



OPEN

## *S100a9* might act as a modulator of the Toll-like receptor 4 transduction pathway in chronic rhinosinusitis with nasal polyps

Nasibeh Khayer<sup>1</sup>✉, Maryam Jalessi<sup>1,2</sup>✉, Mohammad Farhadi<sup>2</sup> & Zahra Azad<sup>1</sup>

Chronic rhinosinusitis with nasal polyp (CRSwNP) is a highly prevalent disorder characterized by persistent nasal and sinus mucosa inflammation. Despite significant morbidity and decreased quality of life, there are limited effective treatment options for such a disease. Therefore, identifying causal genes and dysregulated pathways paves the way for novel therapeutic interventions. In the current study, a three-way interaction approach was used to detect dynamic co-expression interactions involved in CRSwNP. In this approach, the internal evolution of the co-expression relation between a pair of genes (X, Y) was captured under a change in the expression profile of a third gene (Z), named the switch gene. Subsequently, the biological relevancy of the statistically significant triplets was confirmed using both gene set enrichment analysis and gene regulatory network reconstruction. Finally, the importance of identified switch genes was confirmed using a random forest model. The results suggested four dysregulated pathways in CRSwNP, including “positive regulation of intracellular signal transduction”, “arachidonic acid metabolic process”, “spermatogenesis” and “negative regulation of cellular protein metabolic process”. Additionally, the *S100a9* as a switch gene together with the gene pair {*Cd14*, *Tpd52l1*} form a biologically relevant triplet. More specifically, we suggested that *S100a9* might act as a potential upstream modulator in toll-like receptor 4 transduction pathway in the major CRSwNP pathologies.

Chronic rhinosinusitis (CRS) is a highly prevalent disorder that affects approximately 30% of adults. It is characterized by persistent inflammation of the nasal and sinus mucosa<sup>1,2</sup>. CRS is a heterogeneous disease that can be classified into two major subgroups with distinct pathophysiologies: CRS with nasal polyps (CRSwNP) and CRS without nasal polyps (CRSsNP)<sup>1</sup>. It is worth noting that up to 60% of individuals with CRSwNP also experience concurrent lower airway diseases, such as asthma, which typically onset in adulthood<sup>3</sup>. The main therapeutic options for CRSwNP are currently limited to intranasal glucocorticoid administration, sinus surgery, or a combination of both<sup>4</sup>.

Despite the significant morbidity and negative impact on quality of life, our understanding of the underlying molecular mechanisms and specific biomarkers associated with CRSwNP remains incomplete<sup>5</sup>. Therefore, identifying causal genes and dysregulated pathways paves the way for novel therapeutic interventions.

Advances in high-throughput gene expression profiling technologies, such as microarray and RNA sequencing, have provided unprecedented opportunities to produce disease-related transcriptome datasets<sup>6</sup>. Such datasets provide a genome-scale snapshot of gene expression profiles, serving as invaluable resources for deducing gene interactions. Moreover, the identification of disease-related gene interaction plays a pivotal role in unraveling the intricate molecular mechanisms as well as identifying the critical genes in a specific disease. Furthermore, the conclusions drawn from the same biological data can vary based on the computational approach employed<sup>7,8</sup>.

In a few previous studies, transcriptome datasets have been investigated to identify the pathogenicity mechanisms and potential drug targets in CRSwNP. Recently, Peng et al. performed a comprehensive genome-wide gene expression analysis in order to elucidate the pathways and candidate gene sets associated with CRSwNP. They investigated the differentially expressed genes and corresponding pathways in a CRSwNP-related dataset. Their results suggested several biological pathways involved in CRSwNP, including “defective host defences”,

<sup>1</sup>Skull Base Research Center, The Five Senses Health Institute, School of Medicine, Iran University of Medical Sciences, Tehran, Iran. <sup>2</sup>ENT and Head and Neck Research Center and Department, The Five Senses Health Institute, Rasoul Akram Hospital, School of Medicine, Iran University of Medical Sciences, Tehran, Iran. ✉email: khayer.n@iums.ac.ir; dr.jalessi@gmail.com

“inflammation” and “abnormal metabolism of extracellular matrix”<sup>9</sup>. In another recent study, Hao et al. employed an integrated analysis approach to investigate the dysregulated pathway as well as the crucial genes in CRSwNP from a diagnostic and therapeutic perspective. Their results indicated several CRSwNP-related pathways, including “immune effector process”, “leukocyte migration”, “regulation of the inflammatory response”, “*Staphylococcus aureus* infection”, and “cytokine-cytokine receptor interaction”. Moreover, they suggested *Alox5ap*, *Bcl2a1*, *Btk*, *Cybb*, *Ncf2*, *Hck*, and *Hk3* genes, which significantly increase in CRSwNP patients, as potentially crucial genes in the pathogenesis<sup>10</sup>.

A predominant feature of gene interactions is often highly dynamic, invariably linked to alterations in cellular conditions as a reaction to diverse external stimuli and signals<sup>11</sup>. In other words, a pair of genes participating in a comparable biological pathway during a particular biological condition might lose their connection under a different biological circumstance. As a result, the degree and pattern of gene expression correlation between a pair of genes can be influenced by internal changes and cellular conditions<sup>12</sup>. Therefore, pairwise interaction analyses, in the classical sense, may be too simplistic to explain complex molecular relationships<sup>13,14</sup>.

The three-way interaction approach describes the internal evolution of the co-expression relation between a pair of genes (X, Y)<sup>13,15</sup>. If it is supposed that a state change within a cell is associated with the expression levels of a third gene, say Z, the expression profile of Z can capture the dynamic nature of the co-expression relation of two genes X and Y. Indeed, the Z gene as a surrogate for the intrinsic-state variable control the evolution of the co-expression relation of X and Y genes. Therefore, such an approach possesses the capability to decipher sophisticated molecular relations at an elevated level of complexity compared to the classical pairwise interaction approach<sup>16,17</sup>. From now on, such triplet will be shown as Z/{X, Y}.

So far, the three-way interaction approach has been employed to deciphering the pathogenesis of several diseases, including inflammatory diseases<sup>18,19</sup>. The importance of *RT-DOb*/{*Csf1r*, *Milr1*} triplet in onset of Alzheimer’s disease has been suggested and validated in previous studies. These studies indicated the change in expression level of *RT-DOb* gene, as a switch gene, can disturb the communication between mast cell and microglia<sup>19,20</sup>. Another study suggested the *Rps27a* gene can act as a switch gene for {*Il-18*, *Cx3cl1*} gene pair and be associated with pathogenesis of the major neurodegenerative diseases by controlling the microglia activation<sup>21</sup>.

The objective of the present study is a comprehensive analysis of CRSwNP-related transcriptome data, aiming to shed light on dysregulated molecular mechanisms as well as to trace the critical therapeutic targets in the CRSwNP. For such purpose, we used a novel computational method, i.e., liquid association (LA), as a powerful tool to capture dynamic co-expression relationships<sup>15</sup>.

## Results

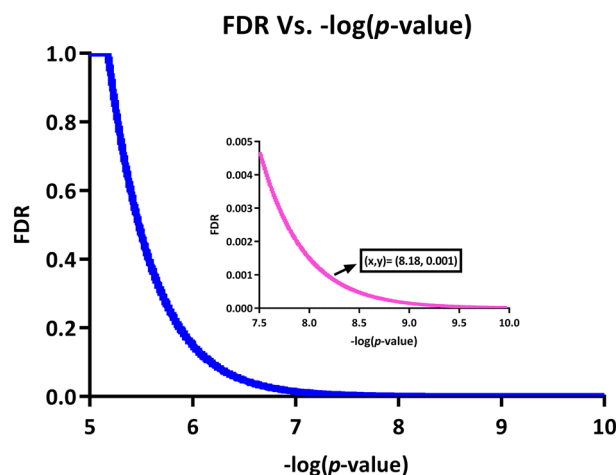
### Statistically significant three-way interaction determination

The liquid association measure was computed for every possible Z/{X, Y} combination in the dataset. The top 200,000 triplet combinations with the highest significance based on *p*-value were selected as primary outputs.

