

Research article

Open Access

A molecular signature of epithelial host defense: comparative gene expression analysis of cultured bronchial epithelial cells and keratinocytes

Joost B Vos*¹, Nicole A Datson⁴, Antoine H van Kampen³, Angela C Luyf³, Renate M Verhoosel¹, Patrick L Zeeuwen², Diana Olthuis², Klaus F Rabe¹, Joost Schalkwijk² and Pieter S Hiemstra¹

Address: ¹Dept. of Pulmonology, Leiden University Medical Center, Leiden, Albinusdreef 2, 2333 ZA Leiden, The Netherlands, ²Dept. of Dermatology, Nijmegen Center for Molecular Life Sciences, Radboud University Nijmegen Medical Center, PO Box 9101, 6500 HB Nijmegen, The Netherlands, ³Bioinformatics Laboratory, Academic Medical Center, PO Box 22700, 1100 DE, Amsterdam, The Netherlands and ⁴Dept. of Medical Pharmacology, Leiden Amsterdam Center for Drug Research, Leiden University Medical Center, Leiden, Einsteinweg 55, 2333 CC Leiden, The Netherlands

Email: Joost B Vos* - j.b.vos@lumc.nl; Nicole A Datson - n.datson@chem.leidenuniv.nl; Antoine H van Kampen - a.h.vankampen@amc.uva.nl; Angela C Luyf - a.c.luyf@amc.uva.nl; Renate M Verhoosel - r.m.verhoosel@lumc.nl; Patrick L Zeeuwen - p.zeeuwen@derma.umcn.nl; Diana Olthuis - d.olthuis@derma.umcn.nl; Klaus F Rabe - k.f.rabe@lumc.nl; Joost Schalkwijk - j.schalkwijk@derma.umcn.nl; Pieter S Hiemstra - p.s.hiemstra@lumc.nl

* Corresponding author

Published: 18 January 2006

Received: 11 April 2005

BMC Genomics 2006, 7:9 doi:10.1186/1471-2164-7-9

Accepted: 18 January 2006

This article is available from: <http://www.biomedcentral.com/1471-2164/7/9>

© 2006 Vos et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Epithelia are barrier-forming tissues that protect the organism against external noxious stimuli. Despite the similarity in function of epithelia, only few common protective mechanisms that are employed by these tissues have been systematically studied. Comparative analysis of genome-wide expression profiles generated by means of Serial Analysis of Gene Expression (SAGE) is a powerful approach to yield further insight into epithelial host defense mechanisms. We performed an extensive comparative analysis of previously published SAGE data sets of two types of epithelial cells, namely bronchial epithelial cells and keratinocytes, in which the response to pro-inflammatory cytokines was assessed. These data sets were used to elucidate a common denominator in epithelial host defense.

Results: Bronchial epithelial cells and keratinocytes were found to have a high degree of overlap in gene expression. Using an *in silico* approach, an epithelial-specific molecular signature of gene expression was identified in bronchial epithelial cells and keratinocytes comprising of family members of keratins, small proline-rich proteins and proteinase inhibitors. Whereas some of the identified genes were known to be involved in inflammation, the majority of the signature represented genes that were previously not associated with host defense. Using polymerase chain reaction, presence of expression of selected tissue-specific genes was validated.

Conclusion: Our comparative analysis of gene transcription reveals that bronchial epithelial cells and keratinocytes both express a subset of genes that is likely to be essential in epithelial barrier formation in these cell types. The expression of these genes is specific for bronchial epithelial cells and keratinocytes and is not seen in non-epithelial cells. We show that bronchial epithelial cells, similar to keratinocytes, express components that are able to form a cross-linked protein envelope that may contribute to an effective barrier against noxious stimuli and pathogens.

Background

Epithelial tissues in the mammalian airways and skin are among the largest organs and form the interface between the internal milieu of the host and the outside world. They not only protect the host against invading pathogens but also provide an effective barrier to noxious external (chemical and physical) stimuli and dehydration [1,2]. The effectiveness of the epithelial barrier is demonstrated by the rare incidence of severe infections to the lung or skin in healthy individuals. It has become clear that epithelia also play an active role in innate and adaptive immunity [3,4]. Epithelial tissues display three main mechanisms to protect the organism from infection. First, epithelial cells form an impermeable physical barrier which both prevents pathogen entry and minimizes dehydration. Second, epithelial cells are capable of producing defense molecules such as antimicrobial peptides and proteinase inhibitors. Finally, these cells are able to produce signaling molecules such as cytokines and chemokines. These molecules may attract or activate cells of the innate and adaptive immune system [5,6]. Interaction between cells of the immune system is mediated by adhesion molecules and cytokine receptors [7,8] that are present on epithelial cells.

Host defense mechanisms in epithelial cells are coordinated by a complex program of gene expression. Very powerful and sophisticated laboratory techniques such as Serial Analysis of Gene Expression (SAGE) [9] and DNA microarrays [10] have been developed to assess the expression of thousands of genes at the mRNA level in a single experiment. To delineate the barrier function of epithelial cells, the transcriptional change induced by pro-inflammatory cytokines was recently assessed by means of SAGE in two well-established culture models of epithelial inflammation using subcultures of primary bronchial epithelial cells [11] and primary keratinocytes [12]. These independent studies showed a marked overlap in gene families expressed in response to pro-inflammatory cytokines in both cell types. Upon cytokine exposure, in particular genes associated with cytoskeletal architecture and epidermal barrier function such as keratins, S100 calcium-binding proteins and various antimicrobial proteinase inhibitors were differentially expressed. These studies indicated that bronchial epithelial cells and keratinocytes might respond similarly to external influences to ultimately provide effective host protection. This is especially of interest because the epithelia of the skin and conducting airways are markedly different in morphology. The potential functional resemblance of these types of epithelia is also demonstrated by comparative analysis of genetic studies in patients with asthma and atopic dermatitis showing that similar patterns of gene expression may contribute to susceptibility to these diseases [13]. This prompted us to conduct a comparative analysis of our

previously generated gene expression in culture models of epithelial inflammation. The aim was to test the hypothesis whether bronchial epithelial cells and keratinocytes employ similar mechanisms for providing effective host defense at these epithelia.

