a 和经典的 SLA II类基因), 如纵隔、关节囊、脾、空肠、支气管淋巴结、结肠等, 但是有一些组织如胎盘,一般认为不表达经典的 SLA 基因(本研究也没有检测到该组织有该类基因的 EST), 而 SLA-8 在其中有表达,提示它很可能就是人的 HLA-G 的对应体:表达于胎盘的绒毛上皮,通过抑制 NK 细胞的活性介导母胎耐受.

与 a 类基因相似, 经典的 SLA 类基因的高表达组织也集中分布于免疫系统和消化道, 在生殖系统的子宫和泌初乳乳腺以及关节囊和纵隔中也较高, 而呼吸、循环、肌肉、神经、泌尿、内分泌均呈低或无表达. 丰度最高者为支气管淋巴结的 *SLA-DRB*, 达3.10‰, 但这个值仅为 a 类基因的最高值的 1/3.

为了便于更详细地了解经典的 SLA 基因的组织

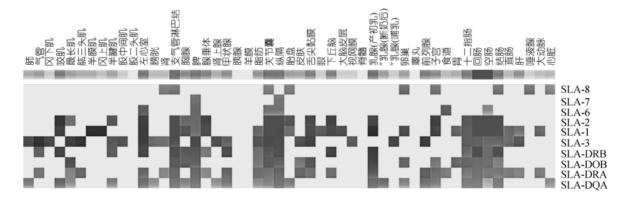


图 1 成年丹麦长白猪 SLA 的 10 个抗原呈递基因在 51 种组织的表达谱 不同颜色代表 EST 丰度, 黑色为高表达, 白色为无表达. 黑色成分越多, 代表丰度越高. 该图用加州大学伯克利分校 和伯克利国家实验室的免费软件 Cluster 绘制(http://rana.lbl.gov/downloads/Cluster.zip)

表达情况, 我们设定出一个"平均丰度", 即以该基因在 51 种组织中的 EST 总数 × 1000/51 种组织的高质量 EST 总数, 相当于该基因在全身的平均值. 7 个经典的 SLA 基因(指 SLA a 和经典的 SLA 类基因, 下同)的平均丰度见表 3.

表 3 经典 SLA 基因的平均丰度和最高丰度

基因	EST 总数	平均丰度(‰)	最高丰度(‰)
SLA-DQA	86	0.24	1.23
SLA- DRA	74	0.20	1.36
SLA- DQB	71	0.19	1.69
SLA- DRB	94	0.26	3.10
SLA-1	271	0.74	9.73
SLA-2	282	0.77	10.51
SLA-3	98	0.27	1.58

从该表中可以看出, 经典的 SLA 类基因具有相似的平均丰度, 并且明显低于 SLA-I 和 SLA-2. SLA-3 的平均丰度更接近于经典的 SLA 类基因, 而不是它所属的 a 类. 以平均丰度的 2 倍为界线, 人为的将 51 种组织分为 2 类: (1) 低表达, 丰度低于 2 倍平均丰度, 包括没有检测到 EST 的情况; (2) 高表达, 丰度值高于 2 倍平均丰度. 各基因具体的组织分类情况见表 4.

从表 4 可以看出, 这 7 个基因的高表达组织数相 仿. 表 5 列出了这些高表达组织及各基因的表达丰度. 14 种高表达的组织中, 3 个来自免疫系统, 6 个来自消化系统. 这 6 个消化系统的组织中, 除肝外, 其他均为肠道(从十二指肠到结肠). 其余 5 个组织包括纵隔、关节囊、乳腺、脂肪、前列腺. 纵隔包含大量的

表 4 按经典的 SLA 基因的 EST 丰度水平进行组织分类

基因	丰度水平							
	<2 平均	>2 平均						
SLA-DQA	44	7						
SLA- DRA	43	8						
SLA- DQB	44	7						
SLA-DRB	45	6						
SLA-1	45	6						
SLA-2	44	7						
SLA-3	42	9						

淋巴结,关节囊的滑膜层含有一种M细胞,电镜下类似巨噬细胞,含有较多溶酶体,有吞噬力,可能也是一种抗原呈递细胞.这2个组织高表达经典的 SLA 基因应该是由所含的免疫细胞造成的.泌初乳的乳腺、脂肪和前列腺均仅有经典的 SLA 类基因高表达,而 SLA a类基因并非高表达.造成这种高表达的原因可能是专职的抗原呈递细胞,也可能是非专职的抗原呈递细胞,如血管内皮细胞.