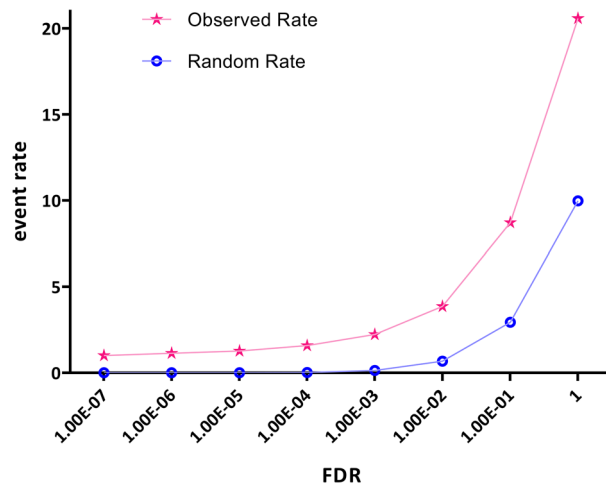
A *p*-value histogram of such triplet combinations is presented in Supplementary Fig. S1 online. Moreover, Fig. 1 demonstrates a downward trend of adjusted *p*-values versus  $-\log(p\text{-value})$ .

In order to assess the accuracy of liquid association analysis, the observed event rate of Z position (switch) genes was compared with the random event rate across a broad range of significant fastLA *p*-values. Figure 2 presents this comparison, which contributes to assessing the analysis’s reliability.

A subset of triplet combinations was chosen for further analysis based on an adjusted *p*-value threshold of less than  $10^{-3}$  (Fig. 1A). This subset consisted of 807 statistically significant triples. An explanation about the chosen cut-off point in this analysis has been presented in the Supplementary Text S1 online. Moreover, the complete list of all statistically significant triplets can be found in the Supplementary Table S1 online.



**Figure 1.** FDR vs.  $-\log(p\text{-value})$ . The changes in FDR (BH-corrected *p*-value) versus  $-\log(p\text{-value})$  for the first 200,000 results of fastLA. As shown, FDR = 0.001 corresponds to  $-\log(p\text{-value}) = 8.18$ .



**Figure 2.** Evaluation of the fastLA analysis accuracy. This figure presents an assessment of the accuracy of the fastLA analysis. The study investigates the relationship between significant fastLA  $p$ -values and the observed event rate of genes located at position Z (switch) in comparison to the random event rate. The findings unequivocally show a substantial deviation between the observed event rate of switch genes and the random rate, providing strong evidence to support the accuracy of the fastLA analysis.

### Biologically relevant three-way interaction identification

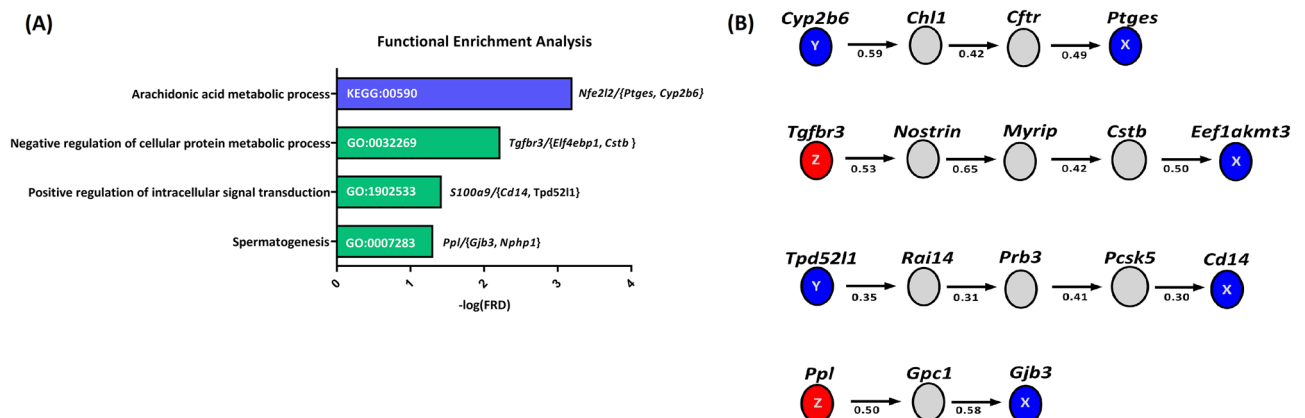
GSEA was conducted as the first step to identify biologically relevant triplets. This analysis focused on the 807 statistically significant triplets. Because of the generality of terms in lower levels of gene ontology, only terms in levels higher than 6 were included. The complete list of enriched terms by considering BH-corrected  $p$ -value  $< 0.05$  is available in the Supplementary Table S2 online.

Consistent with the three-way interaction model concept, it is expected that in a biologically-relevant triplet, the X and Y genes would be involved in the same biological process. In light of this, the analysis focused on identifying statistically significant triplets and examining their association with enriched terms to ascertain their biological relevance.

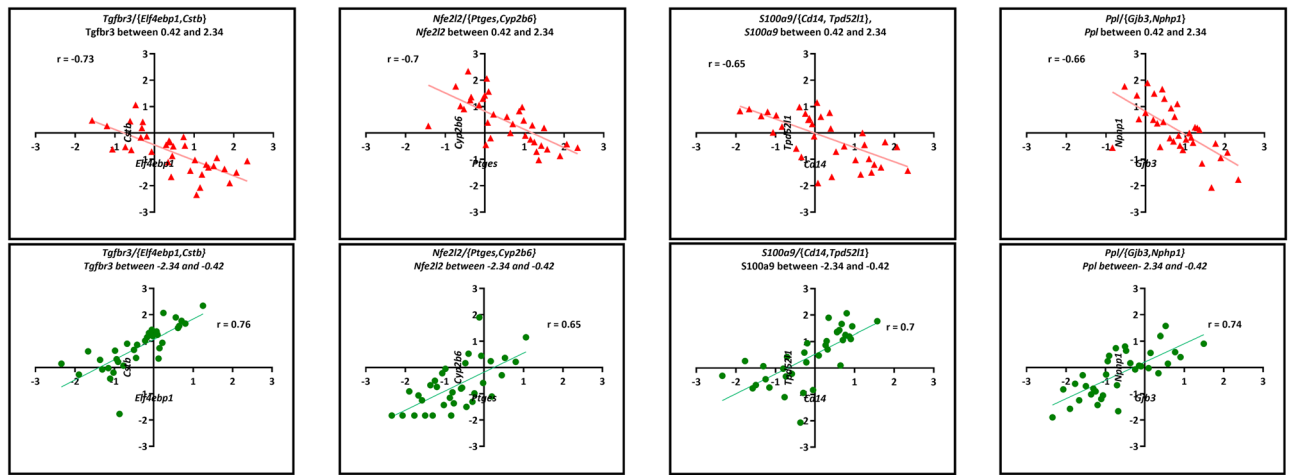
As the next step in uncovering biologically relevant triplets, the reconstruction of the GRN was performed using the ARACNE algorithm. Then, the regulatory relationships involving statistically significant triplets were investigated while disregarding any relationships with more than three mediators.

Collectively, using both GSEA and GRN analyses, the biological relevance of four statistically significant triplets was verified. These triplets consist of *S100a9*/*Cd14*, *Tpd52l1* triplet, *Nfe2l2*/*Ptges*, *Cyp2b6* triplet, *Ppl*/*Gjb3*, *Nphp1* triplet, and *Tgfb3*, *Elf4ebp1*, *Cstb* triplet, which are involved in the biological processes of “positive regulation of intracellular signal transduction”, “arachidonic acid metabolic process”, “spermatogenesis” and “negative regulation of cellular protein metabolic process”, respectively (Fig. 3).

Moreover, the scatter plots of these triplets are presented in Fig. 4, depicting a notable correlation change between X and Y in response to variations in Z expression levels within three different ranges.



**Figure 3.** The biological relevance of four statistically significant triplets. (A) The enriched biological process and pathway terms corresponding with four biologically relevant triplets have been presented in this figure. (B) This figure illustrates the positioning of four biologically relevant triplets within the Gene Regulatory Network (GRN). The red nodes represent the genes located at the Z position within each triplet, while the green nodes represent the genes located at the X and Y positions. Additionally, other genes are depicted as grey nodes.



