Genes and Immunity (2010) 11, 458–466

www.nature.com/gene

ORIGINAL ARTICLE

Both copy number and sequence variations affect expression of human DEFB4

M Groth[1], C Wiegand[2], K Szafranski[1], K Huse[1], M Kramer[1], P Rosenstiel[3], S Schreiber[3], J Norgauer[2] and M Platzer[1]

[1]*Genome Analysis, Leibniz Institute for Age Research—Fritz Lipmann Institute, Jena, Germany;* [2]*Department of Dermatology, University Medical Center Jena, Jena, Germany and* [3]*Institute of Clinical Molecular Biology, Christian Albrechts University, Kiel, Germany*

Copy number variations (CNVs) were found to contribute massively to the variability of genomes. One of the best studied CNV region is the β-defensin cluster (DEFB) on 8p23.1. Individual DEFFB copy numbers (CNs) between 2 and 12 were found, whereas low CNs predispose for Crohn's disease. A further level of complexity is represented by sequence variations between copies (multisite variations, MSVs). To address the relation of DEFB CN and MSV to the expression of β-defensin genes, we analyzed DEFB4 expression in B-lymphoblastoid cell lines (LCLs) and primary keratinocytes (normal human epidermal keratinocyte, NHEK) before and after stimulation with lipopolysaccharide, tumor necrosis factor-α (TNF-α) and interferon-γ (IFN-γ). Moreover, we quantified one DEFB4 MSV in DNA and mRNA as a marker for variant-specific expression (VSE) and resequenced a region of ~2 kb upstream of DEFB4 in LCLs. We found a strong correlation of DEFB CN and DEFB4 expression in 16 LCLs, although several LCLs with very different CNs exhibit similar expression levels. Quantification of the MSV revealed VSE with consistently lower expression of one variant. Costimulation of NHEKs with TNF-α/IFN-γ leads to a synergistic increase in total DEFB4 expression and suppresses VSE. Analysis of the DEFB4 promoter region showed remarkably high density of sequence variabilities (~1 MSV/41 bp).
Genes and Immunity (2010) **11,** 458–466; doi:10.1038/gene.2010.19; published online 6 May 2010

Keywords: *DEFB4; copy number variation; defensin; variant-specific expression; multisite variation; keratinocytes*

Introduction

Defensins are known to have a role in adapted and innate immunity. They have biocidal effects on bacteria, fungi and viruses, as well as cytokine-like functions.[1–9] The majority of defensin genes are located on 8p23.1. The region is subdivided in the α-defensin cluster (DEFA), harboring all α-defensin genes, and the β-defensin cluster (DEFB), containing several β-defensin genes. Both clusters are affected by a common copy number variation (CNV) independently of each other. Within DEFA, the CNV unit of 19 kb encompasses *DEFA1A3* and the pseudogene *DEFTP*.[10–14] However, the 200 kb DEFB varies *en bloc* in its copy number (CN).[11] Individuals carrying 2–12 copies per diploid genome were found.[11–13,15,16] With respect to size, gene content and numeric variability of DEFB, 8p23.1 represents one of the CNV hotspots in the human genome. Genome-wide CNVs contribute massively to the genetic heterogeneity of humans. Whereas ~0.3% of the human genome is covered by single-nucleotide polymorphisms (SNPs), ~12% of the human genome is affected by CNVs.[17] It is known that SNPs may affect gene expression and hence determine phenotypes. Very likely, this is also the case for CNVs.[18,19] Therefore, it is not surprising that the DEFB CNV is associated with clinically important phenotypes. High numbers of DEFB are associated with psoriasis,[20] whereas in the case of Crohn's disease, it is under debate whether low[21] or high CNs[22] are predisposing. Whether and how DEFB CN and expression are correlated is not well known. Considering the gene dosage model, one would expect that transcript level is proportional to CN. This was shown for certain genes on chromosome 21 in patients with Down's syndrome.[23] Recently, a gene dosage effect of *FCGR3A* was reported for individuals harboring one, two or three copies.[24] For *FCGR3B*, this effect was also shown for individuals harboring one or two copies.[25,26] Importantly, low *FCGR3B* CN is associated with glomerulonephritis in the human autoimmune disease systemic lupus erythematosus, whereas the orthologous rat *Fcgr3* gene is a determinant of susceptibility to immunologically mediated glomerulonephritis.[27] In case of β-defensins, Hollox *et al*.[12] found that the transcript level of *DEFB4* is moderately correlated with its CN as determined for seven B-lymphoblastoid cell lines (LCLs). Furthermore, Fellerman *et al*.[21] found a positive correlation considering *DEFB4* transcript level and CN in mucosal specimens obtained from patients suffering from Crohn's disease and ulcerative colitis. A correlation was also observed for

Correspondence: Dr M Groth, Fritz Lipmann Institute—Leibniz Institute for Age Research, Beutenbergstr. 11, 07745 Jena, Germany. E-mail: mgroth@fli-leibniz.de
Received 19 October 2009; revised and accepted 2 March 2010; published online 6 May 2010

the *DEFB4* protein (hBD-2) in the serum of healthy individuals but not in psoriasis patients with active disease.[28] However, in patients suffering from inflammatory bowel disease, transcription of *DEFB4* is more correlated with disease activity than with the CN.[29]

In addition to CNV, sequence variations between copies (multisite variations, MSVs)[30] represent a further level of genomic complexity and are suggested to determine phenotypes, as SNPs do. An association study of *DEFB104* and prostate cancer considered both CN and four MSVs (haplotypes thereof) around exon 1.[31] Four common haplotypes were found to be associated with the risk of prostate cancer. Moreover, high CNs (>9) were significantly underrepresented in both patient samples. This implies a role of both CNs and sequence variations in shaping putative β-defensin-related phenotypes, and therefore MSVs must also to be taken into account when performing β-defensin association studies.

To investigate the relation of DEFB CN and MSVs to the expression of β-defensin genes, we analyzed *DEFB4* in LCLs and primary keratinocytes (normal human epidermal keratinocytes, NHEK) before and after stimulation with lipopolysaccharide (LPS), tumor necrosis factor-α (TNF-α) and interferon-γ (IFN-γ). To investigate whether *DEFB4* copies differ in their expression, we quantified a coding MSV in exon 2 as a marker for variant-specific expression (VSE). As nuclear factor (NF)-κB transcription factor-binding sites (TFBSs) upstream of *DEFB4* mediate basal and induced expressions,[32] we resequenced a region of ~2 kb upstream of the first exon of *DEFB4* in LCLs.

We found a strong correlation of DEFB CN and *DEFB4* expression in 16 LCLs, although several LCLs with very different CNs exhibit similar expression levels. Quantification of the MSV in DNA and mRNA showed VSE with consistently lower expression of one variant. No induction of *DEFB4* expression was detectable in LCLs. In NHEKs, *DEFB4* expression can be stimulated using TNF-α or IFN-γ. Furthermore, costimulation with TNF-α and IFN-γ lead to a synergistic increase in total *DEFB4* expression and suppresses VSE. Analysis of the *DEFB4* promoter region showed remarkably high sequence variability (~1 MSV/41 bp).