Therefore, in the present study, our previously generated SAGE data sets derived from bronchial epithelial cells [11] and keratinocytes [12] that were exposed to pro-inflammatory cytokines were compared to identify a common denominator in host defense in the different types of epithelial cells. SAGE libraries of resting and IL1 β /TNF α -exposed primary bronchial epithelial cells (~28.000 tags in each library) were compared to SAGE libraries of resting and TNF α -exposed human primary keratinocytes (~13.000 tags in each library). The *in silico* method Tissue Preferential Expression (TPE) [14] was used for the recognition of putative cell-specific gene expression in these SAGE libraries. Previously, this method has been successfully applied to identify novel specific markers for disease [14,15]. To verify the *in silico* prediction analysis of tissue specific gene expression, polymerase chain reaction was performed on seven target genes that were identified by the TPE algorithm in a panel of nine different cell types of which seven are normally present in the airways or lungs. The airway- and lung-derived NCI-H292 and A549 cell lines were included since these cell lines are frequently used to study epithelial cell function. We have identified and validated a signature of specific gene expression for bronchial epithelial cells and keratinocytes. The majority of genes in this signature was previously not associated with host defense or inflammation. These results indicate that epithelia of the airways and skin exploit unified host defense strategies to protect the host, despite their morphological differences.

Results

Transcriptional overlap between PBEC and KC and epithelial-specific gene expression upon cytokine exposure was characterized. By comparing the four SAGE libraries of primary bronchial epithelial cells (PBEC) and keratinocytes (KC), an overlap in tags of approximately 80% was observed indicating a high similarity in the repertoire of genes expressed by these types of epithelial cells. Although remarkable commonalities were found in gene families found to be expressed by PBEC and KC, the repertoire of transcribed family members differed among the two cell types (table 2). To extract a pattern of genes that is specifically expressed in epithelial cells that could likely be involved in epithelial host defense we explored which of the genes are preferentially expressed by PBEC and KC using the TPE algorithm. The scatter plot in figure 1 displays the individual tags observed in the cytokine-exposed PBEC and KC libraries. Each dot represents a single tag with the corresponding TPE values for PBEC and KC. In

Table 1: Primer sequences and conditions for RT-PCR. Gene sequences used for primer design were retrieved from the Ensembl website [36], annealing temperature and MgCl₂ concentration used in PCR reactions are listed.

Target	Ensembl ID	Sense	Antisense	Annealing T	MgCl ₂ (mM)
KRT6A	ENSG00000074729	5'-CTG AGG CTG AGT CCT GGT AC-3'	5'-GTT CTT GGC ATC CTT GAG G-3'	56	2.5
SPRR1A	ENSG00000169474	5'-ACA CAG CCC ATT CTG CTC CG-3'	5'-TGC AAA GGA GCG ATT ATG ATT-3'	52	2
SPRR1B	ENSG00000169469	5'-AGA CCA AGC AGA AGT AAT GTG-3'	5'-AGA CCT TCA GCT TCA TTC AGA G-3'	61	4
SPRR2A	ENSG00000163212	5'-TGG TAC CTG AGC ACT GAT CTG CC-3'	5'-CCA AAT ATC CTT ATC CTT TCT TGG-3'	58	2
IL1F9	ENSG00000136688	5'-TGG GAA TCC AGA ATC CAG-3'	5'-TTG GCA CGG TAG AAA AGG-3'	61	3.5
S100A2	ENSG00000160675	5'-CAA GAG GGC GAC AAG TTC-3'	5'-GCC CAT CAG CTT CTT CAG-3'	59	3
CALML5	ENSG00000178372	5'-GGT TGA CAC GGA TGG AAA CG-3'	5'-AAC CTC GGA GAT GAG TTT CCT TAG-3'	60	3
ACTB	ENSG00000075624	5'-AAG GAA GGC TGG AAG AGT GC -3'	5'-CTA CAA TGA GCT GCG TGT GG -3'	56	2

this analysis, four groups of tags were identified: epithelial non-specific tags (*i*), tags preferentially expressed by either PBEC (*ii*) or KC (*iii*) and tags that were preferentially expressed by both PBEC and KC (*iv*). The expression of the 30 tags observed in the latter group represents putative epithelial-specific genes because a TPE score ≥ 9 was observed in both PBEC and KC (table 3). Almost half of these tags corresponded to genes encoding for keratins, small proline-rich proteins, kallikreins and proteinase inhibitors (table 3). Interestingly, the expression of a large proportion of these genes was found to be affected by cytokine exposure in PBEC or KC (or both) as observed in the initial SAGE studies (as indicated by underlined tag numbers in table 3). A similar picture in preferential tag expression was obtained when using the libraries of resting PBEC and KC since the majority of genes do not show an on/off expression profile upon stimulation with cytokines (data not shown).