X	Y	Z	MLA value	Wald's measure	p-value	Corrected p-value	r <sub>low</sub>	r <sub>high</sub>	P <sub>diff</sub>
<i>Elf4ebp1</i>	<i>Cstb</i>	<i>Tgfb3</i>	-0.53	35.46	2.59848E-09	3.90E-04	0.76	-0.73	-1.49
<i>Ptges</i>	<i>Cyp2b6</i>	<i>Nfe2l2</i>	-0.49	34.93	3.41476E-09	5.12E-04	0.65	-0.7	-1.35
<i>Cd14</i>	<i>Tpd52l1</i>	<i>S100a9</i>	-0.48	34.54	4.17839E-09	6.27E-04	0.7	-0.65	-1.35
<i>Gjb3</i>	<i>Nphp1</i>	<i>Ppl</i>	-0.5	33.91	5.78519E-09	8.69E-04	0.74	-0.66	-1.4

**Figure 4.** Scatter plots of four triplets that are biologically relevant. There is a notable change in the correlation between X and Y due to a change in the expression level of Z. The observed alterations emphasize the significant influence of Z expression on the relationship between X and Y, highlighting its biological relevance within the context of the studied system.

### Random forest and selecting important genes

The random forest algorithm offers various metrics for assessing variable importance, among which Mean Decrease Accuracy (MDA) is considered the most reliable. MDA determines importance by evaluating the reduction in classification accuracy when the expression values of a particular gene are randomly permuted<sup>18,19</sup>. In our study, we utilized MDA to identify the important switch genes, as presented in Fig. 5A. The results showed that *Tgfb3*, *Tob1* and *S100a9* genes are more important in classifying such groups.

Furthermore, the area under the receiver operating characteristic (ROC) curve (AUC) is widely employed to evaluate the performance of supervised classification models<sup>20</sup>. In our analysis, we utilized ROC curves to assess the sensitivity and specificity of the Random Forest model. The results, illustrated in Supplementary Fig. S3 online, indicate an AUC of 0.87, a sensitivity of 95, and a specificity of 46 for the classifier.

The ANOVA results reveal significant differences in the gene expression profiles of three specific genes (*Tgfb3*, *Tob1*, and *S100a9*) across various groups (Supplementary Tables S4–S6). Tukey's HSD analysis further indicates that the expression levels of these genes significantly differ between the NP-NP group and the CS-IT group, as well as between the NP-IT group and the CS-IT group. Notably, the *S100a9* gene also demonstrates a significant difference between the NP-IT and CS-IT groups, as shown in Fig. 5B. Detailed gene expression profiles for these genes can be found in Supplementary Table S7.

### Identifying novel putative drugs

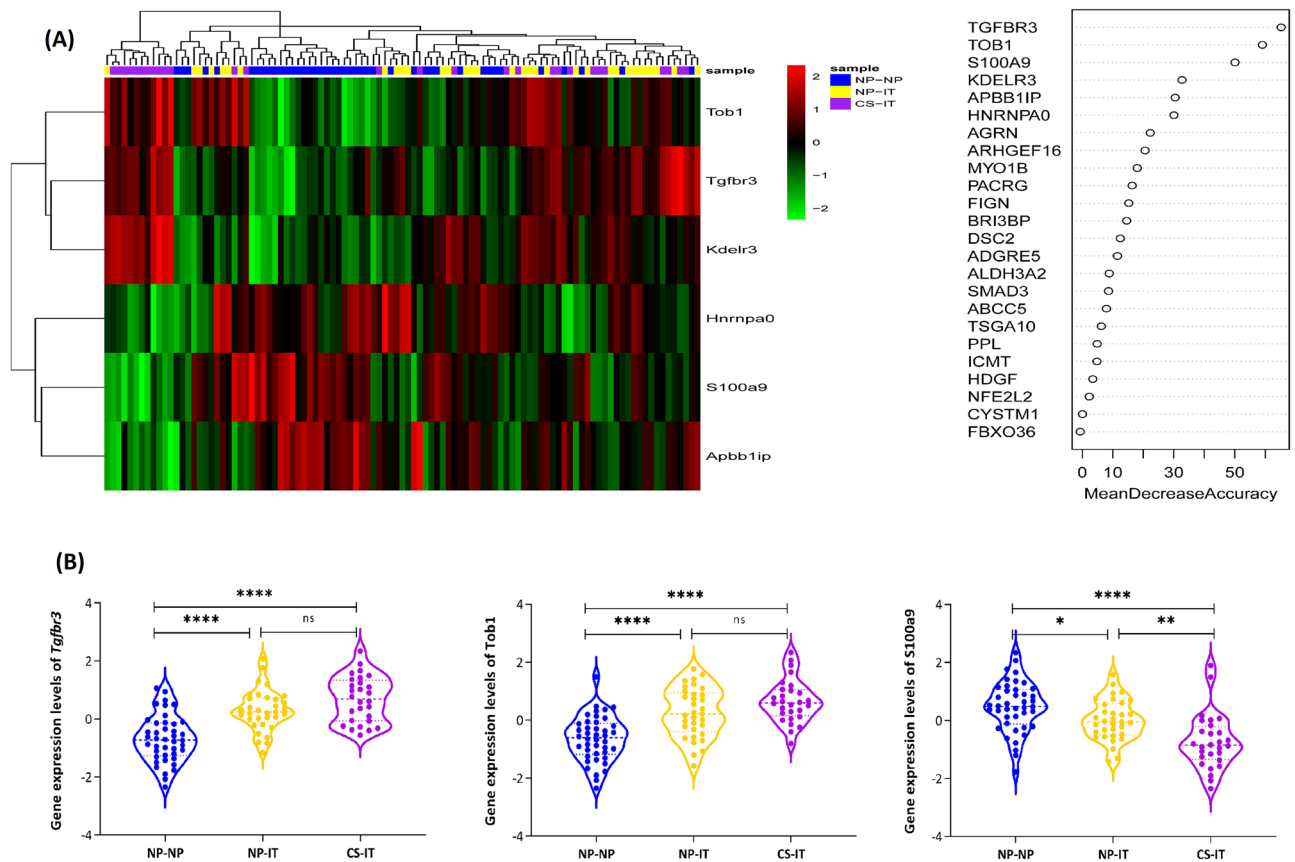
In a biologically relevant triplet, the Z gene is prone to be used as a drug target because this gene can control the intrinsic state changes associated with a disease. Hence, in order to determine novel potential drugs, four switch genes belonging to the biologically relevant triples, including *S100a9*, *nfe2l2*, *Ppl* and *Tgfb3*, were explored in drug-related databases.

The results identified two drugs, namely “tasquinimod” and “paquinimod”, are associated with the *S100a9* genes and two corticosteroid drugs, namely “dexamethasone” and “regramostim”, are related to the *Tgfb3* gene.

### Discussion

Chronic rhinosinusitis (CRS) is an inflammatory disease severely affecting the nasal mucosa<sup>2</sup>. Nasal polyps (NP) are a common comorbidity, impacting around 30% of CRS patients, leading to significant socioeconomic burdens and reduced quality of life<sup>22</sup>. The clinical management of CRSwNP could be more efficacious by increasing the knowledge of the underlying pathogenic factors and subsequently entering effective therapeutic interventions<sup>4</sup>. Therefore, the LA analysis was employed in this study to capture critical triplets associated with CRSwNP pathogenesis. To our knowledge, the CRSwNP transcriptome data has yet to be investigated through the 3WI model.

The accuracy of the LA assay was verified by comparing the observed occurrence rate of Z position genes within a wide range of significant fastLA p-values, juxtaposed against a randomly determined occurrence rate. It is expected that there will be a notable decrease in the number of genes occupying the Z position compared to what would be expected by chance. This expectation arises from the fact that a relatively small set of genes governs the majority of biological processes. As shown in Fig. 2, the observed event rate for switch genes is markedly



**Figure 5.** (A) Random forest classification and selection of important genes. The genes were chosen based on the Mean Decrease Accuracy measure, which serves as a reliable metric for determining their significance in the classification model. The 6 top genes that exhibit the highest importance in contributing to the accuracy of the random forest classification have been presented. (B) Results of Tukey's Honest Significant Difference (HSD) analysis. This bar graph illustrates the outcomes of the Tukey's HSD test on the gene expression levels of *Tgfb3*, *Tob1*, and *S100a9* across three experimental groups: NP-NP, CS-IT, and NP-IT. Asterisks above the error bars indicate significant differences between group means as determined by Tukey's HSD test, with significance levels marked as follows: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ . The gene expression data for each group is represented in distinct colors to enhance visual distinction.

far from what would be expected by chance. The results showed that specific genes predominantly occupy the Z positions within statistically significant triplets.

### The relationship between enriched pathways and CRSwNP

The pathway enrichment analysis results (Fig. 3) show that four pathways are involved in CRSwNP. These pathways include "positive regulation of intracellular signal transduction", "arachidonic acid metabolic process", "spermatogenesis", and "negative regulation of cellular protein metabolic process". In the following, the relationship between these pathways and CRSwNP is explained in detail. See below.