Results

Basal DEFB4 *expression in LCLs*

Among 16 LCLs with DEFB CNs from 2 to 9 determined by multiplex ligation-dependent probe amplification (MLPA), the expression of *DEFB4* varied within four orders of magnitude (Figure 1). Among cell lines with the same DEFB CN, the expression varied up to ~50-fold. DEFB CN and *DEFB4* expressions are significantly correlated (Spearman's rank correlation coefficient $r_s = 0.77$, $P = 0.0001$) and regression analysis shows an exponential dependency between the two parameters. Nevertheless, 9 out of the 16 LCLs show a very similar *DEFB4* expression (between 10- and 100-fold change, $r_s = 0.43$, $P = 0.223$), although varying in their DEFB CNs from 2 to 7 (Figure 1).

Variant-specific DEFB4 *expression in LCLs*

To characterize sequence variations that change *DEFB4* expression, we quantified as a proxy the silent polymorphism rs2740090 in the protein-coding

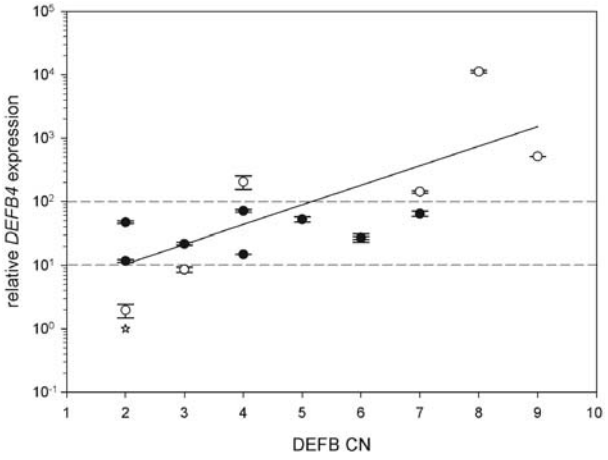


Figure 1 Relative *DEFB4* expression vs DEFB CN. Regression analysis shows exponential dependency (straight line). Relative *DEFB4* expression was calculated by normalizing to the cell line GM18972 with the lowest expression (star symbol). Error bars indicate s.d. Filled circles indicate fold-change *DEFB4* expression between 10- and 100-fold (dotted line), open circles indicate fold change <10 and >100, respectively. The two samples harboring six DEFB CNs show similar expression of *DEFB4*. This results in overlying symbols and error bars.

sequences of *DEFB4* exon 2 at both genomic and mRNA levels. Only cell lines were included, which were identified by direct sequencing to differ between copies at this position. Knowing the DEFB CN, the genomic frequency of the variants allows the estimation of the absolute number of copies harboring the respective variant in the respective cell lines (Table 1). For instance, in the simplest case, LCL GM12864 harbors two DEFB copies and showed frequencies of 50% for each variant, one DEFB copy contains the C and the other the T variant (similar to a common biallelic locus). Comparative analysis to the chimpanzee genome revealed T as the ancestral variant (UCSC Genome Browser, human genome hg18, chimpanzee genome panTro2). Until now, there is no information in dbSNP 130 about the frequency of the variants in populations. The maximum number of copies harboring the C variant per diploid genome is two and was observed in LCLs with DEFB CNs ranging from 4 to 9. With regard to all LCLs analyzed, the amount of the C variant is 32.8% (C/T 19/39). As the cell lines were chosen regarding the DEFB CN, it is questionable whether the observed frequencies of the variants reflect the frequency in human populations. If both variants were transcribed at the same level, one would expect the same C and T fractions in mRNA as in genomic DNA. However, the frequency of the C variant in *DEFB4* transcripts is always reduced compared with its genomic occurrence, indicating VSE with C as the less transcribed variant (two-way ANOVA (analysis of variance), Holm–Sidak method, $P < 0.001$, Figure 2).

Remarkably, the mRNA variant frequency is only weakly correlated with the *DEFB4* expression level ($r_s = 0.25$). In relation to the expression of the T variant (100%), the expression of the C variant varies among cell lines from 34 to 83% (Figure 2). Thereby, the relative expression of the C variant correlates with the DEFB CN (Pearson's product–moment correlation $r = 0.66$; $P = 0.014$).

Table 1 DEFB CN and MSV rs2740090 diplotypes of cell lines

Cell line ID	*Cell type*	*DEFB CN*[a]	*rs2740090*[b]	*No. of clones of promoter analysis*
GM18972[HM]	B-lymphoblastoid cell line	2	TT	28
GM12716	B-lymphoblastoid cell line	2	TC	61
GM18552[HM]	B-lymphoblastoid cell line	2	TC	60
GM12864	B-lymphoblastoid cell line	2	TC	30
C0766	B-lymphoblastoid cell line	3	TTC	—
C0913	B-lymphoblastoid cell line	3	TTC	—
GM15386	B-lymphoblastoid cell line	4	TTCC	42
GM15324	B-lymphoblastoid cell line	4	TTCC	—
GM15213	B-lymphoblastoid cell line	4	TTTC	53
GM15215	B-lymphoblastoid cell line	5	TTTTC	—
GM12760	B-lymphoblastoid cell line	6	TTTTCC	87
GM19140[HM]	B-lymphoblastoid cell line	6	TTTTTTT	84
GM10847	B-lymphoblastoid cell line	7	TTTTTCC	—
GM19204[HM]	B-lymphoblastoid cell line	7	TTTTTCC	—
GM18858[HM]	B-lymphoblastoid cell line	8	TTTTTTTT	—
GM18502[HM]	B-lymphoblastoid cell line	9	TTTTTTTCC	109
NHEK306	Primary keratinocytes	4	TTTC	—
NHEK506	Primary keratinocytes	5	TTTCC	—
NHEK606	Primary keratinocytes	6	TTTCCC	—
NHEK707	Primary keratinocytes	9	TTTTTCCCC	—

Abbreviations: CN, copy number; DEFB, β-defensin cluster; HM, cell lines included in the HapMap project;[56] MSV, multisite variation. Bold items indicate cell lines used for promoter analysis.
[a]For more information on how DEFB CN was obtained, see Supplementary Information and Groth *et al*.[11]
[b]No. of T or C = no. of respective DEFB/*DEFB4* variants (for example, TT = 2 × T variants; TTTCC = 3 × T+2 × C variants).

