Gene expression analysis in predicting the effectiveness of insect venom immunotherapy

Marek Niedoszytko, MD, PhD,a,b Marcel Bruinenberg, PhD,c,d Jan de Monchy, MD, PhD,b,c Ewa Jassem, MD, PhD,a and Joanne N. G. Oude Elberink, MD, PhDb Gdansk, Poland, and Groningen, The Netherlands

Background: Venom immunotherapy (VIT) enables longtime prevention of insect venom allergy in the majority of patients. However, in some, the risk of a systemic reaction increases after completion of treatment. No reliable factors predicting individual lack of efficacy of VIT are currently available. Objective: To determine the use of gene expression profiles to predict the long-term effect of VIT.

Methods: Whole genome gene expression analysis was performed on RNA samples from 46 patients treated with VIT divided into 3 groups: (1) patients who achieved and maintained long-term protection after VIT, (2) patients in whom insect venom allergy relapsed, and (3) patients still in the maintenance phase of VIT.

Results: Among the 48,071 transcripts analyzed, 1401 showed a >2 fold difference in gene expression (P < .05); 658 genes (47%) were upregulated and 743 (53%) downregulated. Forty-three transcripts still show significant differences in expression after correction for multiple testing; 12 of 43 genes (28%) were upregulated and 31 of 43 genes (72%) downregulated. A naive Bayes prediction model demonstrated a gene expression pattern characteristic of effective VIT that was present in all patients with successful VIT but absent in all subjects with failure of VIT. The same gene expression profile was present in 88% of patients in the maintenance phase of VIT.

Conclusion: Gene expression profiling might be a useful tool to assess the long-term effectiveness of VIT. The analysis of differently expressed genes confirms the involvement of immunologic pathways described previously but also indicates novel factors that might be relevant for allergen tolerance.

(J Allergy Clin Immunol 2010; 126; 801-809.)

Key words: Insect venom allergy, venom immunotherapy, gene expression, microarray assessment, prediction of treatment efficacy

Insect venom allergy (defined as at least 1 systemic IgE mediated reaction in a lifetime after an insect sting) is present in approximately 1% to 3% of general population.1

Venom immunotherapy (VIT) with bee, yellow jacket, or Polistes venom is the treatment of choice in patients with insect venom allergy (IVA). At reaching maintenance dose, the risk of a systemic reaction to a subsequent sting is reduced from 70% (ie, before the start of VIT) to 3% to 15%.2 To reach long-term protection, the maintenance phase has to be continued for at least 3 years in patients with mild systemic reactions and at least 5 years in patients with severe systemic reactions.3 This procedure probably enables lifelong prevention of anaphylactic reactions in the majority of patients.3

However, in some patients, the risk of a systemic reaction to a re-sting reappears and increases after stopping the treatment. Currently there is no certain way to predict the individual efficacy of VIT except for deliberate sting challenges, but it is known that a number of factors are associated with a worse outcome of immunotherapy. First is the duration of treatment. The risk of a systemic reaction after 2 years of VIT is higher than in patients who stopped after 3 to 5 years (30% vs 3%).1,2,4 Second, it is known that patients with side effects during treatment are more prone to a lower degree of protection.1,2 Hence, prolongation of VIT may reduce the risk for systemic reaction in these patients.1,2 Third, the amount of allergen routinely administered might not be sufficient to stimulate full protection in all individuals. It has been shown that continuation of VIT with higher dose (eg, 200 ug) is able to reduce this risk.5 Fourth, it was demonstrated that the risk at a systemic reaction after completing the treatment is related to the culprit insect. In patients with yellow jacket venom allergy, the long-term effectiveness of therapy is assessed to be 85% to 95%, whereas in patients allergic to bee venom, this is 75% to 85%.1 Fifth, coexistence of mastocytosis and even elevated serum tryptase level might increase the risk of inefficacy of VIT.6,7 The current guidelines of European
Academy of Allergy and Clinical Immunology indicate that patients with negative skin tests and undetectable specific IgE to insect venom have a diminished risk of relapse after stopping VIT.2-4 Finally, it is known that less severe sting reactions are associated with better protection after completing the treatment.1