The "intracellular signal transduction" plays a crucial role in the regulation of cellular processes such as proliferation, differentiation, and apoptosis<sup>23</sup>. Dysregulation of intracellular signaling pathways has been implicated in the pathogenesis of several diseases, including chronic rhinosinusitis with nasal polyps (CRSwNP)<sup>24–26</sup>. Several studies have suggested the role of intracellular signal transduction pathway in the development and progression of CRSwNP. Spleen tyrosine kinase (Syk) plays a significant role in intracellular signal transduction in various types of hematopoietic cells<sup>27</sup>. According to a study, the expression of the Syk gene is increased in patients with allergic nasal polyps<sup>28</sup>. Another study suggested that the inhibition of the Syk gene prevents mast cell activation, and therefore, it may be an important therapeutic strategy for the treatment of allergic nasal polyps<sup>29</sup>. Furthermore, the Wnt signaling pathway is another intracellular signal transduction that has been suggested to contribute in the pathogenesis of NPs. A recent study found that the expression levels of some Wnt signaling pathway-related genes are dysregulated in CRSwNPs, and also such a pathway is activated<sup>30</sup>. Another study suggested that the Wnt signaling pathway may contribute to the pathogenesis of NPs through epithelial-to-mesenchymal transition<sup>24</sup>. In addition, some studies have suggested that p38 Mitogen-activated protein kinase (MAPK) and Jun N-terminal kinase (JNK) signal transduction pathways are involved in CRSwNP by mediating the expression of glucocorticoid receptor isoforms<sup>36</sup>. Finally, MUC1, a transmembrane protein that plays a role in intracellular signaling, acts as an anti-inflammatory molecule in some airway infections and mediates the



expression of anti-inflammatory genes in CRS<sup>31</sup>. Taken together, the role of “intracellular signal transduction” in the pathogenesis of CRSwNPs is sensible.

The “arachidonic acid” pathway is an intricate metabolic pathway that constitutes a complex biochemical cascade resulting in the synthesis of pro-inflammatory eicosanoids, including prostaglandins, thromboxanes, and leukotrienes<sup>32,33</sup>. Arachidonic acid and its metabolites have received considerable interest in the context of the pathogenesis of various inflammatory diseases, including CRSwNP<sup>34–36</sup>. According to a recent study, the severity of chronic rhinosinusitis (CRS) is associated with the levels of eicosanoids, arachidonic acid-derived lipid mediators, in nasal polyp secretions<sup>37</sup>. Another study has indicated that the expression levels of CXCL1 and CXCL8 genes, which are regulated by thromboxane A2, are up-regulated in the CRSwNP mucosa compared to controls<sup>38</sup>. Additionally, a previous study suggested that prostaglandin E2 was elevated in CRSwNP patients, especially in polyps from aspirin-sensitive individuals<sup>36</sup>. Therefore, substantial evidence suggests that the “arachidonic acid” pathway may well contribute to the pathogenesis of CRSwNP.

Although it may not seem likely at first glance, there is an association between the “spermatogenesis” pathway and nasal polyp pathogenesis; previous studies have established such a relationship. Young’s syndrome, also named sinusitis-infertility syndrome, is a rare and inherited syndrome characterized by chronic sinusitis, bronchiectasis, and decreased fertility due to azoospermia. Various studies have suggested that mutation in the *Cfr* gene may be involved in the pathogenesis of CRS as well as Young’s syndrome<sup>39–42</sup>. Furthermore, a previous study suggested that the up-regulation of a member of matrix metalloproteinase, namely disintegrin and metalloproteinase (ADAM)-33 protein, is associated with the pathogenesis of nasal polyps<sup>43</sup>. On the other hand, another study has indicated that such protein is expressed in spermatogenesis cells and plays a central role during sperm maturation<sup>44,45</sup>. Overall, the above studies demonstrate the existence of common genes in spermatogenesis process and nasal polyp pathogenesis.

The “protein metabolic processes” refer to the various pathways involved in synthesizing, breaking down, and modifying proteins. These processes are tightly regulated to ensure proper protein homeostasis and cellular function. Evidence suggests that “protein metabolic processes” may be disrupted in nasal polyp disease, leading to abnormal protein synthesis and degradation. A transcriptome study of nasal epithelial cells from patients with Anti-Inflammatory Drug-Exacerbated Respiratory Disease (AERD) indicated that the leukotrienes metabolic process actively participates in such disease<sup>46</sup>. Furthermore, proteome studies of nasal mucosa from healthy and CRS patients have shown that the disruption of, either positive or negative, regulation of cellular protein metabolic processes can be involved in the pathogenesis of CRS<sup>47,48</sup>.

The biological relevancy of the four triplets was confirmed using both GSEA and GRN (see Fig. 3). Moreover, the importance of two switch genes, *S100a9* and *Tgfb3*, belonging to such triplets was confirmed using Random Forest Analysis (see Fig. 5). Accordingly, it is suggested that these triplets may play a central role in CRSwNP.

We investigated the biological relationship among genes involved in above mentioned triplets in the literature. A biological relevancy in the *S100a9*{*Cd14*, *Tpd521*} triplet was found. We discuss this triplets in more detail.

### The relationships among involved genes in *S100a9*{*Cd14*, *Tpd521*} triplet

The *S100a9*{*Cd14*, *Tpd521*} triplet was selected to be explained in more detail because it includes several features that distinguish it from other triplets. First, this triple is not only statistically significant but also is biologically relevant based on both GSEA and GRN. Second, according to Random Forest analysis, the expression of *S100a9* gene is significantly different in NPs patients (Fig. 5). Finally, we could find significant evidence in the literature to confirm such relations.

However, to the best of our knowledge, no major evidence found for three others suggested triplets in the literature, these triplets should be studied further.

In the following, we discussed the importance of *S100a9*, *Cd14* and *Tpd521* genes in CRSwNP. See below.

The *S100a9* gene encodes a calcium-binding protein called *S100a9*, a member of the S100 protein family. This gene is predominantly expressed by myeloid cells, including neutrophils and monocytes, and plays a role in the regulation of inflammation and immune responses<sup>49</sup>. Crombruggen and colleagues showed that the protein expression level of S100A9 was significantly increased in CRSwNP patients, resulting in increased deposition on extracellular matrix (ECM) structures of CRSwNP tissue compared to CRS without NP and controls. It suggested that the inflammatory/remodeling, as the major characteristic of CRSwNP, empowers the retention of S100A9 protein in the ECM CRSwNP tissue. In this way, the S100A9 protein acts as a local danger signal-inducing inflammatory mediator, predominantly via TLR-4 activation<sup>50</sup>. Brouk and colleagues showed that the expression level of *S100a9* is evaluated in CRSwNP at both mRNA and protein levels. Furthermore, they indicated that such a phenomenon coincides with increased matrix metalloproteinases (MMPs) production in CRSwNP, implying the contribution of *S100a9* and MMPs in elevated nasal cell proliferation<sup>51</sup>. Moreover, the up-regulation of *S100a9* expression in the nasal polyp tissues of patients with CRSwNP compared to the inferior turbinate tissue of both healthy control and CRS without NP patients was reported in recent studies<sup>52,53</sup>.

The *Cd14* gene encodes a glycosylphosphatidylinositol-anchored membrane protein expressed on various immune cells’ surfaces, including monocytes, macrophages, and dendritic cells. The *Cd14* acts as a co-receptor for lipopolysaccharide (LPS), a component of the outer membrane of Gram-negative bacteria. Its principal function is facilitating the recognition and binding of LPS by Toll-like receptor 4 (TLR4), initiating an inflammatory response. Several studies have investigated the role of the *Cd14* gene in CRSwNP pathogenesis. Yazdani and colleagues investigated genetic variations in the *Cd14* gene and their association with CRSwNP susceptibility. They suggested that specific single nucleotide polymorphisms (SNPs) in the *Cd14* promoter gene might be associated with NP pathogenesis and asthma incidence<sup>17</sup>. Yao and colleagues suggested that *Cd14*, as an inflammatory mediator, is significantly up-regulated in noneosinophilic CRSwNP as compared to eosinophilic CRSwNP and controls<sup>54</sup>. These findings could provide further insights regarding the diverse nature of CRSwNP. Furthermore,

CD14 is known as a marker for monocyte-derived dendritic cells (moDCs). O'Connell and colleagues demonstrated that CRSwNP is distinguished by a significant increase in circulating moDCs, associated with systemic inflammation exhibiting a Th2 bias as well as mucosal inflammation<sup>55</sup>.