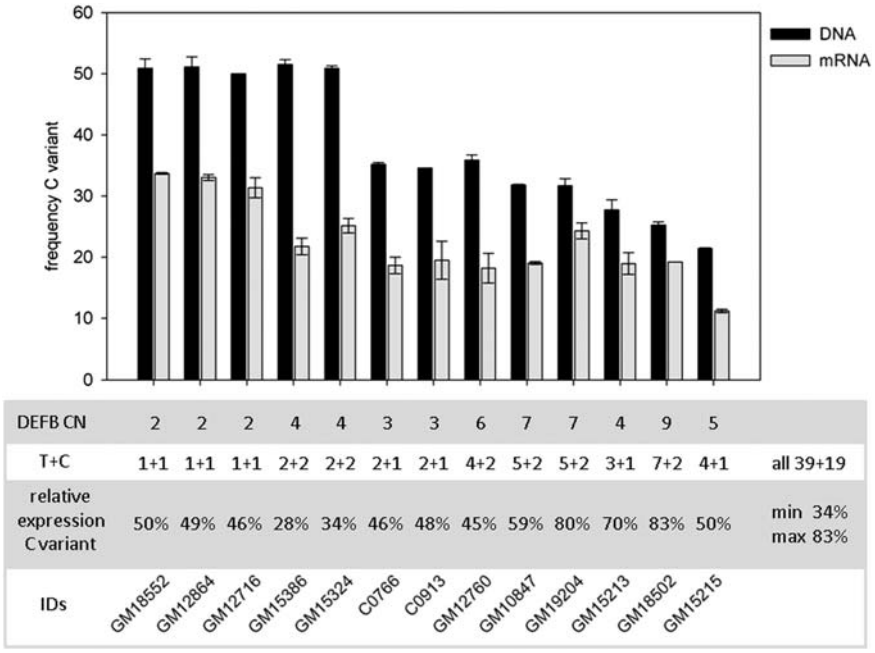


Figure 2 LCLs: frequency of the C variant (*y* axis) of MSV rs2740090 in genomic DNA (black bars) and mRNA (gray bars). The lines below show: DEFB CNs, distribution of variants to DEFB copies (T + C), the calculated relative expression of the C variant and IDs of the cell lines. Two-way ANOVA test shows significant differences of DNA vs mRNA ($P < 0.001$). Error bars indicate s.d.

Expression of *DEFB4* is inducible by LPS, TNF-α and IFN-γ.[32–34] In our experiments, LCLs did not show any increase in *DEFB4* expression after treatment (data not shown). Therefore, we switched to NHEKs known to exhibit basal and inducible *DEFB4* expressions.[34]

Inducible and variant-specific DEFB4 *expression in keratinocytes*

Using four NHEKs, DEFB CNs were determined by MLPA in genomic DNA, and MSV rs2740090 was quantified in both genomic DNA and RNA. *DEFB4* expression was analyzed with and without treatment by LPS, TNF-α, IFN-γ and a mixture of TNF-α and IFN-γ.

In contrast to LCLs, the basal *DEFB4* expression in the four NHEKs is weakly correlated with the DEFB CN ($r_s = 0.4$). To check whether this result could be an effect of the small sample size, a permutation test was performed calculating the correlation coefficient of 4 randomly selected out of the 16 LCLs as described in the section 'Materials and methods'. It showed that in

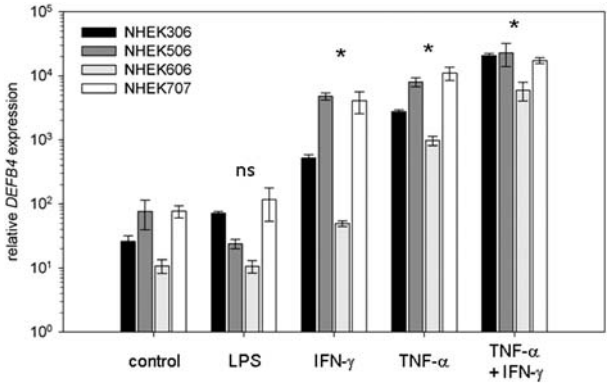


Figure 3 Relative *DEFB4* expression in NHEKs, basal and after stimulation with LPS, TNF-α, IFN-γ and a mixture of TNF-α and IFN-γ. Error bars indicate s.d. The asterisks (*) indicate significant increase in *DEFB4* expression using two-way ANOVA analysis ($P<0.001$), NS, not significant.

22.38% of the cases, a correlation $\leqslant 0.4$ was obtained by using only four cell lines.

In contrast to LPS that did not affect *DEFB4* expression (two-way ANOVA, Holm–Sidak method, $P=0.507$), the expression of *DEFB4* was significantly induced with TNF-α ($P<0.001$), IFN-γ ($P<0.001$) and with a mixture of TNF-α and IFN-γ ($P<0.001$) (Figure 3). The induction by TNF-α is higher than by IFN-γ. Combined treatment by TNF-α and IFN-γ showed a synergistic effect. Furthermore, the expression reached a plateau after induction with a mixture of $20\,ng\,ml^{-1}$ TNF-α and $20\,ng\,ml^{-1}$ IFN-γ and did not increase significantly using $40+20\,ng\,ml^{-1}$ or $20+40\,ng\,ml^{-1}$, respectively (data not shown). Basal and TNF-α, as well as basal and IFN-γ-induced *DEFB4* expressions are highly correlated (both $r_s=0.8$), whereas basal and costimulated *DEFB4* expressions (TNF-α and IFN-γ) are only weakly correlated ($r_s=0.2$).

DEFB CN and genomic variant distribution of the NHEKs are summed up in Table 1 and Figure 4. Similar to LCLs, the C variant is consistently lower expressed in NHEKs (between 26 and 66%; two-way ANOVA, Holm–Sidak method, $P<0.001$). Furthermore, the amount of the C variant in the mRNA does not change after treatment with LPS, TNF-α or IFN-γ. The amount of the C variant increases after treatment with a mixture of TNF-α and IFN-γ (between 57 and 98%; two-way ANOVA, Holm–Sidak method, $P<0.001$). With regard to NHEK306 and NHEK707, the relative expression of the C variant increases up to ~98%, and therefore the variant specificity is almost abrogated and both variants are equally expressed.

Promoter analysis of DEFB4

To shed further light on sequence variations potentially causative for the observed expression phenotype, we resequenced a 2-kb region upstream of the *DEFB4* transcription start site of 9 LCLs, selected from the 16 LCLs analyzed previously. According to DEFB CN data, these 9 LCLs were expected harboring a total of 37 DEFB copies. The samples were chosen to reflect the whole range of *DEFB4* expression and DEFB CN. Therefore, LCLs were included harboring the same β-defensin CN, but differ in their *DEFB4* expression and *vice versa*.

The analyzed region contains 43 annotated variations (dbSNP 130). The assembly (hg18) contains two DEFBs flanking a gap (but only in one *DEFB4* is annotated). In some cases, different annotations describe the same variation, one in each DEFB. Therefore, some variations have two accession numbers (27 annotations are unique in hg18). Altogether, we identified 49 variations (1/41 bp, 43 single-nucleotide variations, 6 indels) of which 20 confirmed the dbSNP annotations. The applied resequencing strategy and the extraordinary variation density allowed us to unambiguously reconstruct 25 haplotypes (Figure 5). The majority of these haplotypes (20 out of 25) were unique among the 37 DEFB copies. Two haplotypes occur twice, two three times and only one four times. Except for GM18502 (DEFB CN: 9), the number of haplotypes per LCL reflects the DEFB CN, which means that in 8 out of 9 cell lines, each *DEFB4* copy harbors a different haplotype.