Overall, this means that 10% to 20% of subjects remain vulnerable to the culprit insect venom in spite of completing the treatment.1,3,6,7

The aim of this study was to determine whether gene expression profiles may predict the efficacy or inefficacy of VIT. We determined whole genome gene expression profiles of patients who successfully completed treatment and compared their gene expression profiles with patients who had repeated systemic sting reactions in spite of VIT. On the basis of these results, we built a naive Bayes prediction model that subsequently was evaluated in a group of patients still on a maintenance dose of VIT.8,9

METHODS

Patients
A total of 46 patients treated with VIT were included. All patients experienced 1 or more severe systemic reactions before starting VIT. Inclusion criteria were the diagnosis of IVA on the basis of medical history (grade III or IV systemic reaction according to Mueller10 before VIT) and positive skin tests or specific immunoglobulin E. Exclusion criteria were lack of consent, pregnancy, severe chronic or/and malignant disease, or mastocytosis. Patients started immunotherapy at the day ward, reaching 1/10 of the maintenance dose every 6 weeks for 3 to 5 years. The study was approved by the Medical Ethical Committee of the University Medical Center Groningen (METC 2008/340).

The following 3 groups of patients were included (Table I):

Group 1 included patients who did not experience a systemic reaction in spite of being stung at least 3 times with the relevant insect after stopping VIT (n = 17). There were 9 (53%) men and 8 (47%) women, with a mean age of 53 years (range, 28-70) in this group.

Group 2 included patients who experienced at least 2 systemic reactions after field re-stings with the relevant insect (n = 12). There were 4 (33%) men and 8 (67%) women, with a mean age of 56 years (range, 42-75) in this group. The severity of the reaction to the re-sting was assessed as grade III in 80% (before VIT, 58%) and grade IV in 20% (before VIT, 42%) of patients according to the Mueller10 scale. The restart of venom immunotherapy was offered to all patients from this group.

Group 3 included patients who were still in the maintenance phase of VIT (3-5 years) and had not been stung since the start of the therapy (n = 17). There were 6 (35%) men and 11 (65%) women, with a mean age of 55 years (range, 21-75) in this group.

Collection of blood samples
From all patients, RNA was isolated from the whole blood by using the PAXgene Blood RNA Tubes (Qiagen, Valencia, Calif). All tubes were immediately frozen and stored at −20°C until RNA isolation (maximum period, 2 months). RNA was isolated by using the PAXgene Blood RNA Kit CE (Qiagen, Venlo, The Netherlands). All RNA samples were stored at −80°C until labeling and hybridization.

The quality and concentration of RNA were determined by using the 2100 Bioanalyzer (Agilent, Amstelveen, The Netherlands) with the Agilent RNA 6000 Nano Kit. Samples with a RNA integrity number >7.5 were used for further analysis on expression arrays.

Gene expression
For amplification and labeling of RNA the Illumina TotalPrep 96 RNA Amplification Kit was used (Applied Biosystems, Nieuwerkerk ad IJssel, The Netherlands). For each sample, we used 200 ng RNA. The Human HT-12_V3 expression arrays (Illumina, San Diego, Calif) were processed according to the manufacturer’s protocol. Slides were scanned immediately using an Illumina BeadStation iScan (Illumina).

Image and data analysis
First line check, background correction and quantile normalization of the data were performed with Genomestudio Gene Expression Analysis module v 1.0.6 Statistics. Entities containing at least 75% of samples with a signal intensity value above the 20th percentile in 100% of the samples in at least 2 groups were included for the further analysis.

Data analysis was performed by using the GeneSpring package version 8.0.0 (Agilent Technologies, Santa Clara, Calif). Genes for which expression was significantly different between compared groups were chosen based on a log2 fold change >2 in gene expression, t test P value <0.05 and a Benjamin-Hochberg false discovery rates <0.11,12 The naive Bayes prediction model was used to build a prediction model assessing the effectiveness of VIT.8,9 The naive Bayesian classifier is a mathematical process computing the probability of classifying the patient from group 3 as a treatment success or treatment failure based on the results of gene expression.8,9 The selection of genes and their influence on classification in a particular group is based on results obtained in groups 1 and 2. The naive Bayesian classifier assumes that the impact of single gene expression is unrelated to other genes in the prediction model. The method does not take into account the interactions of the genes composing the model or gene-environmental interactions.