*Tpd5211*, also known as Tumor protein D52-like 1, is a member of the TPD52 family of proteins. It has been implicated in various cellular processes, including cell proliferation, apoptosis, and vesicular trafficking<sup>56</sup>. The role of *Tpd5211* has been extensively studied in breast cancer<sup>57</sup>, childhood leukemia<sup>58</sup>, colorectal cancer<sup>59</sup> and recurrent implantation failure<sup>60</sup>, while its role in CRSwNP pathogenesis has not been reported yet.

In the current study, the expression level of *S100a9*, as a switch gene of *S100a9*{*Cd14*, *Tpd5211*}, was up-regulated in NP-NP samples (see Fig. 5), in agreement with previous studies mentioned earlier. Such a gene controls the co-expression relationship between *Cd14* and *Tpd5211* gene pairs. As shown in Fig. 4, when the normalized expression level of the *S100a9* gene is between  $-2.34$  and  $-0.42$  (as in most CS-IT samples (see Fig. 5)), there is a direct correlation between *Cd14* and *Tpd5211* expression levels ( $r_{\text{low}} = 0.70$ ). In contrast, when the normalized expression level of the *S100a9* gene is between  $0.42$  and  $2.34$  (as in most NP-NP samples (see Fig. 5)), there is an inverse correlation between *Cd14* and *Tpd5211* expression levels ( $r_{\text{high}} = -0.65$ ).

To our knowledge, the switch role of the *S100a9* gene for the *Cd14* and *Tpd5211* gene pair has not been reported. However, the coordination in the changes of expression level of such genes has been reported in previous studies. See below.

It has been reported that *Cd14* and *Tpd5211* genes are up-regulated in primary lung cancer compared to the adjacent normal tissue. Interestingly, under treatment with Romidepsin, expression levels of both these genes are down-regulated<sup>61</sup>. Similarly, the expression levels of such genes are down-regulated under treatment with Triglitzone in stomach cell cancers<sup>62</sup>, and Emodlin in Hepatocellular Carcinoma<sup>63</sup>. Furthermore, the up-regulation of gene expression levels of both such genes has been reported in some chronic diseases, including Type 2 Diabetes Mellitus<sup>64</sup> and chronic obstructive pulmonary disease<sup>65</sup>.

### On the role of toll-like receptor 4 signal transduction pathway in CRSwNPs

The results of GSEA suggested that the *S100a9*{*Cd14*, *Tpd5211*} triplet participates in the “positive regulation of intracellular signal transduction” pathway (Fig. 3A). On the other hand, an in-depth review of literature indicated that the *S100a9*, *Cd14* and *Tpd5211* genes may be involved in CRSwNP through the toll-like receptor 4 (TLR4) signal transduction pathway, which is surprisingly consistent with our results of GSEA. See below.

The term “Positive regulation of intracellular signal transduction” refers to the biological processes that enhance or amplify signal transmission from cell surface receptors to intracellular targets. This amplification facilitates specific biochemical cascades that alter cellular functions. For instance, the activation of TLR4 by its ligands, such as LPS, triggers a cascade of intracellular signaling that leads to the transcriptional activation of genes involved in immune and inflammatory responses<sup>66</sup>. Some central intracellular signaling pathways positively regulated by TLR4 include the NF- $\kappa$ B<sup>67</sup>, MAPK<sup>68</sup>, IRF3<sup>69</sup> and PI3K-Akt<sup>70</sup> pathways. Additionally, according to data in the Gene Ontology (GO) database, the term “positive regulation of intracellular signal transduction” [GO:1902533] is linked to “positive regulation of the toll signaling pathway” [GO:004572] through the “positive regulation of signal transduction” [GO:009967] term<sup>71</sup>.

In the following, we discuss the role of TLR4 in CRSwNP, and then the role of the *S100a9*, *Cd14* and *Tpd5211* genes in the TLR4 pathway.

Toll-like receptors (TLRs) represent transmembrane receptors characterized by two domains: an extracellular domain responsible for pathogen interaction, coupled with an intracellular signaling domain. These receptors play a central role in triggering inflammatory responses against invading microorganisms through detecting pathogen-associated molecular patterns (PAMPs) associated with foreign pathogens such as bacteria, viruses, and fungi. Specifically, TLR4 is known as a receptor for lipopolysaccharide (LPS), a PAMP found in the outer membrane of gram-negative bacteria as a main pathogenic factor in NP<sup>72</sup>. Some evidence suggest that the TLR4 signaling pathway is involved in the pathogenesis of CRSwNP by remodeling of nasal polyp. Cho and colleagues suggested that high expression levels of the TLR4 gene induce MAPK and PI3K/Akt signaling pathways, contributing to nasal polyps remodeling<sup>73</sup>. Other studies have reported that steroids inhibit the remodeling of airways in NP by decreasing the expression levels of TLR4 in both mRNA and protein levels. Indeed, expression levels of vascular endothelial growth factor (VEGF) as an airway remodeling factor in NPs can be efficiently inhibited via TLR4/Akt/NF- $\kappa$ B signaling pathway<sup>74,75</sup>.

Previous studies indicated that *S100a9* induces the NF- $\kappa$ B activation through the binding to the TLR4, which in turn triggers pro-inflammatory cytokine response in monocytes<sup>76–78</sup>. In addition, the protein levels of S100A9 is significantly increased in ECM structures of CRSwNP patients compared to CRS without NP and control ones. It has been suggested that extracellular S100A9 proteins induce the release of diverse inflammatory mediators through TLR-4 engagement<sup>50</sup>. Ehrchen et al. have been suggested that S100A8/ S100A9 complex can act as endogenous activators of TLR4 and promote lethal, endotoxin-induced shock<sup>79</sup>. Riva et al. indicated that *S100a9* acts as a TLR4 agonist and induces nuclear factor- $\kappa$ B responses<sup>76</sup>.

The above studies indicated that the *S100a9* gene has been shown to modulate TLR4 signaling in certain contexts. Interestingly, such a modulator role is in accordance with the concept of a switch gene.

It has been indicated that *Cd14* plays various functional roles in LPS-induced TLR4 activation. First, as the most important role, *Cd14* prompts the internalization of the LPS/TLR4 complex into endosomes<sup>80</sup>, where TLR4 engages cytosolic TRIF-TRAF3 signaling to stimulate IFN $\beta$  response<sup>81</sup>. Second, the *Cd14* binds and transfers LPS to the TLR4-MD2 complex as a long-established role, triggering the myddosome assembly and signal transduction<sup>82</sup>. Finally, the CD14 protein acts as an essential co-receptor for *S100a9*-mediated TLR4-stimulation<sup>83</sup>. Interestingly, the cooperation between *Cd14* and *S100a9* in the TLR4 signal transduction pathway is consistent with the results of the current study.

A recent study suggested that expression of *Tpd52* gene at the protein level is significantly increased in LPS-stimulated macrophage cells. In such a process, TLR4 is activated by LPS and induces the MyD88 pathway, which subsequently produces pro-inflammatory cytokines through activation of transcriptional nuclear factor (NF)- $\kappa$ B. Furthermore, it was indicated that the expression levels of *Tpd52* gene are significantly decreased upon treatment by statin as an anti-inflammatory agent<sup>84</sup>.

Although the above evidence emphasizes the role of *S100a9*, *Cd14* and *Tpd52l1* genes in the pathogenesis of CRSwNP through the TLR4 signal transduction pathway, a conclusive association is yet to be found between the expression level of such genes and the progression of CRSwNP.

The two investigational drugs, Paquinimod and Tasquinimod, are quinoline-3-carboxamide derivatives known for their potential to inhibit *S100a9* biological function<sup>85</sup>. Although their mechanisms of such an action have not yet been fully understood, an explanation is that these drugs can directly bind to *S100a9*, preventing formation of *S100a9*–*S100a8* complex, and its further interaction with TLR4 and the receptor for advanced glycation end-products<sup>86,87</sup>. The blockage of *S100a9*-mediated pathways by these drugs leads to several immune modulatory and tumor microenvironment impairing effects; the reduced production of the key pro-inflammatory cytokines and chemokines for immune cell activation and recruitment to inflammation or tumor growth sites<sup>88</sup>, anti-angiogenic effects (caused by Tasquinimod)<sup>87</sup>, and the diminished myeloid-derived suppressor cells and tumor-associated macrophages, which can suppress the development and metastasis of tumor<sup>88,89</sup>. Paquinimod has been under clinical trials for the treatment of autoimmune disorders, including systemic lupus erythematosus and systemic sclerosis<sup>90,91</sup>. Furthermore, Tasquinimod has been studied as an immunotherapeutic agent in solid tumors, especially prostate cancer<sup>92</sup>, and hematological malignancies like multiple myeloma<sup>93</sup> and Myelofibrosis<sup>94</sup>. Studies are going on to confirm the use of these investigational medications in practice. However, to the best of our knowledge, the applicability of these drugs in the treatment of CRSwNP has not been scientifically reported so far.