To validate this *in silico* TPE prediction analysis, expression of seven putative epithelial-specific genes by PBEC and KC was assessed by reverse transcriptase polymerase chain reaction (RT-PCR) in nine different cell types. Each cell type in the panel was exposed to medium alone or to IL1 β /TNF α . KC were exposed to medium or TNF α alone to maintain comparability with the original SAGE experiment. In concordance with the SAGE data and TPE analysis, expression for SPRR2A was only observed in PBEC. On the other hand, CALML5 was expected to be expressed by KC alone. However, PBEC were shown to be positive for this transcript as well and weak expression was observed in NCI-H292 cells. As demonstrated by the TPE analysis KRT6A, SPRR1A, SPRR1B, IL1F9, S100A2 all showed TPE values of ≥ 9 in both PBEC and KC libraries. The RT-PCR results in figure 2 demonstrates that preferential expression of SPRR1B was found in PBEC, KC and NCI-H292 cells, whereas moderate to weak expression was also detected in fibroblasts, HUVEC, HASM and monocytes. Expression of KRT6A is restricted to PBEC, KC and the bronchial epithelial cell line NCI-H292, whereas expression of this transcript was negative in all other cell types. Transcription of SPRR1A, IL1F9 and S100A2 was only

detected in primary cultures of PBEC and KC and was completely absent in all other cell types.

Discussion

Comparative genomics approaches have the potential to gain additional insight into a biological process at the mRNA expression level by integrating and combining data obtained from similar model systems. Particularly, SAGE is excellent for this purpose since digital, scalable expression data is generated that allows comparison without the need for complex mathematical normalization methods. Although the SAGE libraries used in the present analysis were not initially intended for comparative genomic research, remarkable commonalities in epithelial-specific gene expression were found that related to host defense.

The tissue preferential expression (TPE) algorithm was employed to recognize specific tag expression by PBEC and KC under inflammatory conditions (pane iv; figure 1, table 3). Experimental verification of selected epithelial-specific genes by RT-PCR showed a good correlation between the *in silico* approach and RT-PCR (figure 2). The PCR setup was designed to detect true presence or absence of validation genes and was not intended to be quantitative. The observed discrepancies between SAGE and PCR results can be explained by the difference in detection sensitivity between techniques: RT-PCR is far more sensitive than SAGE in detecting low abundant gene expression.

The majority of tags of the molecular signature corresponded to genes encoding structural components of the cytoskeleton (keratins, small proline-rich proteins, elafin) and for proteins that are involved in the assembly/disassembly (transglutaminase 1, kallikreins and matrix metalloproteinases) of the cornified cell envelope in keratinocytes (reviewed in [16]). Components of the cross-linked or cornified envelope are linked by transglutaminases (reviewed in [17]). The observation that bronchial epithelial cells express components of and assembly/disassembly enzymes forming a cross-linked envelope is relevant to our understanding of epithelial

host defense in the airways. Additional support for this observation is provided by abundant transcription of genes that are known to be involved in cornification in skin, including the S100 calcium-binding proteins [18], annexins [18] and cystatins [19,20] (table 2). So far, only few studies provide evidence for the existence of a protein envelope in bronchial epithelial cells. Components such as small proline-rich proteins (SPRR) have been suggested to be associated with squamous differentiation [21,22]. Low SPRR expression has been associated with squamous cell carcinoma [23], whereas high expression of SPRR1B enhanced G0-arrest resulting in growth arrest [24]. Interestingly, the families of small proline-rich proteins and S100 calcium-binding proteins are encoded in the epidermal differentiation complex (EDC) [25,26]. Proteins encoded in this region share significant sequence similarities, particularly in the glutamine- and lysine-rich regions that are involved in the cross-linking by transglutaminases [17]. This indicates that PBEC and KC not only share structural characteristics, but may also share functional characteristics.

A disadvantage of the present study might be the differences in type of cytokine-exposure and duration of the treatment. The opposite directional changes in expression in gene families (table 2) observed could be explained either by dissimilarities in the initial model systems or by the inherent differences between PBEC and KC. By using the TPE algorithm, highly cell-specific tag expression can be predicted largely independently from transcriptional levels because the more unique a tag is to a particular tissue, the less important is its level of expression. Therefore, we are confident that the signature of epithelial host defense that was extracted is representative for bronchial epithelial cells and keratinocytes.

Computational subtraction methods such as the TPE algorithm allow functional clustering of genes derived from large and complex genome-wide expression profiles without having full knowledge of the repertoire of genes involved in biological processes of interest. Although the identified molecular signature of host defense is characteristic for bronchial epithelial cells and keratinocytes, it would be of great interest to study whether this gene expression pattern is also applicable to other types of epithelial cells, a finding that would greatly enhance our understanding of epithelial defense strategies.

Conclusion

In summary, our comprehensive comparison of overlapping genes across bronchial epithelial cells and keratinocytes provides novel insights in epithelial host defense strategies, in particular of the airway epithelium. Combining *in silico* and experimental approaches is very valuable in accelerating the interpretation of genomics data and

defining follow-up research. We identified an expression signature of genes that were specifically expressed by bronchial epithelial cells and keratinocytes. These genes are likely to fulfill an eminent function in epithelial host defense. Based on the present findings we propose that formation of a cross-linked protein envelope by bronchial epithelial cells is an effective host defense strategy of the mucosal epithelium in the human airways. This function would be analogous to the host defense function of cornifying keratinocytes. Finally, a better understanding of unified host defense strategies in different epithelia may lead to the identification of novel therapeutic targets for epithelial inflammatory disorders such as asthma and atopic dermatitis.

Methods

SAGE data

The previously published SAGE libraries that were compared in this study were derived from two models of epithelial inflammation using primary bronchial epithelial cells [11] and primary keratinocytes [12]. The SAGE data from the original studies is accessible through NCBI's Gene Expression Omnibus [27] with GEO accessions GSM37337 (PBEC_unstimulated), GSM37339 (PBEC_IL1beta/TNFalpha), GSM1121 (NormCultKC_Diff) and GSM1122 (TNF_AlphaCultKC).

For tag mapping, after discarding tags occurring only once, the libraries were compared with NCBI's "reliable Unigene cluster to SAGE tag map" [28] and with SAGEgenie of the Cancer Genome Anatomy Project [29]. Both maps were based on Unigene build#171. Additionally, to enhance the reliability of tag identity we included the virtual tag classification as used in SAGEgenie to assess the location of each tag within the corresponding transcript. Reliable tags can be discriminated from tags that are not isolated from the 3'-end such as internally primed transcripts and tags derived from internal NlaIII restriction sites [29,30]. In the data set comparisons, only tags were included that were derived from the most 3'-restriction site of NlaIII, tags that matched to undefined 3'-end transcripts and tags for which no additional information was available. The last category may contain tags that correspond to novel transcripts.