Functional annotation of genes was described by using the Go Process analysis tool. Biological processes were obtained from the GO database. An enrichment P value <0.05 was considered significant. The GeneCards functional annotation web based tool10,11 was used to find the corresponding PubMed articles.

Clinical data for this study were analyzed with Statistica 8.0 (StatSoft, Tulsa, Okla).

RESULTS
Whole genome gene expression analysis was performed on RNA samples isolated from all blood cells in whole blood of 46 patients with IVA treated with VIT. From all 48,804 probes
present on the array, 48,071 transcripts had sufficient data for further analysis.

**Comparison of gene expression profiles between patients with long-term protection of VIT versus failure of treatment**

Of the analyzed transcripts, 1401 showed at log2 >2 fold difference in gene expression ($P < .05$), of which 658 genes (47%) were upregulated and 743 genes (53%) were downregulated. Significant differences ($P < .05$) in single gene expression were found for 978 transcripts. Correction for multiple testing reduced the number of significantly expressed genes to 43, of which 12 (28%) were upregulated and 31 (72%) downregulated. We identified a group of 18 genes with the most discriminative change in gene expression that fulfilled the following conditions: (1) log2 >3 fold change in gene expression, (2) $P < .0015$, (3) confirmed by correction for multiple testing with a $P < .005$. These genes were used to build the prediction model (Table II). A hierarchical dendrogram of those genes is presented in Fig 1.

**Functional annotation of genes differentially expressed**

Functional annotation of 978 genes with log2 >2 fold change and a significant difference ($P < .05$) was assigned by GeneCodis$^{16,17}$ (Table III). The main functions of the differently expressed genes were signal transduction, ion transport, multicellular organism development, transcription, cell proliferation, cell-cell signaling, and cytoskeletal organization. The most important signaling transduction pathways identified were the FceRI signaling pathway, the mitogen-activated protein kinase (MAPK) signaling pathway, the Wnt signaling pathway, the Jak–signal transducer and activator of transcription (STAT) signaling pathway, and the calcium signaling pathway.

Among the 18 most differentially expressed transcripts used for the prediction model (Table II) were actin filament associated protein 1-like 1 (AFAP1L1), which is involved in intracellular signaling and is a constituent of the cytoskeleton; claudin 1 (CLDN1) and protocadherin β 10 (PCDHB10), which are involved in cell adhesion; the prolactin receptor (PRLR), which is involved in signal transduction; and transcription factor twist homolog 2 (TWIST2), which increases the expression of the anti-inflammatory cytokine IL-10, which in turn is related to the success of immunotherapy.$^{18}$ For a majority of the transcripts, the function is unknown.

We also analyzed the expression profiles of leukocyte-specific genes expressed in dendritic cells, B cells, effector memory T cells, mast cells, and basophils as described by Liu at al.$^{19}$ A statistically significant difference in expression between patients with long-term protection compared with the group of patients with failure of VIT was found for the mast cell–specific gene follistatin (FST; $P = .003$), the memory T-cell gene galactokinin (GALK; $P = .008$), and B-cell specific Fc receptor-like 5 (FCRL5; $P = .04$).

Comparing our data set with the set of genes as described by Konno et al.$^{20}$ in a similar group of patients treated with bee venom allergy demonstrated statistically significant differences ($P = .04$) in gene expression for IL-1 receptor (IL1R1) and IL-1 receptor antagonist (IL1RN).