## Conclusion and further work

Advances in producing “omics” disease-related datasets have provided worthwhile research opportunities about disease-related pathways and genes. In this study, the three-way interaction approach was utilized for the first time to trace critical biomarkers and dysregulated biological pathways involved in CRSwNP pathogenesis. The three-way interaction approach can cope with the dynamic nature of co-expression relations by introducing a switch gene as a surrogate for the intrinsic state variable in the cell. Hence, such an approach describes a more comprehensive and precise comprehension of the underlying reasons for cellular changes. Moreover, the switch genes as the controller of the evolution in gene interactions can be considered potential therapeutic targets. Our study results revealed four dysregulated pathways in CRSwNP, including “positive regulation of intracellular signal transduction”, “arachidonic acid metabolic process”, “spermatogenesis” and “negative regulation of cellular protein metabolic process”. Additionally, the *S100a9* as a switch gene, together with the gene pair {*Cd14*, *Tpd52l1*} form a statistically significant and biologically relevant triplet. More specifically, we suggested that *S100a9* might act as a potential upstream modulator in the TLR4 transduction pathway in the major CRSwNP pathologies. Moreover, drug exploring results suggested that “Tasquinimod” and “Paquinimod” can potentially influence CRSwNP treatment by modulating the *S100a9* expression profile.

While our study provides new clues into the pathogenesis of CRSwNPs through computational approaches, further trials are required to validate such findings. In the next step, it is imperative to experimentally confirm the relationship between the *S100a9* gene and the {*Cd14*, *Tpd52l1*} gene pair.

## Methods

### Gene expression profiling dataset

The dataset used in this study comprises gene expression data from various samples. Specifically, it includes 42 samples of chronic rhinosinusitis with nasal polyps (CRSwNP-NP), 33 paired non-polyp inferior turbinate samples (CRSwNP-IT), and 28 samples from inferior turbinate controls without chronic rhinosinusitis (CS-IT). The gene expression data which was generated by the Illumina HiSeq 4000 platform is accessible in the Gene Expression Omnibus (GEO) database under the accession number GSE136825<sup>9</sup>.

To ensure data comparability and reliability, the raw RNA-Seq expression profiles underwent normalization using the reads per kilobase million (RPKM) values. This normalization method was implemented using the edgeR R package<sup>95</sup>. After normalization, 6018 genes were selected for further investigation and analysis.

Gene expression data were analyzed using one-way Analysis of Variance (ANOVA)<sup>96</sup> to determine the significance of differences among the study groups. Upon establishing significant differences from the ANOVA, post-hoc comparisons between group means were conducted using Tukey’s Honest Significant Difference (HSD) test<sup>97</sup>. This test was chosen with the objective of comparing multiple groups simultaneously. Statistical significance was set at a p-value of less than 0.05.

### Liquid association analysis

The liquid association measure was calculated to capture three-way interactions among all genes in the dataset. It was accomplished using the fastMLA R package<sup>14</sup>, which employs a modified liquid association algorithm. The algorithm assesses changes in co-expression relationships between two genes, X and Y, based on the expression level of a third gene, Z.

Specifically, the fast modified liquid association algorithm assigns an MLA score to each gene triplet, quantifying the magnitude of the liquid association. In detail, MLA ( $Z/\{X, Y\}$ ) can be estimated using the following formula:



$$\widehat{MLA} = \frac{\sum_i^M \widehat{\rho}_i \bar{Z}_i}{M}$$

Here,  $M$  represents the number of bins over  $Z$ ;  $\widehat{\rho}_i$  is the Pearson's correlation coefficient of  $X$  and  $Y$  in samples of the  $i$ th bin, and  $\bar{Z}_i$  is the mean expression value of  $Z$  in the  $i$ th bin.

Before running the fastMLA algorithm, two preprocessing steps are required. Firstly, the marginal distribution of each variable should follow a normal distribution, as determined by Li's approach<sup>11</sup>. Secondly, each variable should be standardized to have a mean of 0 and a variance of 1, which was achieved using the CTT package<sup>15</sup>. The first preprocessing step was performed using an in-house implementation, while the second one used the CTT package<sup>98</sup>.

To establish more conservative significance thresholds, considering the large number of tests conducted, the corrected p-values were estimated using the Bonferroni method<sup>99</sup>. Liquid association triplets with a corrected p-value less than 0.001 were deemed statistically significant. These thresholds were applied to ensure the reliable identification of significant associations amidst multiple tests.

### Functional enrichment analysis

Functional enrichment analysis, also called gene set enrichment analysis (GSEA), is a valuable statistical method used to classify genes or proteins over-represented within a specific gene set based on predefined annotations<sup>100</sup>. In the context of this study, the GSEA was employed for two purposes; (i) to trace the biologically relevant triplets and (ii) to identify key biological processes and pathways associated with the pathogenesis of CRSwNP in both NP and IT tissue.

This analysis focused on "biological processes" defined in the gene ontology (GO) database<sup>71</sup>, as well as "pathways" from the KEGG database<sup>101</sup>. Moreover, it was performed using the ClueGO tool<sup>102</sup> within the Cytoscape environment<sup>103</sup>. The right-sided hypergeometric test (with a Kappa threshold of 0.4) and the Benjamini–Hochberg (BH) correction method were used for the validation of enrichment<sup>104</sup>.

### Gene regulatory network construction

A gene regulatory network (GRN) is employed to conceptualize intricate regulatory mechanisms that control gene expression levels within cells. It is represented as a direct graph, composed of nodes (genes) and directed edges (regulatory connections) that exhibit either activatory or inhibitory interactions. Such a graph enables the prediction of gene expression patterns in various conditions<sup>105</sup>.

The GRN was reconstructed on the geWorkbench (genomics Workbench) platform<sup>106</sup> utilizing ARACNE (Algorithm for the Reconstruction of Accurate Cellular Networks)<sup>107</sup>. This algorithm employs gene expression data to reconstruct a cellular network based on a reverse engineering approach. It captures directed regulatory relationships between each transcriptional regulator and its potential target genes based on mutual information measures. Furthermore, this network was reconstructed for every gene implicated in the statistically significant triplets by considering Bonferroni-corrected p-value < 0.05<sup>99</sup>.

### Random forests classification

The Random Forest (RF) is an advanced classification method based on machine learning principles. It involves constructing a collection of independent decision trees using techniques like bagging and feature randomness. Each decision tree recursively partitions the data into more homogeneous subsets based on specific features, ultimately leading to accurate combined classification outputs.

A random forest classifier was constructed using the randomForest R package<sup>108</sup>. The algorithm was configured with a specific set of parameters. The "number of decision trees" parameter was set to 10,000 trees, ensuring a large ensemble to capture complex patterns and improve overall performance. The "mtry" parameter, determining the number of randomly selected features considered at each node, was set to the square root of the total number of features. This choice helps maintain diversity among the trees and prevent overfitting<sup>109</sup>. Furthermore, the gene importance measure is calculated by assessing the average increase in error rate across all the trees when a specific gene is randomly permuted.

A receiver operating characteristic (ROC) curve was generated<sup>110</sup> to assess the classification performance of the random forest classifier.

### Putative drugs exploring

This study employed two prominent drug databases, DrugBank<sup>111</sup> and DGI<sup>112</sup>, to identify potential drugs. DrugBank is a comprehensive and extensively curated database that provides detailed information about various drugs, including their chemical structures, pharmacological activities, mechanisms of action, and therapeutic indications. On the other hand, DGI is a specialized database that focuses on drug-gene interactions. It provides a comprehensive collection of genes and their drug interactions, including information on drug targets, mechanisms of action, and therapeutic associations.

### Data availability

The datasets generated and/or analysed during the current study are available in the Gene Expression Omnibus (GEO) repository [GSE136825] and in the Supplementary Information files.