TPE analysis

Epithelial-specific gene expression in PBEC and KC was identified using the Tissue Preferential Expression (TPE) algorithm [14]. The calculated Tissue Preferential Expression (TPE) value is based both on the presence of a particular tag and its level of expression in the SAGE library of interest in comparison to a panel of reference SAGE libraries derived from a range of different whole tissues. To allow calculation of TPE values, each of the PBEC and KC SAGE libraries as well as the reference libraries were nor-

Table 2: Inventory of gene families and their members expressed by PBEC and/or KC.

SAGE Tag Sequence	SAGE Tag count#				Symbol	Description	Tissue Preferential Expression values			
	PBEC CTRL	PBEC IL1 β /TNF α	KC CTRL	KC TNF α			PBEC CTRL	PBEC IL1 β /TNF α	KC CTRL	KC TNF α
Keratins										
ACATTTCAAA	0	0	145	127	KRT1	keratin 1	*	*	13,4	12,9
GCCCCTGCTG	96	180	108	115	KRT5	keratin 5	10,2	10,9	11,4	11,6
AAAGCACAAG	267	252	136	235	KRT6A	keratin 6A	12,0	11,9	12,3	12,6
CGAATGCCT	21	60	84	97	KRT6B	keratin 6B	9,7	10,7	12,2	12,3
GATGTGCACG	12	26	419	514	KRT14	keratin 14	8,6	9,2	13,5	13,5
CAGCTGTCCC	21	25	26	29	KRT16	keratin 16	9,9	10,0	11,1	11,3
CTTCCTTGCC	181	216	431	380	KRT17	keratin 17	9,8	10,0	12,7	12,4
GACATCAAGT	20	48	0	0	KRT19	keratin 19	5,6	6,5	*	*
Small proline-rich proteins										
CTGTCACCT	9	25	126	75	SPRR1A	small proline-rich protein 1A	10,3	11,1	13,8	12,9
CCCTTGAGGA	27	56	264	176	SPRR1B	small proline-rich protein 1B	10,1	10,8	13,5	12,8
ATGATCCCTG	3	12	2	0	SPRR2A	small proline-rich protein 2A	9,1	9,9	10,4	*
TTCCTGCTC	91	75	0	2	SPRR3	small proline-rich protein 3	9,1	8,9	*	6,3
Calcium binding proteins										
GATCTCTTGG	123	170	103	113	S100A2	S100 calcium binding protein A2	8,5	9,0	9,0	9,0
CCCCCTGGAT	61	133	82	32	S100A6	S100 calcium binding protein A6	<4	4,1	<4	<4
TACCTGCAGA	139	214	28	77	S100A8	S100 calcium binding protein A8	5,8	6,4	4,2	5,5
GTGGCCACGG	121	419	77	156	S100A9	S100 calcium binding protein A9	5,6	7,3	6,2	7,2

Table 2: Inventory of gene families and their members expressed by PBEC and/or KC. (Continued)

AGCAGATCAG	114	86	35	70	SI00A10	SI00 calcium binding protein A10	<4	<4	<4	<4
CAGGCCCCAC	15	29	9	11	SI00A11	SI00 calcium binding protein A11	<4	<4	<4	<4
TGGGGAGAGG	53	51	30	32	SI00A14	SI00 calcium binding protein A14	7,7	7,7	8,7	8,7
AGCAGGAGCA	26	54	40	18	SI00A16	SI00 calcium binding protein A16	4,3	5,3	5,2	4,2
ATCCGCGAGG	0	0	30	29	CALML5	calmodulin-like 5	*	*	12,0	11,6
Annexins										
AGAAAGATGT	102	62	7	7	ANXA1	annexin A1	4,5	<4	<4	<4
CTTCCAGCTA	16	43	77	109	ANXA2	annexin A2	<4	<4	4,5	5,0
AAGGGCGCGG	4	8	12	0	ANXA3	annexin A3	<4	4,2	4,7	*
TTGTTATTGC	5	0	2	5	ANXA7	annexin A7	<4	*	<4	<4
CCCTCAGCAC	3	2	0	14	ANXA8	annexin A8	7,9	7,6	*	9,9
Proteinase inhibitors										
ATCCTTGCTG	94	85	150	91	CSTA	cystatin A	8,7	8,6	10,6	9,9
ATGAGCTGAC	64	135	37	38	CSTB	cystatin B	5,0	6,1	4,1	4,1
GTGGAGGGCA	2	2	12	27	CST6	cystatin E/M	4,7	4,7	6,6	7,3
CATTGTAAAT	16	12	26	14	SERPINB5	serine proteinase inhibitor, member 5	9,6	9,4	11,3	10,4
TTGAATCCCC	30	80	56	50	PI3	elafin, protease inhibitor 3, (SKALP)	8,0	9,0	9,9	9,6
TGTGGGAAAT	18	40	101	124	SLPI	secretory leukocyte protease inhibitor	5,3	6,3	7,7	8,0

Expression of these gene families was observed in both PBEC and KC, whereas these cells differ in the expression pattern of the individual family members. From left to right the SAGE tag sequence, SAGE tag counts as observed in the SAGE libraries (unstimulated PBEC, IL1 β /TNF α -stimulated PBEC, unstimulated KC and TNF α -stimulated KC), HUGO approved gene symbol and gene description and the calculated Tissue Preferential Expression Values (unstimulated PBEC, IL1 β /TNF α -stimulated PBEC, unstimulated KC and TNF α -stimulated KC) are indicated. No TPE values could be calculated for those tags that were absent in one or more libraries and are indicated by (*) in the table. #A selection of these SAGE tag counts were represented in previously published tables [11,12]; all tag counts are available online through the NCBI Gene Expression Omnibus (GEO) website [37] with GEO accessions as listed in the methods.