根据这 10 个基因在 51 种组织中的表达丰度和表达模式, 利用 cluster 软件对它们进行聚类的结果见图 2.

前面的分析中已经显示, SLA-1 与 SLA-2 具有相似的表达丰度和组织表达模式, 因此这 2 个基因最先聚类, 编码经典 SLA 类分子 α 链的 2 个基因 SLA-DRA 和 SLA-DQA 在平均丰度、最高丰度、无-低-高表达组织数方面均很相似,但高表达的组织种类相差较大,它们以较远的距离聚在一起,然后与编码 β 链的 2 个基因以及 SLA-3 再聚到一起; SLA-3 是个特殊的 a 类基因、它的丰度水平与 SLA 类基因更相

表 5 经典的 SLA 基因的高表达组织及其表达丰度(‰)

组织	基因													
20.57	SLA-DQA	SLA-DRA	SLA-DQB	SLA-DRB	SLA-1	SLA-2	SLA-3							
支气管淋巴结	0.98	1.27	1.69	3.10	2.95	3.38	0.84							
	1.23	1.23	1.23	2.11	7.90	8.78	1.58							
十二指肠	1.03	0.51	1.37		1.54	1.71	1.54							
回肠		0.39	0.88	1.07	9.73	10.51	0.68							
结肠	0.96			2.11	2.88	2.31	0.58							
关节囊			0.45	0.69	1.93	2.04	0.79							
纵隔	1.02	0.91		0.91		2.49	0.68							
脾	1.06	1.06					0.66							
胸腺			0.41											
乳腺(产初乳)	1.08	0.41												
脂肪			0.66											
前列腺		1.36												
肝							1.44							

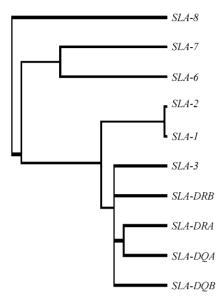


图 2 SLA a, b 和经典 SLA 类基因的表达聚类

似, 而组织分布更倾向 a 类, SLA-6 和 SLA -7 因为极 低的表达水平自成一类, 而 SLA-8 以其独特的表达组 织成为最后一个被聚类进来的成员.

() 丹麦长白猪部分组织的不同发育阶段表达 水平变化. 中丹家猪基因组计划还获得了若干组织 不同发育阶段的 EST 测序数据(表 1). 通过比较同一 个组织的不同发育阶段/生理时期 SLA a SLA b 及经典的 类基因的表达水平变化、将有助于了解 这些基因的表达动态. 发育阶段包括胚胎 50, 100 和 幼猪 115 d 以及成年猪. 此外、乳腺组织的产初乳、 常乳和断奶后 7 d 的数据也进行了比较.

SLA b 的 3 个基因在所有组织中表达水平都 很低、未发现有不同发育时期的差异表达、SLA a 和经典的类基因在脑干、小脑、大脑皮层、半腱肌、 表皮等 5 种组织的各发育阶段表达水平很低, 未检测

到显著性差别。在其他6种组织中则有不同程度的表 达差异, 总体来说, 这些基因在胸腺的胚胎 100 d, 肺 的出生后 115 d 表达为最高、而结肠、肝和空肠以成 年为最高. 泌初乳的乳腺的表达水平要显著高于泌 常乳和断奶后的乳腺(表 6).

() 二花脸猪与长白猪的表达差异。 丹麦长白 猪和中国二花脸猪作为2个不同的品系,有不同的畜 牧学特性. 它们的这 10 个基因的表达是否也存在差 异?