As defensin expression is mediated by the NF-κB pathway,[35,36] we checked the analyzed region regarding NF-κB TFBSs. Five potential NF-κB TFBS were found at following positions: chr8: 7 788 252–7 788 261; chr8: 7 789 028–7 789 036; chr8: 7 789 374–7 789 313; chr8: 7 789 405–7 789 414 and chr8: 7 789 414–7 789 423. NF-κB TFBS at position chr8: 7 789 374–7 789 313 is destroyed (C variant) or created (G variant) by a variation found in five analyzed LCLs (rs2698827). Almost no correlation between the frequency of TFBS (G variant) and relative expression of *DEFB4* was found ($r_s=0.15$). Furthermore, no variation or haplotype could be assigned to or correlated with either high or low *DEFB4* expression.

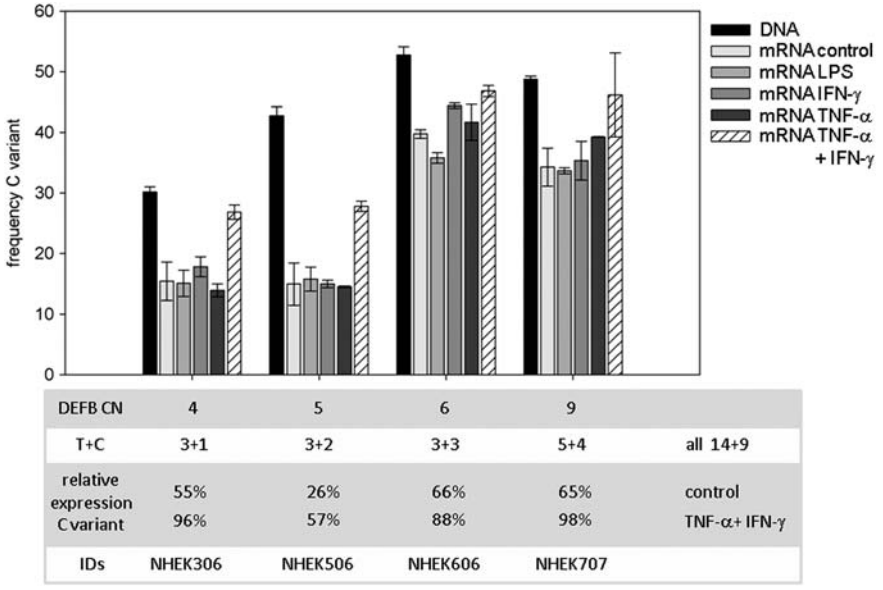
All the four NHEKs are polymorphic for rs2698827. As for the LCLs, we cannot assign the C or G variant either to the total expression or to the inducibility or VSE of *DEFB4* (data not shown).

The expression of genes is also regulated by methylation of CpGs. CpGPlot did not show CpG islands within the analyzed region, but 11 out of the 49 variations create or destroy single CpGs. However, no variation was found correlating with high or low expressed *DEFB4*.

Discussion

Although β-defensins within DEFB are important components of innate immunity, it seems that their variability is an underappreciated source of phenotypic variation. Only few association studies, remarkably designed as candidate gene approaches, have been published. This may in part be due to methodological problems in the characterization of the superimposed types of polymorphisms: structural variation (CNV) and sequence variation (MSV). Although scalable methods to determine CNV in high-throughput screenings lack accuracy,[37] MSVs are resistant to common genotyping. However, both are important to know in order to correlate the genomic state of β-defensins with the phenotypes. CNVs are responsible for different gene dosage in individuals, and MSVs might influence gene expression or mRNA/protein properties.

The observed concordance of DEFB gene numbers[11] means that the CN determined for a one gene or locus serves as a proxy for any other gene/locus of the cluster. However, it is not known whether DEFB MSVs are in linkage equilibrium or whether they are forming only a



DEFB CN	4	5	6	9	
T+C	3+1	3+2	3+3	5+4	all 14+9
relative expression C variant	55%	26%	66%	65%	control
	96%	57%	88%	98%	TNF-α+ IFN-γ
IDs	NHEK306	NHEK506	NHEK606	NHEK707	

Figure 4 NHEKs: frequency of the C variant (*y* axis) of MSV rs2740090 in genomic DNA (black bars), mRNA with or without treatment (gray bars and streaked bars, respectively). The lines below show: DEFB CNs, distribution of variants to DEFB copies (T + C), the calculated relative expression of the C variant (control and costimulation) and IDs of the cell lines. Two-way ANOVA test shows significant differences of DNA vs mRNA control and mRNA control vs mRNA TNF-α + IFN-γ (each $P < 0.001$). Error bars indicate s.d.