Received: 20 November 2023; Accepted: 19 April 2024

Published online: 27 April 2024

## References

- Hopkins, C. Chronic rhinosinusitis with nasal polyps. *N. Engl. J. Med.* **381**(1), 55–63 (2019).
- Van Crombruggen, K. *et al.* Pathogenesis of chronic rhinosinusitis: Inflammation. *J. Allergy Clin. Immunol.* **128**(4), 728–732 (2011).
- Ragab, A., Clement, P. & Vincken, W. Objective assessment of lower airway involvement in chronic rhinosinusitis. *Am. J. Rhinol.* **18**(1), 15–21 (2004).
- Bachert, C., Zhang, L. & Gevaert, P. Current and future treatment options for adult chronic rhinosinusitis: Focus on nasal polyposis. *J. Allergy Clin. Immunol.* **136**(6), 1431–1440 (2015).
- Hastan, D. *et al.* Chronic rhinosinusitis in Europe—An underestimated disease. A GA2LEN study. *Allergy* **66**(9), 1216–1223 (2011).
- Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: Ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* **17**(6), 333–351 (2016).
- Tamayo, P. *et al.* Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. USA* **96**(6), 2907–2912 (1999).
- Eisen, M. B. *et al.* Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**(25), 14863–14868 (1998).
- Peng, Y. *et al.* Whole-transcriptome sequencing reveals heightened inflammation and defective host defence responses in chronic rhinosinusitis with nasal polyps. *Eur. Respir. J.* **54**(5), 1900732 (2019).
- Hao, Y. *et al.* Transcriptomic signatures and functional network analysis of chronic rhinosinusitis with nasal polyps. *Front. Genet.* **12**, 609754 (2021).
- Li, K.-C. Genome-wide coexpression dynamics: Theory and application. *Proc. Natl. Acad. Sci.* **99**(26), 16875–16880 (2002).
- Shokati Eshkiki, Z. *et al.* Novel insight into pancreatic adenocarcinoma pathogenesis using liquid association analysis. *BMC Med. Genom.* **15**(1), 30 (2022).
- Khayer, N. *et al.* Three-way interaction model with switching mechanism as an effective strategy for tracing functionally-related genes. *Expert Rev. Proteom.* **16**(2), 161–169 (2019).
- Gunderson, T. & Ho, Y. Y. An efficient algorithm to explore liquid association on a genome-wide scale. *BMC Bioinform.* **15**(1), 371 (2014).
- Ho, Y. Y. *et al.* Modeling liquid association. *Biometrics* **67**(1), 133–141 (2011).
- Alavi Majd, H. *et al.* Two-way gene interaction from microarray data based on correlation methods. *Iran. Red Crescent Med. J.* **18**(6), e24373 (2016).
- Acharyya, S., Zhou, X. & Baladandayuthapani, V. SpaceX: Gene co-expression network estimation for spatial transcriptomics. *Bioinformatics* **38**(22), 5033–5041 (2022).
- Khayer, N. *et al.* Nkx3-1 and Fech genes might be switch genes involved in pituitary non-functioning adenoma invasiveness. *Sci. Rep.* **11**(1), 20943 (2021).
- Khayer, N. *et al.* Three-way interaction model to trace the mechanisms involved in Alzheimer's disease transgenic mice. *PLoS One* **12**(9), e0184697 (2017).
- Khayer, N. *et al.* RT-DOb, a switch gene for the gene pair {Csf1r, Mir1}, can influence the onset of Alzheimer's disease by regulating communication between mast cell and microglia. *PLoS One* **18**(7), e0288134 (2023).
- Khayer, N. *et al.* Rps27a might act as a controller of microglia activation in triggering neurodegenerative diseases. *PLoS One* **15**(9), e0239219 (2020).
- Stevens, W. W., Schleimer, R. P. & Kern, R. C. Chronic rhinosinusitis with nasal polyps. *J. Allergy Clin. Immunol.* **4**(4), 565–572 (2016).
- Li, H. *et al.* Src family kinases and pulmonary fibrosis: A review. *Biomed. Pharmacother.* **127**, 110183 (2020).
- Bae, J.-S. *et al.* Effects of Wnt signaling on epithelial to mesenchymal transition in chronic rhinosinusitis with nasal polyp. *Thorax* **75**(11), 982–993 (2020).
- Milara, J. *et al.* Mucin 1 downregulation associates with corticosteroid resistance in chronic rhinosinusitis with nasal polyps. *J. Allergy Clin. Immunol.* **135**(2), 470–476 (2015).
- Wang, Z. *et al.* Interleukin-1 $\beta$  regulates the expression of glucocorticoid receptor isoforms in nasal polyps in vitro via p38 MAPK and JNK signal transduction pathways. *J. Inflamm.* **12**(1), 3–12 (2015).
- Mócsai, A., Ruland, J. & Tybulewicz, V. L. The SYK tyrosine kinase: A crucial player in diverse biological functions. *Nat. Rev. Immunol.* **10**(6), 387–402 (2010).
- Hamajima, Y. *et al.* Expression of Syk is associated with nasal polyp in patients with allergic rhinitis. *Auris Nasus Larynx* **34**(1), 49–56 (2007).
- Patou, J. *et al.* Syk-kinase inhibition prevents mast cell activation in nasal polyps. *Rhinology* **49**(1), 100–106 (2011).
- Böscke, R. *et al.* Wnt signaling in chronic rhinosinusitis with nasal polyps. *Am. J. Respir. Cell Mol. Biol.* **56**(5), 575–584 (2017).
- Ballester, B., Milara, J. & Cortijo, J. The role of mucin 1 in respiratory diseases. *Eur. Respir. Rev.* **30**(159), 200149 (2021).
- Wang, B. *et al.* Metabolism pathways of arachidonic acids: Mechanisms and potential therapeutic targets. *Signal Transduct. Target. Ther.* **6**(1), 94 (2021).
- Funk, C. D. Prostaglandins and leukotrienes: Advances in eicosanoid biology. *Science* **294**(5548), 1871–1875 (2001).
- Sala, A. *et al.* Two-pronged approach to anti-inflammatory therapy through the modulation of the arachidonic acid cascade. *Biochem. Pharmacol.* **158**, 161–173 (2018).
- Bahia, M. S. *et al.* Inhibitors of microsomal prostaglandin E2 synthase-1 enzyme as emerging anti-inflammatory candidates. *Med. Res. Rev.* **34**(4), 825–855 (2014).
- Jung, T. T. *et al.* Prostaglandins, leukotrienes, and other arachidonic acid metabolites in nasal polyps and nasal mucosa. *Laryngoscope* **97**(2), 184–189 (1987).
- Nordström, A. *et al.* Levels of eicosanoids in nasal secretions associated with nasal polyp severity in chronic rhinosinusitis. *Prostaglandins Leukot. Essent. Fatty Acids* **184**, 102474 (2022).
- Tsai, Y.-J. *et al.* Thromboxane A2 regulates CXCL1 and CXCL8 chemokine expression in the nasal mucosa-derived fibroblasts of chronic rhinosinusitis patients. *PLoS One* **11**(6), e0158438 (2016).
- Hsu, J. *et al.* Genetics of chronic rhinosinusitis: State of the field and directions forward. *J. Allergy Clin. Immunol.* **131**(4), 977–993.e5 (2013).
- Lannou, D. L. *et al.* Genetics: Obstructive azoospermia with agenesis of vas deferens or with bronchiectasia (Young's syndrome): A genetic approach. *Hum. Reprod.* **10**(2), 338–341 (1995).
- Hwang, K. *et al.* Mendelian genetics of male infertility. *Ann. N. Y. Acad. Sci.* **1214**(1), E1–E17 (2010).
- Hinzpeter, A. *et al.* Identification of a novel 5' alternative CFTR mRNA isoform in a patient with nasal polyposis and CFTR mutations. *Hum. Mutat.* **35**(7), 805–808 (2014).
- Erbek, S. S. *et al.* Expression of a disintegrin and metalloproteinase 33 protein in nasal polyposis: An immunohistochemical study. *Am. J. Rhinol. Allergy* **24**(3), 79–82 (2010).
- Wolfsberg, T. G. *et al.* The precursor region of a protein active in sperm-egg fusion contains a metalloprotease and a disintegrin domain: Structural, functional, and evolutionary implications. *Proc. Natl. Acad. Sci. USA* **90**(22), 10783–10787 (1993).