对二花脸猪进行 EST 测序的组织有 14 个(表 1). 前面的分析已经显示了SLA b类的3个基因在丹麦 长白猪中是低表达的、它们在中国二花脸猪中表达 水平也不高, 因此, 大多数组织没有发现品系间表达 差异, 仅 SLA-6 略有例外. 在下列 4 个组织中, 二花 脸的表达水平均高于长白猪(P<0.05). 它们是: 肾 (3 0)、空肠(3 1)、淋巴腺(6 0)、脾(5 1). SLA a 和经典的 类基因在表皮、大脑皮层、睾丸、 子宫和肾这 5 种组织中均低水平表达(各基因的 EST 数均不超过3个), 未发现显著性差别, 其他9个组织 中、有的基因表达水平存在显著性品系间差别。总体 来说、除了肝的SLA-3和空肠中丹麦长白猪的表达水 平高于中国二花脸猪外, 其他组织中情况均相反, 详 情见表 7.

2.2 SLA 区域其他基因的表达概况

除了前面讨论的 10 个基因, SLA 区域还有 74 个 其他基因. 它们的组织表达见图 3. 其特点是:() 大部分的基因表达丰度很低:()少数高丰度的基 因表达也没有明显的组织特异性,不像前面的 10 个 基因一样集中于免疫组织和消化道. 根据在 51 种组 织的 EST 总数的多少可以将这 74 个基因分成 3 类: EST 数为 0 的有 20 个(包括 DPA1 和 SEEK1, 一直没有

			র	₹6 £	麦长日	3猪不同	可发育的	ì段经典	!的 SLA	A 基因	差异表 :	达的组织	Д ^{а)}			
组织	结	结肠 肝				空肠			腺		肺		乳腺			
发育阶段	LN115	LA	F50	F100	LA	F50	LN115	LA	F100	LA	F50	LN115	LA	初乳	常乳	断奶
EST 总数	6242	5201	6678	5008	6967	10210	9503	10276	5771	7247	7549	5241	6762	7406	5785	6045
SLA-DQA	7	5	0	0	0	0	2	1	7	2	0	0	1	8**	0	2
SLA-DQB	1	1	0	1	0	1	0	9**	4	3	0	0	0	3	0	0
SLA-DRA	2	2	0	0	1	1	3	4	3	0	0	2	0	3	0	0
SLA-DRB	4	11**	0	0	0	0	1	11**	5	3	0	3	0	3	0	0
SLA-1	6	15**	0	0	1	2	6	100**	5	2	2	5**	0	3	0	0
SLA-2	6	12*	0	1	0	0	7	108**	8**	1	1	14**	0	7**	0	0
SLA-3	4	3	0	0	10**	0	1	7**	1	0	0	5	1	0	1	0

a) 数字为基因对应的 EST 数. **示 P<0.01, *示 P<0.05

	空肠		淋巴腺 卵巢		朋	脾		肝		乳腺		上腺	胃		心脏			
	E	L	Е	L	Е	L	Е	L	Е	L	Е	L	Е	L	Е	L	Е	L
EST 总数	5834	10276	8397	7107	7731	7574	6297	7525	6637	6967	3642	5785	6680	8287	6870	5153	7055	4883
SLA-DQA	0	1	36**	7	4	1	23**	8	2	0	0	0	15**	1	0	0	0	1
SLA-DQB	1	9*	50**	12	1	1	30**	3	0	0	0	0	13**	0	2	2	2	0
SLA-DRA	3	4	34**	9	4	0	22**	8	2	1	1	0	13**	0	0	0	1	0
SLA-DRB	2	11*	53**	22	3	0	34**	3	3*	0	1	0	18**	0	0	0	0	0
SLA-1	15	100**	29	21	31**	1	32**	1	22**	1	5**	0	46**	0	7*	1	5*	0
SLA-2	4	108**	22	24	23**	0	41**	1	19**	0	7**	0	35**	0	4	0	1	0
SLA-3	2	7*	9	6	2	0	13**	5	2	10**	1	1	10**	2	2	0	4	0