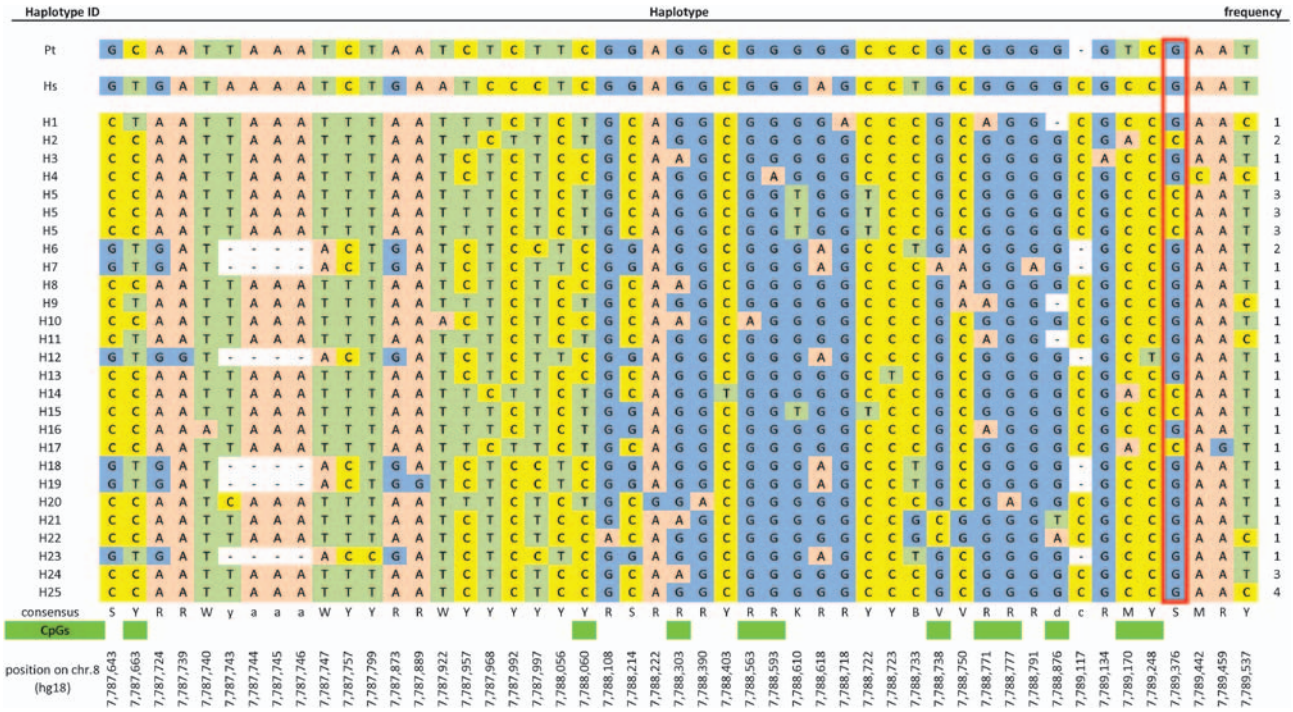


Figure 5 Haplotypes in the putative *DEFB4* promoter region in nine LCLs. Human haplotypes are numbered from H1 to H25. A chimpanzee haplotype (Pt) was build on the basis of UCSC Genome Browser[52] data (*Pan troglodytes* assembly March 2006 (panTro2). Nucleotide color code: A—pink; C—yellow; G—blue; T—green). The rightmost column contains the number of haplotypes. The variation marked by a red box destroys/creates an NF-κB TFBS. The IUPAC consensus sequence summarizes the found haplotypes (capital letters indicate nucleotide exchanges, lower case letters indicate nucleotide exchanges in combination with an insertion/deletion). The variations marked in green in the lower line destroy or create CpGs. Positions of the MSVs on chromosome 8 are based on UCSC Genome Browser, human genome hg18, and are shown in the lower line.

few common haplotypes. Moreover, determining the diploid DEFB CN of a sample (diplotype), we do not know the constituent CN alleles (genotype; terminology according to the study by Hollox[38]). Consequently, it is currently completely unknown whether particular MSVs and/or MSV haplotypes occur preferentially in 1, 2, 3 or 4 copy alleles or with the same frequency in all DEFB CN allele classes.

The study presented herein is an integrated approach, which investigates the *DEFB4* expression by regarding both DEFB CN and one MSV (rs2740090). *DEFB4* expression was studied in LCLs and NHEKs characterized by their MLPA-determined total DEFB CN subdivided into C and/or T variants. We determined *DEFB4* expression by TaqMan assays as basal expression and, in NHEKs only, as stimulated expression after challenging the cells with known *DEFB4*-inducing agents (such as LPS, TNF-α and IFN-γ). The ratios of rs2740090 variants were also determined in *DEFB4* transcripts to evaluate their possible differential expression.

In LCLs, we confirmed CN-dependent *DEFB4* basal expression ($r_s = 0.76$)[12,29] in an exponential manner. This was also shown for the expression of *DEFB4* and *DEFB103* in mucosal specimen.[21] Considering a gene dose-dependent model, one would expect a linear dependency between CN and expression level, as reported for *FCGR3B*.[25] In hemizygous individuals, half the protein amount was found in the plasma. In contrast, we found that the *DEFB4* mRNA expression does not change in a linear but in an exponential manner regarding DEFB CN. AS DEFB is amplified as an entity[39] and each copy of the cluster is flanked by low copy repeats,[16] the exponential dependency might be explained by chromatin remodeling due to stringed arrangement of DEFB copies even cytogenetically visible by fluorescent in situ hybridization.[12] If so, in the future not only the DEFB diplotype has to be taken into account but also the number of DEFB copies per chromosome (genotype). Possible distributions of DEFB copies in one individual sample harboring six DEFBs can be: 1 + 5, 2 + 4, 3 + 3 (we never found an individual sample with one copy, therefore we suggest that the null allele is at least very rare). Furthermore, by comparing original and derived loci in segmental duplication events, Zheng[40] showed that different copies have different histone modifications, which in turn could lead to different expressions of certain copies and consequently to differences in transcript levels.

Among the analyzed 16 LCLs, we identified samples harboring the same DEFB CN, but showing differences in *DEFB4* expression, and *vice versa*, showing similar expression while harboring different DEFB CNs (Figure 1). Therefore, expression of *DEFB4* is not only controlled by CN but also by other factors. *DEFB4* expression was not inducible in LCLs. This may be attributable to ligand-independent activation of NF-κB in these cells by EBV proteins,[41] which may result in a permanent induction of *DEFB4* expression in this cell lines.

Investigation of MSV rs2740090 in exon 2 of *DEFB4* revealed VSE in all analyzed LCLs and in NHEKs. In general, the allele-specific expression (ASE) of a non-CNV biallelic locus, was shown to be widespread in the human genome[42–44] and associated with disease.[45] Surprisingly, regarding all analyzed cell lines, we found that *DEFB4* harboring the C variant is consistently less transcribed. There is no evidence that the analyzed MSV is causative for VSE. We rather suggest a linked variant, which causes the expression difference. Interestingly, the relative expression of the C variant varies in different cell lines more than twofold from 34 to 83%. Furthermore, the relative expression of the C variant correlates with DEFB CN; this means the higher the CN, the lesser the variant specificity during transcription. A possible chromatin remodeling due to stringed arrangement of the DEFB copies may be suspected to affect regulatory elements causative for VSE in the same manner as for CN-dependent *DEFB4* expression.

With regard to *IL10*, it was shown that different haplotypes (microsatellites) are strongly associated with differences in interleukin-10 secretion.[46] Expression of a specific variant was also observed regarding *KRT1*. Cis-regulatory SNPs in the promoter region of this gene were found to modulate promoter activity and to cause ASE.[47] Treatment of NHEKs with a mixture of TNF-α and IFN-γ (the best stimulus in our hands) led to the suppression of the VSE nearly equalizing the amount of the C variant in the *DEFB4* transcripts to those on the genomic level. This may result either from expression saturation of the T variant or preferential induction of the C variant. Such a shift in expression of specific variants was reported in barley hybrids, changing the allelic ratio in some genes after drought stress (for example, ABC02113, ABC03499, ABC10029).[48]

To screen for possible cis-regulatory variations causative for total *DEFB4* expression or VSE, we resequenced ~2 kb upstream of *DEFB4* transcription start site in selected cell lines. We found five NF-κB-binding sites in the putative promoter region just upstream of the transcription start. These sites have been described previously and four of them were identified to mediate LPS-induced and basal expressions of *DEFB4* in macrophages.[32] Our results do not show any clues for transcription-relevant variations within the TFBS or the complete analyzed region of ~2 kb. If there is a variation linked to the consistently low-expressed C variant that causes VSE, it may not be localized within the analyzed upstream region. In addition to that, because VSE is shifted by costimulation with TNF-α and IFN-γ, we suggest that there are different regulatory elements for basal and induced *DEFB4* expressions. The variant specificity in basal expression will be diminished by regulatory elements that mediate expression by costimulating with TNF-α and IFN-γ. This fits to the findings, that cotreatment with TNF-α and IFN-γ not only has a synergistic effect on VSE but also on the total *DEFB4* transcript level. Such a synergistic effect of *DEFB4* expression was previously described by Joly *et al*.[49] However, they also report that induced expression correlates with basal expression. This is somehow contrary to our results as we found that after stimulation with either TNF-α or IFN-γ, but not with both cytokines, the induced expression correlates with the basal expression. In agreement with Joly *et al*.,[49] we also found that *DEFB4* expression reached a plateau after costimulation with TNF-α and IFN-γ, suggesting some kind of saturation. Similar results were also described by Fellermann *et al*.[21] for *DEFB4* expression in mucosa specimen. Taken together, there is a synergistic effect of the mixture of TNF-α and IFN-γ on both VSE and total *DEFB4* expressions.