Table 3: 30 Epithelial-specific genes as identified by the TPE analysis.

SAGE Tag Sequence	SAGE Tag count [#]				Symbol	Description	Barrier formation
	PBEC CTRL	PBEC ILI β /TNF α	KC CTRL	KC TNF α			
AAAGCACAAG	267	252	<u>136</u>	<u>235</u>	KRT6A	keratin 6A	yes
CTTCCTTGCC	181	216	431	380	KRT17	keratin 17	yes
GCCCCGTGCTG	<u>96</u>	<u>180</u>	108	115	KRT5	keratin 5	yes
TAAACCTGCT	<u>40</u>	<u>68</u>	<u>758</u>	<u>439</u>	LGALS7	lectin, galactoside-binding, soluble, 7 (galectin 7)	
TTGAATCCCC	<u>30</u>	<u>80</u>	56	50	PI3	protease inhibitor 3, skin-derived (SKALP), Elafin	yes
CCCTTGAGGA	<u>27</u>	<u>56</u>	<u>264</u>	<u>176</u>	SPRR1B	small proline-rich protein 1B (cornifin)	yes
CGAATGTCCT	<u>21</u>	<u>60</u>	84	97	KRT6B	keratin 6B	yes
CAGCTGTCCC	21	25	26	29	KRT16	keratin 16	yes
AGCTTCTACC	17	22	<u>14</u>	<u>25</u>	HCG9	HLA complex group 9	
CATTGTAAAT	16	12	<u>26</u>	<u>14</u>	SERPINB5	serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 5	
GATGTGCACG	<u>12</u>	<u>26</u>	419	514	KRT14	keratin 14	yes
CTGTCACCCCT	<u>9</u>	<u>25</u>	<u>126</u>	<u>75</u>	SPRR1A	Small proline-rich protein 1A	yes
CCCTGTTGAT	8	8	<u>7</u>	<u>18</u>	KLK7	kallikrein 7 (chymotryptic, stratum corneum)	
GGCTTCTAAC	<u>4</u>	<u>16</u>	35	38	SPRR2B	small proline-rich protein 2B	yes
GAAGCACAAG	<u>4</u>	<u>13</u>	<u>9</u>	<u>18</u>		Transcribed sequences	
CCAGCGCCAA	<u>3</u>	<u>16</u>	16	14	C4.4A	GPI-anchored metastasis-associated protein homolog	
GCTTCCTCGG	<u>2</u>	<u>11</u>	2	5	RHCG	Rhesus blood group, C glycoprotein	
TCTCTGGGG	2	4	0	2	FLJ11036	hypothetical protein FLJ11036	
AAAGCACAAT	1	1	0	5	TRA1	tumor rejection antigen (gp96)	
TCCTGGATCA	1	1	0	2	KLK10	kallikrein 10	
ATCCCTTGCT	1	2	<u>9</u>	<u>2</u>		Transcribed sequences	
AGAGCACAAG	1	1	2	5		Transcribed sequences	
CTTGCCCTGC	0	2	0	2	ZDHHC9	zinc finger, DHHC domain containing 9	
ACCTCCACTG	0	2	<u>54</u>	<u>34</u>	UNQ467	KIPV467	
CTGCTCAATG	0	3	9	11	TGMI	transglutaminase I	yes
TTCCCTTACC	0	2	<u>5</u>	<u>16</u>	SPRL6A	small proline rich-like 6A	yes
ACCTGGAGGG	<u>0</u>	<u>12</u>	<u>28</u>	23	PCBP1	poly(rC) binding protein 1	
GGGCCACGGC	0	1	7	<u>14</u>	MMP11	matrix metalloproteinase 11 (stromelysin 3)	
ACTAGCACAG	<u>0</u>	<u>7</u>	0	2	ILIF9	interleukin 1 family, member 9	
TAGACCTGCT	0	3	2	2		CDNA FLJ32217 fis, clone PLACE6003771	

Putative epithelial-specific tags as presented in pane iv of the scatter plot in figure 1 as identified by the TPE analysis are listed in table 3. From left to right are indicated the SAGE tag sequence, the normalized SAGE tag counts (unstimulated PBEC, ILI β /TNF α -stimulated PBEC, unstimulated KC and TNF α -stimulated KC), HUGO approved gene symbol and gene description and the putative involvement in epithelial barrier formation. Underlined tag counts indicate statistical significant difference in expression in the original SAGE data sets of PBEC (unstimulated vs. ILI β /TNF α) or KC (unstimulated vs. TNF α). # A selection of these SAGE tag counts were represented in previously published tables [11,12]; all tag counts are available online through the NCBI Gene Expression Omnibus website [37] with GEO accessions as listed in the methods.

malized to a equal number of tags per library. TPE values were determined for each tag in the individual PBEC and KC libraries by selecting each of the libraries as library of interest before applying the TPE algorithm. After calculation, the TPE values were ranked according to their value. Large positive TPE values represent tissue-specific genes that are overexpressed in the PBEC and/or KC libraries.