表 7 成年中国二花脸猪和丹麦长白猪的经典 SLA 基因差异表达组织 a)

发现前者转录本证据,后者已经明确是假基因),40 条以下的有 42 个,所有的框架基因都属于这 2 类. EST 总数超过 40 条的基因只有 10 个. 它们是(括号中为 51 种组织中的 EST 总数): OK-SW-CL.56(273), HSPA1B(194), PRDMI(194), HSPA1A(182), COL11A2 (111), BATI(98), TNX(54), PSMB8(56), CLICI(48), PSMB9(41). 值得说明的是,在人的基因组上,这些基因中只有 PSMB9, PSMB8, TNX 和 PRDMI 的序列是惟一的,其他基因有至少 2 个相似性很高的同源物, OK-SW-CL.56 甚至与来自 13 条染色体的 17 个 contig可以比对上. 因此,这些基因的 EST 数反映的是它们所属同源基因家族的转录水平,而不是该基因的表达水平. 由于同源基因之间的高度相似性,目前没有办法将 EST 准确地对应到各个基因,因此对这些基因本研究不拟作进一步的分析.

3 讨论

本研究利用 EST 数据构建了丹麦长白猪 51 种组织的 SLA 表达谱,并进行了部分组织的不同发育阶段和品系之间的比较. 总体来看, SLA 区域的基因大部分为低水平表达. 表达丰度较高的少数基因包括 2部分:() 直接参加抗原呈递的 SLA a 和经典的 SLA 类基因,选择性地在某些组织高表达.() 非抗原呈递基因,表达不具组织特异性.

经典的SLA基因高表达的组织主要为免疫系统和消化系统,尤其在消化道的中段.免疫系统的脾、胸腺、支气管淋巴结富含免疫细胞,SLA a和经典SLA 类基因的高表达正是它们免疫功能的反映.对消化道来说,肠道黏膜的固有层中含有大量分散的淋巴细胞和淋巴小结(在十二指肠和空肠多为孤立淋巴小结,在回肠多为集合淋巴小结,结肠的黏膜固有层中也含有散在的孤立淋巴小结.肠道作为消

化吸收食物的主要场所, 是接触异种蛋白和病原物 的前线、大量的抗原呈递细胞参与对外来抗原的处 理和呈递可能是肠道高表达经典SLA 类基因的原 因. Oliver等人[18]利用单克隆抗体对人的研究发现, 小肠黏膜固有层和集合淋巴结有大量DR+、DP+和 DQ+细胞、而且主要集中在消化道的中段。而消化道 高表达SLA a类基因除了免疫细胞的成分、肠道管 腔上皮也是重要原因. 首先, 在回肠和空肠中, SLA-1、SLA-2 的表达水平为免疫组织(支气管淋巴结) 的 2~3 倍, 并且这 2 个组织中经典的SLA 类基因表 达水平与SLA-1 和SLA-2 的表达水平极不相称. 其次、 Metzger^[19]以经典MHC / 类分子的单克隆抗体对 大鼠的多种组织MHC表达谱进行研究已经发现、除 了免疫系统的组织和某些血管内皮细胞高表达这两 类分子外, 类分子还可见于多种上皮细胞,包括胃 肠道的管腔上皮.

在对不同发育阶段的组织比较中, 肝、结肠与回肠的 SLA 基因高表达都是成年以后的事, 说明这些组织的 SLA 的表达与其功能发育过程相平行; 胸腺是 T细胞发育和分化的场所, 在性成熟后逐渐萎缩退化, 本研究中成年猪的胸腺 SLA 表达水平低于胚胎 100 d 时, 也正反映了这种趋势. 初乳中除了高浓度的免疫球蛋白外, 其脂肪、维生素等营养成分的含量也有别于常乳, 这个时期的乳腺合成和分泌活动有它自己的特点, 但此时 SLA 的表达水平也不同于其他时候未见文献报道, 有何生理学意义有待进一步研究. 此外, 出生后 115 d 的猪肺 SLA 的表达达高峰, 可能与该时期肺组织较快的生长速度有一定的关系.