As defensins and cytokines that mediate defensin expression are involved in inflammatory diseases, it is not surprising that Aldhous *et al*.[29] and Jansen *et al*.[28] found that *DEFB4* expression reflects the disease activity more than the respective DEFB CN. Therefore, not only cis- but also trans-regulatory elements have to be taken into account when considering *DEFB4* expression. In this context, it is of value to possess a sequence tag in *DEFB4* (rs2740090), which is associated with a molecular phenotype: One variant (T) shows higher basal expres-

sion, whereas the other (C) indicates better inducibility. It would be of interest to investigate whether the variant distributions differ in patients suffering from Crohn's disease or psoriasis.

Altogether, the basal *DEFB4* expression correlates with the DEFB CN, *DEFB4* copies harboring the C variant are less expressed than the T variant, and there is a synergistic effect of TNF-α and IFN-γ on the VSE and the total *DEFB4* expression. Therefore, performing an association study considering DEFB CN and disease with inflammatory background, one has to keep in mind that defensin gene variants seem to be induced differentially. For the future, it would be a challenge to identify sequence variations between copies affecting the induced *DEFB4* expression.

Materials and methods

Cells and cell culture

Human LCLs were purchased from Coriell Institute for Medical Research (Camden, NJ, USA; naming: GMxxxx) and ECACC (Wiltshire, UK; naming: Cxxxx) (Table 1). LCLs were cultured in RPMI 1640 with GlutaMAX (Gibco, Carlsbad, CA, USA) supplemented with 15% fetal bovine serum 'GOLD' (PAA Laboratories, Pasching, Austria) in 25 cm^2 BD Falcon flasks (Becton Dickinson, Franklin Lakes, NJ, USA) at 37 °C in a humidified 5% CO_2 atmosphere in a total volume of 10 ml. Cells were grown to a density of 6×10^5 cells per ml and split in a ratio of 1:4. The cell lines characterized in this study were selected by their DEFB CNs covering the range from 2 to 8 copies, as previously determined by Groth *et al*.[11]

NHEKs have been established at the Department of Dermatology using keratinocytes isolated from the foreskins of male infants undergoing routine surgery due to phimosis (circumcision) through enzymatic treatment of the epidermis.[50] Written parental consent was obtained for all patients. NHEKs were cultured in keratinocyte basal medium (PromoCell, Heidelberg, Germany) augmented with Supplement Pack C-39011 (PromoCell) at 37 °C and in a 5% CO_2 atmosphere.

For experiments, the cells were harvested by trypsin-EDTA (Invitrogen, Karlsruhe, Germany) treatment and seeded into 12-well culture plates (Greiner, Frickenhausen, Germany) at a density of 4×10^5 cells per cm^2. Seventy-two hours after plating, the cell medium was replaced by medium containing LPS (resolved in phosphate-buffered saline, PromoCell) from *Salmonella typhimurium* (Sigma-Aldrich, St Louis, MO, USA) at concentration of 1 μg ml^{-1} or TNF-α (R&D Systems, Minneapolis, MN, USA; 210-TA/CF, concentration 20 ng ml^{-1}) or IFN-γ (R&D Systems; 285-IF/Cf, concentration 20 ng ml^{-1}) or with a mixture of TNF-α and IFN-γ (each 20 ng ml^{-1}). The cells were then further incubated for 24 h before mRNA was isolated. Parallel incubation with medium alone was performed as control.

Genomic DsNA and mRNA/cDNA

Genomic DNA from cell cultures was isolated from 5×10^6 cells using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Total mRNA was isolated using the Qiagen RNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. The isolated mRNAs were reverse transcribed using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA). The reaction was performed in 200 μl PCR tubes (Brand, Wertheim, Germany) according to the manufacturer's instructions in a Peltier Cycler PTC-200 (MJ Research, Waltham, MA, USA). A mixture of anchored Oligo-dT primers (5′ T_{15}VN 3′) (Metabion, Martinsried, Germany) was used for priming reverse transcription.

MLPA

MLPA was carried out using the MLPA-Kit P139 (MRC Holland, Amsterdam, The Netherlands) as described by Groth *et al*.[11]

Expression analysis

Determination of transcription level of *DEFB4* was performed using TaqMan assays (TaqMan Gene Expression Assay and TaqMan Gene Expression Mastermix) purchased from Applied Biosystems. The reactions were performed in 96-well PCR plates (Bio-Rad Laboratories, Hercules, CA, USA) according to the manufacturer's instructions using an iCycler instrument (Bio-Rad Laboratories). The following TaqMan Gene Expression Assays were used: *ACTB* (Hs99999903_m1), *DEFB4* (00823638_m1) and *TBP* (99999910_m1). All reactions were performed at least in triplicates. Relative *DEFB4* expression was determined as described previously.[51]

Variant-specific transcription

PCR was performed with primers flanking rs2740090 on genomic DNA and cDNA. Amplification of genomic DNA was carried out with 50 ng DNA template in a volume of 25 μl using BioMix Red (Bioline, Luckenwalde, Germany) according to the manufacturer's instructions using forward primer 5′-GGCGATACTGACACAGGG TT-3′ and reverse primer 5′-GGCGATACTGACACAGG GTT-3′. The amplification of cDNA was performed under the same conditions using forward primer 5′-ATCAGCC ATGAGGGTCTTGT-3′ and reverse primer 5′-TTTGGTTT ACATGTCGCACG-3′. Quantification of rs2740090 variants was performed using a pyrosequencing approach as described previously[39] using sequencing primer 5′-GTGTTTTTGGTGGTATAGGCGA-3′ on a PSQ 96 MA System (Qiagen).