Tags with TPE values of <4 were excluded from further analysis since these tags occur very frequently in other cell types as well (See supplement for a detailed description of the TPE algorithm). The threshold value for the TPE analysis that is indicative for tissue-specific expression was chosen very high to prevent possible false positives. Tags with a corresponding TPE value of ≥ 9 is indicative for tis-

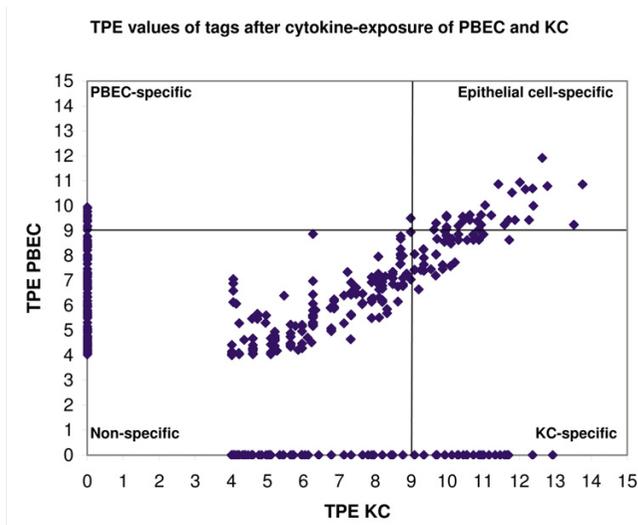


Figure 1
TPE scatter plot of SAGE tags of PBEC and KC libraries after cytokine exposure. Reliable 3'-end tags with TPE>4 and tag frequency of = 2 in at least one library were plotted. Tags with corresponding TPE values = 9 in both libraries were considered to be potential epithelial cell-specific tags as indicated by the threshold lines in the figure. A similar picture was obtained when TPE values of tags from the resting libraries were plotted.

sue preferential expression in the epithelial cells used in this analysis. See additional file for a detailed description of the TPE algorithm.

Cell culture

Primary Bronchial Epithelial Cells (PBEC) and Primary Keratinocytes (KC)

Subcultures of human primary bronchial epithelial cells and human primary keratinocytes were derived and cultured as described previously [31,32].

A549 and NCI-H292 cells

the lung derived epithelial cell lines A549 (CCL-185, American Type Culture Collection, Rockville, MD, USA) and NCI-H292 (CRL-1848, American Type Culture Collection) were cultured according to the supplier's recommendation. Prior to stimulation, cells were cultured overnight in serum free medium.

Human airway smooth muscle cells (HASM)

Human airway smooth muscle cells (HASM) from two donors were purchased from Stratagene (La Jolla, CA) and were cultured as described previously [33].

Human mast cells (HMC-1)

HMC-1 were kindly provided by J.H. Butterfield [34] and were cultured in IMDM medium containing 25 mM

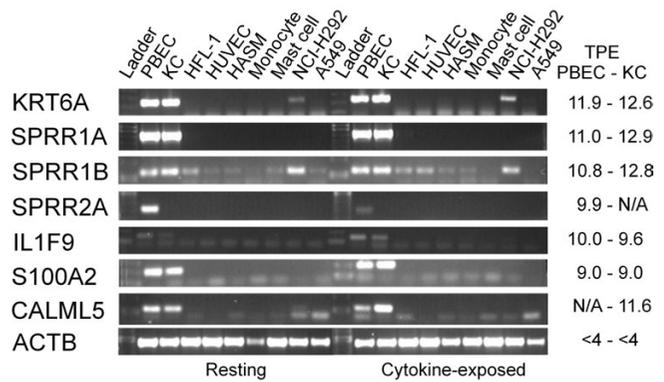


Figure 2
PCR verification of 7 potential preferentially expressed tags identified using the TPE algorithm. The expression of these seven genes was assessed both under resting conditions and after cytokine exposure in all cell types. On the right, the TPE values of the gene are listed for both PBEC and KC after cytokine exposure. Tags for which no TPE value could be calculated because of absence of the tag in the particular library are indicated by "not available" (N/A). The predicted preferential expression could be verified for all genes. The expression of six of these genes seems to be selective for epithelial cells only. Whereas SPRR1B is preferentially expressed by epithelial cells, moderate to low levels of expression were also detected in other cell types as well. SPRR2A is preferentially expressed by PBEC only. No tags for SPRR2A were found in the KC libraries after TNF α exposure whereas expression for CALML5 was observed by RT-PCR in both PBEC and KC, while no tags for this gene were found in PBEC libraries.

Hepes, 2 mM L-glutamine, 20 U/ml penicillin, 20 μ g/ml streptomycin (all from Bio Whittaker, Walkersville, MD), 5 μ g/ml apo-transferrin, 0,36% β -mercaptoethanol and 10% heat-inactivated Fetal Calf Serum (FCS; GibcoBRL/Life Technologies, Breda, The Netherlands). Prior to stimulation, cells were cultured overnight in serum free medium (same as above, without heat-inactivated FCS)

Human lung fibroblasts (HFL-1)

HFL-1 (CCL-153, American Type Culture Collection) were cultured according to the supplier's recommendations. Prior to stimulation, cells were cultured overnight in serum free medium.

Monocytes

CD14-purified monocytes were kindly provided by the department of Nephrology (Leiden University Medical Center, Leiden, The Netherlands), and were resuspended in RPMI 1640 medium supplemented with 20 U/ml penicillin, 20 μ g/ml streptomycin, 2 mM glutamine (all from Bio Whittaker) and 10% heat-inactivated FCS and were

seeded in 6-wells plates to allow adherence. After 2 hours, the medium was replaced by serum free medium (same as above, without heat-inactivated FCS) for overnight incubation prior to stimulation.

Human umbilical vein endothelial cells (HUVEC)

HUVEC were kindly provided by the department of Nephrology, Leiden University Medical Center, Leiden, The Netherlands. All cell cultures were performed at 37°C, 5% CO₂ and 95% relative humidity.

Stimulation of cell cultures

All cell types, except for the keratinocytes, were stimulated for 6 hours with either medium alone, a mixture of the pro-inflammatory cytokines IL1 β (20 ng/ml; PeproTech, Rocky Hill, NJ) and TNF α (20 ng/ml; PeproTech). Keratinocytes were stimulated with TNF α (25 ng/ml) for 48 hours as described previously [12].