综上所述,本研究的结果与已有的知识是一致的,也说明以 EST 方法来研究 SLA 的组织表达谱是可行的.

a) 数字为基因对应的 EST 数. E 示二花脸猪; L 示长白猪. **示 P<0.01, *示 P<0.05

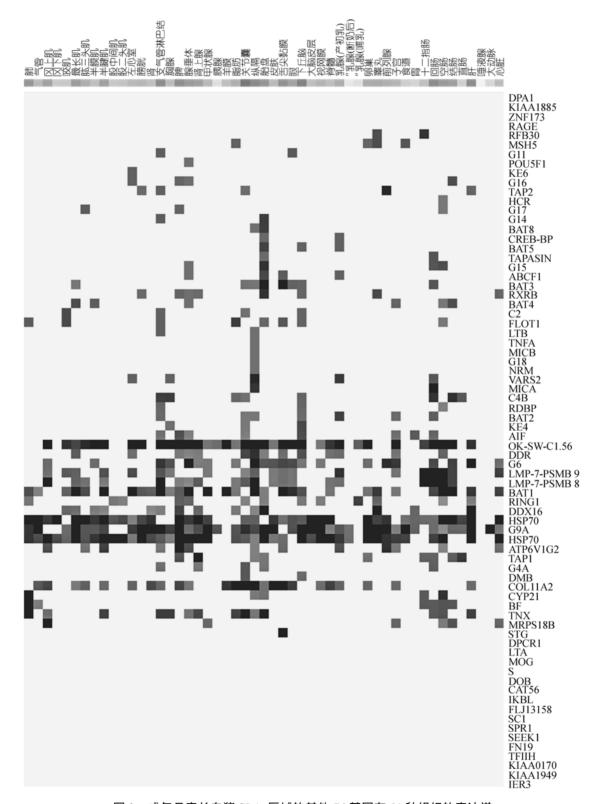


图 3 成年丹麦长白猪 SLA 区域的其他 74 基因在 51 种组织的表达谱 不同颜色代表基因表达丰度,黑色为高表达,白色为无表达,黑色成分越多,代表该基因的丰度越高,该图用加州大学伯克利分校 和伯克利国家实验室(http://rana.lbl.gov/downloads/Cluster.zip)的免费软件 Cluster 绘制

在对成年丹麦长白猪的 51 种组织的SLA表达谱 分析中发现、SLA a, b和经典的SLA 类基因有 不同的表达丰度水平和不同的组织表达模式、在此 基础上进行聚类的结果很好地体现了它们的功能分 类. 值得提出的是, SLA-3 虽然属于 a类, 但表达丰 度及组织表达模式与SLA 类更相似, HLA中也有相 似的现象. HLA a类的 3 个基因(HLA-A, HLA-B, HLA- C)中, HLA-C在 10%~50%的个体中血清学检测 为阴性^[20]、其在细胞膜表面的表达量也只有HLA-A 和-B的 $1/10^{[21]}$, B细胞中, HLA-C的mRNA水平只有 HLA-B的 1/10~1/2. McCutchoen 等人[20]利用HLA-B7 和HLA-Cw3 基因组合成的多种嵌合基因研究了造成 这种现象的原因、推测在HLA-C的外显子3与3'-UTR 之间某一段序列导致HLA-C mRNA容易被降解. SLA-3 与HLA-C的相似性不仅体现在表达水平上、还 表现在多态性位点的分布上. 与HLA-A和-B相比, HLA-C的抗原结合位点(主要在外显子 2)多态性较少 而跨膜区多态性相对更丰富[21], SLA-3 的多态性位点 显示了同样的特征. 虽然HLA和SLA a类的这 6 个 基因的系统发生树(通过Clustal W)显示、种间的差别 比基因间的差别更大、但SLA-3 和HLA-C在表达和等 位基因间多态性位点方面的一致性提示、这2个基因 可能具有某些不同于其他 2 个 a类基因的功能、在 "趋同进化"的作用下彼此相似.