DEFB4 *promoter analysis*

Genomic DNA of 2 kb upstream of exon 1 of *DEFB4* (chr8: 7 787 620–7 789 657; chr8: 7 261 747–7 263 787, UCSC Genome Browser, human genome hg18[52]) was amplified using the primers 5′-AATGGAAGCATTTCTGTGGG-3′ and 5′-ACAAGACCCTCATGGCTGAT-3′. PCR was performed in a volume of 25 μl using BioMix Red (Bioline) according to the manufacturer's instructions. Direct sequencing of the amplicons failed due to insertions/deletions located in the analyzed region. In contrast to Aldhous *et al*.,[29] we therefore performed a cloning/sequencing-based approach. The amplicon was cloned using the TA Cloning Kit (with pCR 2.1 vector) (Invitrogen) according to the manufacturer's instructions. Depending on the DEFB CN of the respective individual sample, certain numbers of clones were sequenced (see Table 1) using two primers located within the amplicon (forward primer 5′-ATGAAGAGGTCAG GCAGGTC-3′, reverse primer 5′-GAAGGAGAATGGGA GGATGG-3′) and vector-specific M13 primers. To avoid false-positive variations, the number of clones to

sequence was chosen on the basis of the assumption that at least 10 clones have to represent the amplified region from one copy, meaning that DEFB CN times 10 was the least number of clones to sequence. The resulting sequences were manually inspected for variations using Genome Assembly Program 4[53] and analyzed for TFBS using Transcription Element Search System[54] and CpG islands using EMBOSS CpGPlot.[55]

Statistic analysis

All statistics were performed using SigmaPlot 11 (Systat Software Inc., Erkrath, Germany). A permutation test for calculating the possibility of finding a correlation with four LCLs was performed as follows. Data (DEFB CN and *DEFB4* expression) of four LCLs were chosen randomly from the set of 16 LCLs. Spearman's rank correlation was calculated for this subset and this procedure (data selection, correlation test) was performed one million times. The frequency of correlation coefficients reaching a threshold (r_s of the four NHEKs) reflects the statistical power of the experiment using only four cell lines.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The skilful and diligent technical assistance of Beate Szafranski and Ivonne Goerlich is gratefully acknowledged. This work was supported by grants from the Wilhelm Sander Stiftung (2005.045.1) and National Bundesministerium für Bildung und Forschung (01GS0809).

References

1 Chang TL, Vargas Jr , DelPortillo A, Klotman ME. Dual role of alpha-defensin-1 in anti-HIV-1 innate immunity. *J Clin Invest* 2005; **115**: 765–773.
2 Daher KA, Selsted ME, Lehrer RI. Direct inactivation of viruses by human granulocyte defensins. *J Virol* 1986; **60**: 1068–1074.
3 Ericksen B, Wu Z, Lu W, Lehrer RI. Antibacterial activity and specificity of the six human {alpha}-defensins. *Antimicrob Agents Chemother* 2005; **49**: 269–275.
4 Ganz T. Defensins: antimicrobial peptides of innate immunity. *Nat Rev Immunol* 2003; **3**: 710–720.
5 Garcia JR, Krause A, Schulz S, Rodriguez-Jimenez FJ, Kluver E, Adermann K *et al.* Human beta-defensin 4: a novel inducible peptide with a specific salt-sensitive spectrum of antimicrobial activity. *FASEB J* 2001; **15**: 1819–1821.
6 Lehrer RI, Ganz T, Szklarek D, Selsted ME. Modulation of the *in vitro* candidacidal activity of human neutrophil defensins by target cell metabolism and divalent cations. *J Clin Invest* 1988; **81**: 1829–1835.
7 Quinones-Mateu ME, Lederman MM, Feng Z, Chakraborty B, Weber J, Rangel HR *et al.* Human epithelial beta-defensins 2 and 3 inhibit HIV-1 replication. *Aids* 2003; **17**: F39–F48.
8 Selsted ME, Szklarek D, Lehrer RI. Purification and antibacterial activity of antimicrobial peptides of rabbit granulocytes. *Infect Immunol* 1984; **45**: 150–154.
9 Zhang L, Yu W, He T, Yu J, Caffrey RE, Dalmasso EA *et al.* Contribution of human alpha-defensin 1, 2, and 3 to the anti-HIV-1 activity of CD8 antiviral factor. *Science* 2002; **298**: 995–1000.
10 Aldred PM, Hollox EJ, Armour JA. Copy number polymorphism and expression level variation of the human alpha-defensin genes DEFA1 and DEFA3. *Hum Mol Genet* 2005; **14**: 2045–2052.
11 Groth M, Szafranski K, Taudien S, Huse K, Mueller O, Rosenstiel P *et al.* High-resolution mapping of the 8p23.1 beta-defensin cluster reveals strictly concordant copy number variation of all genes. *Hum Mutat* 2008; **29**: 1247–1254.
12 Hollox EJ, Armour JA, Barber JC. Extensive normal copy number variation of a beta-defensin antimicrobial-gene cluster. *Am J Hum Genet* 2003; **73**: 591–600.
13 Linzmeier RM, Ganz T. Human defensin gene copy number polymorphisms: comprehensive analysis of independent variation in alpha- and beta-defensin regions at 8p22-p23. *Genomics* 2005; **86**: 423–430.
14 Mars WM, Patmasiriwat P, Maity T, Huff V, Weil MM, Saunders GF. Inheritance of unequal numbers of the genes encoding the human neutrophil defensins HP-1 and HP-3. *J Biol Chem* 1995; **270**: 30371–30376.
15 Armour JA, Palla R, Zeeuwen PL, den Heijer M, Schalkwijk J, Hollox EJ. Accurate, high-throughput typing of copy number variation using paralogue ratios from dispersed repeats. *Nucleic Acids Res* 2007; **35**: e19.
16 Taudien S, Galgoczy P, Huse K, Reichwald K, Schilhabel M, Szafranski K *et al.* Polymorphic segmental duplications at 8p23.1 challenge the determination of individual defensin gene repertoires and the assembly of a contiguous human reference sequence. *BMC Genomics* 2004; **5**: 92.
17 Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. *Nat Rev Genet* 2006; **7**: 85–97.
18 Henrichsen CN, Chaignat E, Reymond A. Copy number variants, diseases and gene expression. *Hum Mol Genet* 2009; **18**: R1–R8.
19 Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N *et al.* Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science* 2007; **315**: 848–853.
20 Hollox EJ, Huffmeier U, Zeeuwen PL, Palla R, Lascorz J, Rodijk-Olthuis D *et al.* Psoriasis is associated with increased beta-defensin genomic copy number. *Nat Genet* 2008; **40**: 23–25.
21 Fellermann K, Stange DE, Schaeffeler E, Schmalzl H, Wehkamp J, Bevins CL *et al.* A chromosome 8 gene-cluster polymorphism with low human beta-defensin 2 gene copy number predisposes to Crohn disease of the colon. *Am J Hum Genet* 2006; **79**: 439–448.
22 Bentley RW, Pearson J, Gearry RB, Barclay ML, McKinney C, Merriman TR *et al.* Association of higher DEFB4 genomic copy number with Crohn's disease. *Am J Gastroenterol* 2009; **105**: 354–359.
23 Ait Yahya-Graison E, Aubert J, Dauphinot L, Rivals I, Prieur M, Golfier G *et al.* Classification of human chromosome 21 gene-expression variations in Down syndrome: impact on disease phenotypes. *Am J Hum Genet* 2007; **81**: 475–491.
24 Breunis WB, van Mirre E, Geissler J, Laddach N, Wolbink G, van der Schoot E *et al.* Copy number variation at the FCGR locus includes FCGR3A, FCGR2C and FCGR3B but not FCGR2A and FCGR2B. *Hum Mutat* 2009; **30**: E640–E650.
25 Koene HR, de Haas M, Kleijer M, Roos D, von dem Borne AE. NA-phenotype-dependent differences in neutrophil Fc gamma RIIIb expression cause differences in plasma levels of soluble Fc gamma RIII. *Br J Haematol* 1996; **93**: 235–241.
26 Willcocks LC, Lyons PA, Clatworthy MR, Robinson JI, Yang W, Newland SA *et al.* Copy number of FCGR3B, which is associated with systemic lupus erythematosus, correlates with protein expression and immune complex uptake. *J Exp Med* 2008; **205**: 1573–1582.
27 Aitman TJ, Dong R, Vyse TJ, Norsworthy PJ, Johnson MD, Smith J *et al.* Copy number polymorphism in Fcgr3 predisposes to glomerulonephritis in rats and humans. *Nature* 2006; **439**: 851–855.