**Prediction of the outcome of treatment in a group of patients in the maintenance phase of VIT**

We subsequently predicted the potential outcome of treatment in patients still treated with VIT (group 3). We built 3 prediction models by using a naive Bayes$^{8,9}$ classifier based on (1) the 978 genes differentially expressed between the groups with failure and success of VIT (log2 fc > 2; $P < .05$), (2) the 56 genes with log2 fc >3 and $P < .05$ withstanding multiple test correction, and (3) the most discriminative 18 genes with $P < .0015$ withstanding multiple testing $P < .005$ (Table II). Because the 3 prediction models gave the same results, the one based on the lowest number of genes was used for further analysis. We were interested how this model would predict the percentage of failure of VIT in patients still in the maintenance phase of VIT. Of this patient group, according to this model, 2 (12%) would have treatment failure of VIT, whereas 15 (88%) would be protected. These percentages are in accordance with the known data of protection,$^{1-3}$ which might substantiate the usefulness of this model in the future.

**DISCUSSION**

In this study, we have shown that there is a gene expression profile that may help differentiate patients with success from those with failure after VIT. The differences in gene expression are related to known mechanisms of T-lymphocyte differentiation.
and mast cell activation, but probably also to other, yet unknown mechanisms. For this study, we used the RNA isolated from the whole blood. This not only is a simple and standardized method that can be used in a routine setting but also reduces the effect of sample handling, thereby making it an easy tool for clinical diagnosis.

Clinical relevance of the results

We were able to identify a gene expression pattern that is characteristic for the success of VIT in 100% of the patients with success of VIT, whereas it was not found in the patients with failure of VIT. We built 3 prediction models based on 978, 56, and 18 genes with the same results in all models. Therefore, we concluded that the number of genes in the prediction model may be reduced to 18. Subsequently we used the final prediction model to see whether this model gives realistic percentages in the group of patients who are yet in the maintenance phase of VIT. The gene profile characteristic for the success of treatment was present in 88% of patients on maintenance treatment of VIT, which is in agreement with the epidemiologic data on the risk of a reaction to a re-sting in this group.4 It will be necessary to follow these patients over time to test the true predictive value of our model, because this is necessary before applying such a model in clinical practice. It is planned to perform sting challenges before the end of VIT to ensure the final outcome of VIT.

A potential selection bias has to be taken into account, because the patient group with failure of treatment was selected on the basis of the medical history given by the patient and available medical records. In cases in which the data were not clear, the general physicians were contacted, and the data in medical records were compared with the anamnesis. In spite of these efforts, it is possible that the severity of the reaction in some patients in fact was different from the classified one. The anamnesis of patients with successful treatment is more reliable, because at least 3 re-stings were observed without side effects, although they could have been stung by a not relevant insect. The current data suggest that patients with successful treatment is more reliable, because at least 3 re-stings were observed without side effects, although they could have been stung by a not relevant insect. The current data suggest that patients with the reaction assessed as III and IV in the Mueller scale should be treated for 5 years, which may increase the effectiveness of VIT. The duration of treatment of the patients described in this study was shorter (3 years) but did not differ between the groups. Therefore, our results should be repeated in independent patient groups to evaluate the validity of the model.

The main question for further studies is whether we can use the gene expression analysis in daily practice. The severity of the reaction to a re-sting not only may depend on intrinsic patient factors but also may be related to the stinging insect, patient comorbidities and condition, and used medication. It is likely that an optimal diagnostic tool should include these factors as well as gene expression profiling. The definitive prediction of the outcome will always be difficult. Prospective studies in larger groups of patients treated in different centers should be performed to evaluate the accuracy of this gene profile. In the future, the analysis of gene expression profiles might also be used for the
TABLE III. Gene co-occurrence annotation found by Genecodis$^{14,15}$ (GO Process) for the genes differentially expressed (FC > 2; $P < .05$) between groups with success and failure of VIT

<table>
<thead>
<tr>
<th>No. of genes</th>
<th>NGR (37435)</th>
<th>NG (3434)</th>
<th>Hyp</th>
<th>Hyp*</th>
<th>Annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>1700</td>
<td>52(434)</td>
<td>2.30e-10</td>
<td>4.37e-09</td>
<td>GO:0007165: signal transduction</td>
</tr>
<tr>
<td>21</td>
<td>503(37435)</td>
<td>21(434)</td>
<td>5.76e-07</td>
<td>5.47e-06</td>
<td>GO:0006811: ion transport</td>
</tr>
<tr>
<td>28</td>
<td>874(37435)</td>
<td>28(434)</td>
<td>1.64e-06</td>
<td>1.04e-05</td>
<td>GO:0007275: multicellular organismal development</td>
</tr>
</tbody>
</table>

Table continues with additional genes and annotations...