对成年丹麦长白猪和中国二花脸猪的SLA Ia和 Ib及经典II 类基因的表达水平比较发现,除了空肠 外, 其他组织二花脸猪的表达水平均高于长白猪. 二 花脸猪属于江海型猪种中的太湖猪、太湖猪以高产 性能蜚声世界、二花脸猪更是太湖猪中繁殖力最高 的系. 与丹麦长白猪相比, 二花脸猪抗逆性较强, 主 要表现在抗寒、耐热性能好、耐粗饲、耐饥饿(对低营 养的耐受力强)。能适应高海拔生态环境。二花脸猪 还有一个特点是早熟早肥、肉质好、但生长速度慢、 瘦肉率低、只有 38.8%~45%[22]. 丹麦长白猪是世界 上著名瘦肉型猪种之一, 其主要特点是, 生长发育快, 省饲料, 胴体瘦肉率高(达 65%), 但抗逆性差, 对饲 料营养要求高. 这两种猪在繁殖力、抗逆性、生长速 度、饲料的回报率等生产性状上的差别的遗传基础可 能有多个,但SLA无疑是很重要的一个^[8]. SLA不但 与抗病力、免疫应答能力之间有明确的相关性¹⁹,其区 域内部和附近也有一些数量性状基因座(quantitative trait locus, QTL), 这些QTL与猪的生长速度、瘦肉率、

脂肪组成、肉质等经济性状相关^[23-26]. 虽然抗病力、免疫应答能力和经济性状之间的正、负相关关系并不是十分确定,但这种不确定的关系可能正是由于SLA区域免疫相关基因与QTL之间的连锁不平衡造成的. 为了培育生长速度快、瘦肉率高、繁殖力强、肉质好,同时又有良好抗病性的猪,SLA区域是一个值得关注的部位.

二花脸猪抗逆性强于丹麦长白猪, 同时, 前者 SLA 区域抗原呈递基因的表达高于后者. 这 2 种现 象之间是否存在一定的关联?本研究结果显示、二 花脸猪淋巴腺的经典SLA 类基因、脾的SLA a和 类基因表达水平均明显高于丹麦长白猪 经典SLA (表 7). 由于这 2 类基因主要表达在免疫细胞中、据此 可以推断, 二花脸的淋巴腺和脾中免疫细胞的数量/ 活性要高于后者. 其次, 二花脸的肾上腺SLA a和 类基因的表达均远远高于丹麦长白猪(表 7) . Marx等人[27]利用免疫组化的方法研究了人的肾 上腺MHC 类分子的分布、发现该类分子主要位于 肾上腺皮质的内层——网状带(该带细胞主要分泌雄 激素和少量糖皮质激素). MHC 类分子的表达水平 与肾上腺网状带细胞的发育相平行、即幼年起逐渐 增强、成年期最旺盛、老年后衰减、并且与肾上腺肿 瘤的恶性程度呈负相关[28], 可能参与肾上腺皮质细 胞的程序性死亡过程. 联合应用免疫组化和原位杂 交的方法、Gonzalez-Hernandez^[29,30]的研究小组证实 了肾上腺皮质表达MHC 类分子的细胞主要是巨噬 细胞、这些细胞分泌IL-6和 IL-1、这2种白细胞介素、 尤其是IL-1, 不但是免疫系统和下丘脑-垂体-肾上腺 轴之间相互作用的一个重要介质,还对肾上腺类固 醇(盐皮质激素、糖皮质激素、雄激素、雌激素)的合 成和分泌有促进作用. 肾上腺合成和分泌的多种激 素在调节免疫功能、应激、糖和蛋白质以及脂肪代谢 中有重要的作用. 因此, SLA a和经典SLA 类基因 在二花脸猪肾上腺的高表达反映了该组织中巨噬细 胞的高数量/高活性, 通过巨噬细胞影响肾上腺的功 能(如类固醇的合成量),从而影响猪的抗逆性. 当然, 不同品系猪之间抗逆性的差别可能存在多个层面上 的原因. 虽然本研究无法判断二花脸和长白猪的经 典SLA基因的表达差异是由于细胞数量的差别还是 每个细胞表达水平的差别造成的、但是、经典SLA基 因在免疫系统和肾上腺的差异表达可能是抗逆性差 异的基础之一,而这种差异表达是品系特异性的还

是等位基因特异性的, 抑或是由于上游调控因素造成的, 有待进一步研究.