28 Jansen PA, Rodijk-Olthuis D, Hollox EJ, Kamsteeg M, Tjabringa GS, de Jongh GJ *et al.* Beta-defensin-2 protein is a serum biomarker for disease activity in psoriasis and reaches biologically relevant concentrations in lesional skin. *PLoS One* 2009; **4**: e4725.
29 Aldhous MC, Noble CL, Satsangi J. Dysregulation of human beta-defensin-2 protein in inflammatory bowel disease. *PLoS One* 2009; **4**: e6285.
30 Fredman D, White SJ, Potter S, Eichler EE, Den Dunnen JT, Brookes AJ. Complex SNP-related sequence variation in segmental genome duplications. *Nat Genet* 2004; **36**: 861–866.
31 Huse K, Taudien S, Groth M, Rosenstiel P, Szafranski K, Hiller M *et al.* Genetic variants of the copy number polymorphic beta-defensin locus are associated with sporadic prostate cancer. *Tumour Biol* 2008; **29**: 83–92.
32 Tsutsumi-Ishii Y, Nagaoka I. NF-kappa B-mediated transcriptional regulation of human beta-defensin-2 gene following lipopolysaccharide stimulation. *J Leukoc Biol* 2002; **71**: 154–162.
33 Harder J, Bartels J, Christophers E, Schroder JM. A peptide antibiotic from human skin. *Nature* 1997; **387**: 861.
34 Nomura I, Goleva E, Howell MD, Hamid QA, Ong PY, Hall CF *et al.* Cytokine milieu of atopic dermatitis, as compared to psoriasis, skin prevents induction of innate immune response genes. *J Immunol* 2003; **171**: 3262–3269.
35 Kaiser V, Diamond G. Expression of mammalian defensin genes. *J Leukoc Biol* 2000; **68**: 779–784.
36 Kota S, Sabbah A, Chang TH, Harnack R, Xiang Y, Meng X *et al.* Role of human beta-defensin-2 during tumor necrosis factor-alpha/NF-kappaB-mediated innate antiviral response against human respiratory syncytial virus. *J Biol Chem* 2008; **283**: 22417–22429.
37 Wain LV, Armour JA, Tobin MD. Genomic copy number variation, human health, and disease. *Lancet* 2009; **374**: 340–350.
38 Hollox EJ. Copy number variation of beta-defensins and relevance to disease. *Cytogenet Genome Res* 2008; **123**: 148–155.
39 Groth M, Huse K, Reichwald K, Taudien S, Hampe J, Rosenstiel P *et al.* Method for preparing single-stranded DNA templates for pyrosequencing using vector ligation and universal biotinylated primers. *Anal Biochem* 2006; **356**: 194–201.
40 Zheng D. Asymmetric histone modifications between the original and derived loci of human segmental duplications. *Genome Biol* 2008; **9**: R105.
41 Gires O, Zimber-Strobl U, Gonnella R, Ueffing M, Marschall G, Zeidler R *et al.* Latent membrane protein 1 of epstein-barr virus mimics a constitutively active receptor molecule. *EMBO J* 1997; **16**: 6131–6140.
42 Lo HS, Wang Z, Hu Y, Yang HH, Gere S, Buetow KH *et al.* Allelic variation in gene expression is common in the human genome. *Genome Res* 2003; **13**: 1855–1862.
43 Palacios R, Gazave E, Goni J, Piedrafita G, Fernando O, Navarro A *et al.* Allele-specific gene expression is widespread across the genome and biological processes. *PLoS One* 2009; **4**: e4150.
44 Serre D, Gurd S, Ge B, Sladek R, Sinnett D, Harmsen E *et al.* Differential allelic expression in the human genome: a robust approach to identify genetic and epigenetic cis-acting mechanisms regulating gene expression. *PLoS Genet* 2008; **4**: e1000006.
45 Yan H, Dobbie Z, Gruber SB, Markowitz S, Romans K, Giardiello FM *et al.* Small changes in expression affect predisposition to tumorigenesis. *Nat Genet* 2002; **30**: 25–26.
46 Eskdale J, Gallagher G, Verweij CL, Keijsers V, Westendorp RG, Huizinga TW. Interleukin 10 secretion in relation to human IL-10 locus haplotypes. *Proc Natl Acad Sci USA* 1998; **95**: 9465–9470.
47 Tao H, Cox DR, Frazer KA. Allele-specific KRT1 expression is a complex trait. *PLoS Genet* 2006; **2**: e93.
48 von Korff M, Radovic S, Choumane W, Stamati K, Udupa SM, Grando S *et al.* Asymmetric allele-specific expression in relation to developmental variation and drought stress in barley hybrids. *Plant J* 2009; **59**: 14–26.
49 Joly S, Organ CC, Johnson GK, McCray Jr PB, Guthmiller JM. Correlation between beta-defensin expression and induction profiles in gingival keratinocytes. *Mol Immunol* 2005; **42**: 1073–1084.
50 Schantz J-T, Ng KW. *A Manual for Primary Human Cell Culture*. World Scientific Publishing Co. Pte. Ltd.: Singapore, 2004.
51 Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001; **29**: e45.
52 Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM *et al.* The human genome browser at UCSC. *Genome Res* 2002; **12**: 996–1006.
53 Bonfield JK, Smith K, Staden R. A new DNA sequence assembly program. *Nucleic Acids Res* 1995; **23**: 4992–4999.
54 Baxevanis AD. *Current Protocols in Bioinformatics*. John Wiley & Sons Inc: Hoboken, 2002.
55 Rice P, Longden I, Bleasby A. EMBOSS: the european molecular biology open software suite. *Trends Genet* 2000; **16**: 276–277.
56 International HapMap Consortium. A haplotype map of the human genome. *Nature* 2005; **437**: 1299–1320.

Supplementary Information accompanies the paper on Genes and Immunity website (http://www.nature.com/gene)