**Immunological mechanisms that might be involved in the long-term effectiveness or ineffectiveness of VIT**

Functional annotation of the genes differentially expressed between patient groups with success and failure of VIT included genes involved in well known mechanisms of immunotherapy, such as FcεRI, JAK-STAT, MAPK, and Wnt, and calcium signaling pathways, cell signaling, or transcription. The function of many other differently expressed transcripts is yet unknown (Table II) or questionable, which is a common situation in whole gene expression studies.$^{21,22}$ Interestingly, genes commonly related to known mechanism of VIT like IL-10, IL-4, and osteopontin were not differentially expressed. This does not exclude significant differences in protein levels or differences in RNA expression in subpopulation of cells, like regulatory T lymphocytes, but they could not be demonstrated by the model chosen in the study examining the whole blood RNA. The description here therefore is based on the genes with known function and on the highest differences in expression composing the prediction model used in the study.

TWIST2 was upregulated in patients who gained success of VIT. It has been shown that this gene promotes the production of the IL-10 and decreases the synthesis of IL-4.$^{13,14}$ In spite of the fact that no difference in expression of IL-10 and IL-4 was observed in our study, the upregulation in TWIST2 expression may be responsible for the differences in cytokine levels and cell subtypes typical for immunotherapy.$^{18}$ Further studies are needed to address this finding.

The downregulation of PRLR in successfully treated patients may also indicate a shift toward T_{H}2. A decrease in serum levels of prolactin is found in patients during sublingual immunotherapy.$^{23}$ Prolactin induces overexpression of γδ T-cell receptor, which increases the IL-4–dependent IgE and IgG1 response essential for the development of T_{H}2 lymphocytes.$^{23}$ The downregulation of PRLR is consistent with this finding.

CLDN1 expression was higher in patients who were protected from re-sting reactions after VIT. This protein is a crucial structural component of tight junctions and plays an important role in adhesion and migration of dendritic cells.$^{24}$ The expression of CLDN1 is increased by TGF-β. Higher expression of CLDN1 in dendritic cells may be related to the role of these cells in regulatory T differentiation.$^{18}$

The function of solute carrier family 16 (SLC16A4), sh3 and px domain containing 3 (SNX33), and MCT5 is known although it is difficult to relate it to the mechanism of immunotherapy.$^{25}$ SNX33 product—sorting nexin 33—modulates endocytosis trafficking.$^{26}$ The function of all other transcripts (AFAP1L1, C16orf13, HS.129980, HS.205446, HS.21177, HS.428102, HS.532515, HS.581554, HS.583392, LOC644019, SLC47A2) contains 8 (COMMD8) gene may play a role in cell proliferation.$^{27}$ The function of all other transcripts (AFAP1L1, C16orf13, HS.129980, HS.205446, HS.21177, HS.428102, HS.532515, HS.581554, HS.583392, LOC644019, SLC47A2) contains 8 (COMMD8) gene may play a role in cell proliferation.$^{27}$
is unknown. The functional studies of the genes described and studies indicating the change in gene expression during immunotherapy may explain the mechanisms of venom immunotherapy in the future.

In conclusion, the use of gene expression profiles might be a useful tool to predict the effectiveness of VIT. The analysis of differentially expressed genes confirms the involvement of immunologic pathways described before but also indicates novel pathways potentially involved in induction of allergen tolerance. Further studies in larger groups of patients are required to confirm this prediction model before it can be used in clinical practice.

Clinical implications: Gene expression profiles may help identify patients who fail to achieve longtime protection by insect venom immunotherapy. The results of this study may be a basis for further studies.

REFERENCES