丹麦长白猪的惟一抗原呈递基因表达水平普遍高于二花脸的组织是空肠. 空肠是食物消化和吸收的重要部位, 丹麦长白猪该组织中抗原呈递基因高表达, SLA a 类基因的表达尤其高. 肠道的管腔上皮很可能是高表达该类基因的主要细胞. 而丹麦长白猪空肠管腔上皮细胞高表达该类基因与该品系猪高饲料回报率之间是否存在一定的关系, 有待进一步研究.

本研究发现,不同猪种之间抗原呈递基因的表达有显著性差异,而且这种差异超过同一猪种的不同发育阶段.但是这种差异究竟是个体水平还是品系水平,有待通过系统的研究得到更多的数据来说明.

总体来说,本研究结果证明用 EST 构建组织表达谱或做"电子 Northern",对于表达水平比较高的基因来说是合适的,组织之间的表达差异可以用数字化的形式表示出来,但对低表达的基因则难以奏效,这也是 EST 方法用于表达谱分析的一个局限.

致谢 在"家猪基因组计划"中, 丹麦农业科学院提供了丹麦长白猪的各发育阶段的 cDNA 文库, 在此表示真诚的感谢. 本工作受科学技术部"863"国际合作项目——家猪基因组计划(批准号: 2002AA229061)、中国科学院知识创新重大项目, 1%人类基因组和中国重要战略资源生物基因组计划(批准号: KSCX1-01)资助.

参 考 文 献

- Geffrotin C, Popescu C P, Cribiu E P, et al. Assignment of MHC in swine to chromosome 7 by in situ hybridization and serological typing. Ann Genet, 1984, 27(4): 213~219
- Shigenari A, Ando A, Renard C. Nucleotide sequencing analysis of the swine 433-kb genomic segment located between the non-classical and classical SLA class gene clusters. Immunogenetics, 2004, 55(10): 695~705[DOI]
- 3 Renard C, Vaiman M, Chiannilkulchai N, et al. Sequence of the pig major histocompatibility region containing the classical class genes. Immunogenetics, 2001, 53(6): 490~500[DOI]
- 4 Chardon P, Rogel-Gaillard C, Cattolico L, et al. Sequence of the swine major histocompatibility complex region containing all nonclassical class genes. Tissue Antigens, 2001, 57(1): 55~65[DOI]
- Velten F W, Renard C, Rogel-Gaillard C, et al. Spatial arrangement of pig MHC class sequences. Immunogenetics, 1999, 49(11-12): 919~930[DOI]
- 6 Peelman L J, Chardon P, Vaiman M, et al. A detailed physical map of the porcine major histocompatibility complex (MHC) class region: Comparison with human and mouse MHC class gions. Mamm Genome, 1996, 7(5): 363~367[DOI]
- 7 Amadou C. Evolution of the Mhc class region: The framework hypothesis. Immunogenetics, 1999, 49(4): 362~367[DOI]
- 8 Chardon P, Renard C, Gaillard C R, et al. The porcine major histocompatibility complex and related paralogous regions: A review. Genet Sel Evol, 2000, 32(2): 109~128[DOI]