Reverse Transcription PCR

RT-PCR was used to verify preferential expression of genes identified in the TPE analysis. Total RNA from the nine different cell cultures was extracted using the RNeasy mini kit (Qiagen, Westburg, Leusden, The Netherlands) and on-column DNA digestion was performed with DNase I (Qiagen, Westburg), all according to the manufacturer's instructions. Single-stranded cDNA was synthesized of total RNA using M-MLV Reverse Transcriptase primed with Oligo-dT (both from Invitrogen/Life Technologies, Breda, The Netherlands) in the presence of a RNase inhibitor (RNaseOUT; Invitrogen/Life Technologies) according to the manufacturer's instructions. Gene-specific primers were designed for keratin 6A (KRT6A), small proline-rich protein and 1B (SPRR1B), the calcium-binding protein S100A2, IL1 family member 9 (IL1F9), calmodulin-like 5 (CALML5) and β -actin (ACTB) as internal control (table 1). Primers for the small proline-rich protein 1A and 2A [35] were kindly provided by Claude Backendorf. Primers were synthesized by Isogen (Maarsse, The Netherlands). All PCR reactions were carried out according to the following PCR conditions: initial denaturation of 3 minutes at 95°C, then 35 cycles of 15 seconds denaturing at 95°C, 15 seconds primer annealing, 30 seconds elongation at 72°C, and a final extension of 3 minutes at 72°C in the last cycle. For β -actin, the cycle number was limited to 25. PCR products were visualized on 1% agarose gels using ethidium bromide. Both the reverse transcription and PCR reactions were performed on a Biometra T-Gradient thermocycler (Biometra GmbH, Goettingen, Germany).

Authors' contributions

JBV was primarily responsible for study design and coordination, building comparative SAGE expression databases, performing computational analyses and drafting the manuscript. NAD and PSH participated in study

design, coordination and manuscript editing. AHvK and ACL advised in study design and performed the TPE computational analyses. DO, JS and PLZ provided the keratinocyte SAGE data, supplied samples and reagents and assisted in editing the manuscript. RMV performed cell culture experiments, PCR analyses and provided comments and discussion. KFR provided useful discussion and manuscript editing. All authors read and approved the final manuscript.

Additional material

Additional File 1

A detailed description of the Tissue Preferential Algorithm has been provided as supplemental material in a Word-document entitled "Materials supplement.doc".

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-7-9-S1.doc>]

Acknowledgements

We would like to thank Marianne van Sterkenburg (Department of Pulmonology, LUMC, Leiden, The Netherlands) for technical support, Melinda Oroszlán for providing the HUVEC, Sandra van der Kooy (Department of Nephrology, LUMC, Leiden, The Netherlands) for providing the CD14+ Monocytes and René Lutter (Department of Pulmonology, AMC, Amsterdam, The Netherlands) for providing the HMC-1 mast cells. We also would like to thank Claude Backendorf (Department of Molecular Genetics, Leiden Institute of Chemistry, Leiden, The Netherlands) for providing the SPRR1A and 2A primers and for critically reading the manuscript. This research was supported by grant 902-11-092 and 903-42-085 from the Dutch Organisation for Scientific Research (NWO)

References

1. Presland RB, Jurevic RJ: **Making sense of the epithelial barrier: what molecular biology and genetics tell us about the functions of oral mucosal and epidermal tissues.** *J Dent Educ* 2002, **66**:564-574.
2. Thompson AB, Robbins RA, Romberger DJ, Sisson JH, Spurzem JR, Teschler H, Rennard SI: **Immunological functions of the pulmonary epithelium.** *Eur Respir J* 1995, **8**:127-149.
3. Bals R, Hiemstra PS: **Innate immunity in the lung: how epithelial cells fight against respiratory pathogens.** *Eur Respir J* 2004, **23**:327-333.
4. Diamond G, Legarda D, Ryan LK: **The innate immune response of the respiratory epithelium.** *Immunol Rev* 2000, **173**:27-38.
5. Takizawa H: **Airway epithelial cells as regulators of airway inflammation (Review).** *Int J Mol Med* 1998, **1**:367-378.
6. Laberge S, El Bassam S: **Cytokines, structural cells of the lungs and airway inflammation.** *Paediatr Respir Rev* 2004, **5 Suppl A**:S41-S45.
7. Pilewski JM, Albelda SM: **Adhesion molecules in the lung. An overview.** *Am Rev Respir Dis* 1993, **148**:S31-S37.
8. Barker JN, Mitra RS, Griffiths CE, Dixit VM, Nickoloff BJ: **Keratinocytes as initiators of inflammation.** *Lancet* 1991, **337**:211-214.
9. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW: **Serial analysis of gene expression.** *Science* 1995, **270**:484-487.
10. Schena M, Shalon D, Davis RW, Brown PO: **Quantitative Monitoring of Gene-Expression Patterns with a Complementary-Dna Microarray.** *Science* 1995, **270**:467-470.
11. Vos JB, van Sterkenburg MA, Rabe KF, Schalkwijk J, Hiemstra PS, Datson NA: **Transcriptional response of bronchial epithelial**

- cells to *Pseudomonas aeruginosa*: identification of early mediators of host defense. *Physiol Genomics* 2005, **21**:324-336.
12. Jansen BJ, van Ruissen F, de Jongh G, Zeeuwen PL, Schalkwijk J: **Serial analysis of gene expression in differentiated cultures of human epidermal keratinocytes.** *J Invest Dermatol* 2001, **116**:12-22.
 13. Cookson W: **The immunogenetics of asthma and eczema: a new focus on the epithelium.** *Nat Rev Immunol* 2004, **4**:978-988.
 14. Moreno JC, Pauws E, Van Kampen AH, Jedlickova M, de Vijlder JJ, Ris-Stalpers C: **Cloning of tissue-specific genes using serial analysis of gene expression and a novel computational substraction approach.** *Genomics* 2001, **75**:70-76.
 15. Pauws E, Veenboer GJ, Smit JW, de Vijlder JJ, Morreau H, Ris-Stalpers C: **Genes differentially expressed in thyroid carcinoma identified by comparison of SAGE expression profiles.** *FASEB J* 2004, **18**:560-561.
 16. Nemes Z, Steinert PM: **Bricks and mortar of the epidermal barrier.** *Exp Mol Med* 1999, **31**:5-19.
 17. Kalinin AE, Kajava AV, Steinert PM: **Epithelial barrier function: assembly and structural features of the cornified cell envelope.** *Bioessays* 2002, **24**:789-800.
 18. Robinson NA, Lopic S, Welter JF, Eckert RL: **S100A11, S100A10, annexin I, desmosomal proteins, small proline-rich proteins, plasminogen activator inhibitor-2, and involucrin are components of the cornified envelope of cultured human epidermal keratinocytes.** *J Biol Chem* 1997, **272**:12035-12046.
 19. Abrahamson M, Alvarez-Fernandez M, Nathanson CM: **Cystatins.** *Biochem Soc Symp* 2003:179-199.
 20. Zettergren JG, Peterson LL, Wuepper KD: **Keratinin: the soluble substrate of epidermal transglutaminase from human and bovine tissue.** *Proc Natl Acad Sci U S A* 1984, **81**:238-242.
 21. An G, Huang TH, Tesfaigzi J, Garcia-Heras J, Ledbetter DH, Carlson DM, Wu R: **An unusual expression of a squamous cell marker, small proline-rich protein gene, in tracheobronchial epithelium: differential regulation and gene mapping.** *Am J Respir Cell Mol Biol* 1992, **7**:104-111.
 22. Tesfaigzi J, Th'ng J, Hotchkiss JA, Harkema JR, Wright PS: **A small proline-rich protein, SPRR1, is upregulated early during tobacco smoke-induced squamous metaplasia in rat nasal epithelia.** *Am J Respir Cell Mol Biol* 1996, **14**:478-486.
 23. Luo A, Kong J, Hu G, Liew CC, Xiong M, Wang X, Ji J, Wang T, Zhi H, Wu M, Liu Z: **Discovery of Ca²⁺-relevant and differentiation-associated genes downregulated in esophageal squamous cell carcinoma using cDNA microarray.** *Oncogene* 2004, **23**:1291-1299.
 24. Tesfaigzi Y, Wright PS, Belinsky SA: **SPRR1B overexpression enhances entry of cells into the G₀ phase of the cell cycle.** *Am J Physiol Lung Cell Mol Physiol* 2003, **285**:L889-L898.
 25. Gibbs S, Fijneman R, Wiegant J, van Kessel AG, van de PP, Backendorf C: **Molecular characterization and evolution of the SPRR family of keratinocyte differentiation markers encoding small proline-rich proteins.** *Genomics* 1993, **16**:630-637.
 26. Mischke D, Korge BP, Marenholz I, Volz A, Ziegler A: **Genes encoding structural proteins of epidermal cornification and S100 calcium-binding proteins form a gene complex ("epidermal differentiation complex") on human chromosome 1q21.** *J Invest Dermatol* 1996, **106**:989-992.
 27. Edgar R, Domrachev M, Lash AE: **Gene Expression Omnibus: NCBI gene expression and hybridization array data repository.** *Nucleic Acids Res* 2002, **30**:207-210.
 28. Lash AE, Tolstoshev CM, Wagner L, Schuler GD, Strausberg RL, Riggins GJ, Altschul SF: **SAGEmap: a public gene expression resource.** *Genome Res* 2000, **10**:1051-1060.
 29. Boon K, Osorio EC, Greenhut SF, Schaefer CF, Shoemaker J, Polyak K, Morin PJ, Buetow KH, Strausberg RL, De Souza SJ, Riggins GJ: **An anatomy of normal and malignant gene expression.** *Proc Natl Acad Sci U S A* 2002, **99**:11287-11292.
 30. Liang P: **SAGE Genie: a suite with panoramic view of gene expression.** *Proc Natl Acad Sci U S A* 2002, **99**:11547-11548.
 31. van Wetering S, van der Linden AC, van Sterkenburg MA, de Boer WI, Kuijpers AL, Schalkwijk J, Hiemstra PS: **Regulation of SLPI and elafin release from bronchial epithelial cells by neutrophil defensins.** *Am J Physiol Lung Cell Mol Physiol* 2000, **278**:L51-L58.
 32. van Ruissen F, de Jongh GJ, Zeeuwen PL, Van Erp PE, Madsen P, Schalkwijk J: **Induction of normal and psoriatic phenotypes in submerged keratinocyte cultures.** *J Cell Physiol* 1996, **168**:442-452.
 33. Zuyderduyn S, Hiemstra PS, Rabe KF: **TGF-beta differentially regulates TH2 cytokine-induced eotaxin and eotaxin-3 release by human airway smooth muscle cells.** *J Allergy Clin Immunol* 2004, **114**:791-798.
 34. Butterfield JH, Weiler DA: **In vitro sensitivity of immature human mast cells to chemotherapeutic agents.** *Int Arch Allergy Appl Immunol* 1989, **89**:297-300.
 35. Cabral A, Voskamp P, Cleton-Jansen AM, South A, Nizetic D, Backendorf C: **Structural organization and regulation of the small proline-rich family of cornified envelope precursors suggest a role in adaptive barrier function.** *J Biol Chem* 2001, **276**:19231-19237.
 36. **Ensembl Genome Browser [www.ensembl.org].** 2005.
 37. **NCBI Gene Expression Omnibus [http://www.ncbi.nlm.nih.gov/projects/geo/].** 2005.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