- 9 Madden K B, Murrell K D, Lunney J K, et al. *Trichinella spiralis*: Major histocompatibility complex-associated elimination of encysted muscle larvae in swine. Exp Parasitol, 1990, 70(4): 443~451[DOI]
- 10 Renard C, Vaiman M. Possible relationships between SLA and porcine reproduction. Reprod Nutr Dev, 1989, 29(5): 569~576
- 11 Chen F X, Tang J, Li N L, et al. Novel SLA class alleles of Chinese pig strains and their significance in xenotransplantation. Cell Res, 2003, 13(4): 285~294
- 12 Xu X C, Naziruddin B, Sasaki H, et al. Allele-specific and peptide-dependent recognition of swine leukocyte antigen class by human cytotoxic T-cell clones. Transplantation, 1999, 68(4): 473~479[DOI]
- 13 Kent W J. BLAT: The BLAST-like alignment tool. Genome Res, 2002, 12(4): 656~664[DOI]
- 14 Audic S, Claverie J M. The significance of digital gene expression profiles. Genome Res, 1997, 7(10): 986~995
- Stekel D J, Git Y, Falciani F. The comparison of gene expression from multiple cDNA libraries. Genome Res, 2000, 10(12): 2055~2061[DOI]
- 16 Claverie J M. Computational methods for the identification of differential and coordinated gene expression. Hum Mol Genet, 1999, 8(10): 1821~1832[DOI]
- 17 Bortoluzzi S, d'Alessi F, Danieli G A. A computational reconstruction of the adult human heart transcriptional profile. J Mol Cell Cardiol, 2000, 32(11): 1931~1938[DOI]
- Oliver A M, Thomson A W, Sewell H F, et al. Major histocompatibility complex (MHC) class antigen (HLA-DR, DQ, and DP) expression in human fetal endocrine organs and gut. Scand J Immunol, 1988, 27(6): 731~737
- 19 Metzger R, Mempel T, Joppich I, et al. Organ-specific distribution of major histocompatibility antigens in rats. Pediatr Surg Int, 2000, 16(4): 285~292[DOI]
- 20 McCutcheon J A, Gumperz J, Smith K D, et al. Low HLA-C expression at cell surfaces correlates with increased turnover of heavy chain mRNA. J Exp Med, 1995, 181(6): 2085~2095[DOI]
- 21 Zemmour J, Parham P. Distinctive polymorphism at the HLA-C locus: Implications for the expression of HLA-C. J Exp Med, 1992, 176(4): 937~950[DOI]
- 22 中国家畜品种志编委会, 编. 中国家畜品种志. 上海: 上海科学技术出版社, 1980.
- 23 Bidanel J P, Milan D, Iannuccelli N, et al. Detection of quantitative trait loci for growth and fatness in pigs. Genet Sel Evol, 2001, 33(3): 289~309[DOI]
- 24 Demeure O, Renard C, Yerle M, et al. Rearranged gene order between pig and human in a QTL region on SSC 7. Mamm Genome, 2003, 14(1): 71~80[DOI]
- 25 Genet C, Renard C, Cabau C, et al. In the QTL region surrounding porcine MHC, gene order is conserved with human genome. Mamm Genome, 2001, 12(3): 246~249[DOI]
- 26 Milan D, Bidanel J P, Iannuccelli N, et al. Detection of quantitative trait loci for carcass composition traits in pigs. Genet Sel Evol, 2002, 34(6): 705~728[DOI]
- 27 Marx C, Bornstein S R, Wolkersdorfer G W, et al. Relevance of major histocompatibility complex class expression as a hall-mark for the cellular differentiation in the human adrenal cortex. J Clin Endocrinol Metab, 1997, 82(9): 3136~3140[DOI]
- Wolkersdorfer G W, Marx C, Brown J W, et al. Evaluation of apoptotic parameters in normal and neoplastic human adrenal. Endocr Res, 1996, 22(4): 411~419
- 29 Gonzalez-Hernandez J A, Bornstein S R, Ehrhart-Bornstein M, et al. Interleukin-6 messenger ribonucleic acid expression in human adrenal gland in vivo: New clue to a paracrine or autocrine regulation of adrenal function. J Clin Endocrinol Metab, 1994, 79(5): 1492~1497[DOI]
- 30 Gonzalez-Hernandez J A, Bornstein S R, Ehrhart-Bornstein M, et al. IL-1 is expressed in human adrenal gland *in vivo*. Possible role in a local immune-adrenal axis. Clin Exp Immunol, 1995, 99(1): 137~141

(2004-10-26 收稿, 2005-03-17 收修改稿)