45. Wolfsberg, T. G. *et al.* ADAM, a novel family of membrane proteins containing a disintegrin and metalloprotease domain: Multipotential functions in cell-cell and cell-matrix interactions. *J. Cell Biol.* **131**(2), 275–278 (1995).
46. Priyadharshini, V. *et al.* Transcriptome analysis identifies doublesex and Mab-3 related transcription factor (DMRT3) in nasal polyp epithelial cells of patients suffering from non-steroidal anti-inflammatory drug-exacerbated respiratory disease (AERD). *Biomolecules* **11**(8), 1092 (2021).
47. Kao, S. *et al.* Scoping review of chronic rhinosinusitis proteomics. *Rhinology* **58**(5), 418–429 (2020).
48. Kim, Y.-S. *et al.* In-depth, proteomic analysis of nasal secretions from patients with chronic rhinosinusitis and nasal polyps. *Allergy Asthma Immunol. Res.* **11**(5), 691–708 (2019).
49. Gonzalez, L. L., Garrie, K. & Turner, M. D. Role of S100 proteins in health and disease. *Biochim. Biophys. Acta (BBA) Mol. Cell Res.* **1867**(6), 118677 (2020).
50. Van Crombruggen, K. *et al.* Differential release and deposition of S100A8/A9 proteins in inflamed upper airway tissue. *Eur. Respir. J.* **47**(1), 264–274 (2016).
51. Boruk, M. *et al.* Elevated S100A9 expression in chronic rhinosinusitis coincides with elevated MMP production and proliferation in vitro. *Sci. Rep.* **10**(1), 16350 (2020).
52. Huang, Y. *et al.* Reduced expression of antimicrobial protein secretory leukoprotease inhibitor and clusterin in chronic rhinosinusitis with nasal polyps. *J. Immunol. Res.* **2021**, 1057186 (2021).
53. Sumsion, J. S., Pulsipher, A. & Alt, J. A. Differential expression and role of S100 proteins in chronic rhinosinusitis. *Curr. Opin. Allergy Clin. Immunol.* **20**(1), 14–22 (2020).
54. Yao, Y. *et al.* Comparative analysis of inflammatory signature profiles in eosinophilic and noneosinophilic chronic rhinosinusitis with nasal polyposis. *Biosci. Rep.* **40**(2), BSR20193101 (2020).
55. O'Connell, B. P. *et al.* Systemic monocyte-derived dendritic cells and associated Th2 skewing in chronic rhinosinusitis. *Otolaryngol. Head Neck Surg.* **150**(2), 312–320 (2014).
56. Boutros, R. *et al.* Alternative splicing as a mechanism for regulating 14-3-3 binding: Interactions between hD53 (TPD52L1) and 14-3-3 proteins. *J. Mol. Biol.* **332**(3), 675–687 (2003).
57. Abba, M. C. *et al.* Breast cancer molecular signatures as determined by SAGE: Correlation with lymph node status. *Mol. Cancer Res.* **5**(9), 881–890 (2007).
58. Barbaric, D. *et al.* Expression of tumor protein D52-like genes in childhood leukemia at diagnosis: Clinical and sample considerations. *Leuk. Res.* **30**(11), 1355–1363 (2006).
59. Hong, Q. *et al.* Transcriptomic analyses of the adenoma-carcinoma sequence identify hallmarks associated with the onset of colorectal cancer. *Front. Oncol.* **11**, 704531 (2021).
60. Zeng, H. *et al.* Integrated analysis of multiple microarrays based on raw data identified novel gene signatures in recurrent implantation failure. *Front. Endocrinol. (Lausanne)* **13**, 785462 (2022).
61. Schrupp, D. S. *et al.* Clinical and molecular responses in lung cancer patients receiving Romidepsin. *Clin. Cancer Res.* **14**(1), 188–198 (2008).
62. Cheon, C. W. *et al.* Effects of ciglitazone and troglitazone on the proliferation of human stomach cancer cells. *World J. Gastroenterol.* *WJG* **15**(3), 310 (2009).
63. Zhou, R.-S. *et al.* Anticancer effects of emodin on HepG2 cell: Evidence from bioinformatic analysis. *BioMed Res. Int.* **2019**, 3065818 (2019).
64. Song, L. *et al.* Transcriptomic analysis of the association between diabetes mellitus and myocardial infarction. *Exp. Clin. Endocrinol. Diabetes* **127**(09), 603–614 (2019).
65. Esquinas, C. *et al.* Gene and miRNA expression profiles in PBMCs from patients with severe and mild emphysema and PiZZ alpha1-antitrypsin deficiency. *Int. J. Chron. Obstr. Pulmon. Dis.* **12**, 3381–3390 (2017).
66. Kim, H. J. *et al.* Toll-like receptor 4 (TLR4): New insight immune and aging. *Immun. Ageing* **20**(1), 67 (2023).
67. Hoessel, B. & Schmid, J. A. The complexity of NF- $\kappa$ B signaling in inflammation and cancer. *Mol. Cancer* **12**, 86 (2013).
68. Arthur, J. S. & Ley, S. C. Mitogen-activated protein kinases in innate immunity. *Nat. Rev. Immunol.* **13**(9), 679–692 (2013).
69. Liu, S. *et al.* Phosphorylation of innate immune adaptor proteins MAVS, STING, and TRIF induces IRF3 activation. *Science* **347**(6227), aaa2630 (2015).
70. Ojaniemi, M. *et al.* Phosphatidylinositol 3-kinase is involved in Toll-like receptor 4-mediated cytokine expression in mouse macrophages. *Eur. J. Immunol.* **33**(3), 597–605 (2003).
71. Ashburner, M. *et al.* Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**(1), 25–29 (2000).
72. Chen, K. *et al.* The active contribution of Toll-like receptors to allergic airway inflammation. *Int. Immunopharmacol.* **11**(10), 1391–1398 (2011).
73. Cho, J. S. *et al.* Lipopolysaccharide induces pro-inflammatory cytokines and MMP production via TLR4 in nasal polyp-derived fibroblast and organ culture. *PLoS One* **9**(11), e90683 (2014).
74. Cho, J. S. *et al.* Activation of TLR4 induces VEGF expression via Akt pathway in nasal polyps. *Clin. Exp. Allergy* **43**(9), 1038–1047 (2013).
75. Cho, J. S. *et al.* Steroids inhibit vascular endothelial growth factor expression via TLR4/Akt/NF- $\kappa$ B pathway in chronic rhinosinusitis with nasal polyp. *Exp. Biol. Med. (Maywood)* **239**(8), 913–921 (2014).
76. Riva, M. *et al.* Induction of nuclear factor- $\kappa$ B responses by the S100A9 protein is Toll-like receptor-4-dependent. *Immunology* **137**(2), 172–182 (2012).
77. Lv, Z., Li, W. & Wei, X. S100A9 promotes prostate cancer cell invasion by activating TLR4/NF- $\kappa$ B/integrin  $\beta$ 1/FAK signaling. *OncoTargets Ther.* **13**, 6443–6452 (2020).
78. Franz, S. *et al.* Overexpression of S100A9 in obesity impairs macrophage differentiation via TLR4-NF $\kappa$ B-signaling worsening inflammation and wound healing. *Theranostics* **12**(4), 1659 (2022).
79. Ehrchen, J. M. *et al.* The endogenous Toll-like receptor 4 agonist S100A8/S100A9 (calprotectin) as innate amplifier of infection, autoimmunity, and cancer. *J. Leukoc. Biol.* **86**(3), 557–566 (2009).
80. Zanoni, I. *et al.* CD14 controls the LPS-induced endocytosis of Toll-like receptor 4. *Cell* **147**(4), 868–880 (2011).
81. Kagan, J. C. *et al.* TRAM couples endocytosis of Toll-like receptor 4 to the induction of interferon- $\beta$ . *Nat. Immunol.* **9**(4), 361–368 (2008).
82. Latz, E. *et al.* Lipopolysaccharide rapidly traffics to and from the Golgi apparatus with the toll-like receptor 4-MD-2-CD14 complex in a process that is distinct from the initiation of signal transduction. *J. Biol. Chem.* **277**(49), 47834–47843 (2002).
83. He, Z. *et al.* CD14 is a co-receptor for TLR4 in the S100A9-induced pro-inflammatory response in monocytes. *PLoS One* **11**(5), e0156377 (2016).
84. Kamal, A. H. M., Chakrabarty, J. K. & Chowdhury, S. M. Lipopolysaccharide and statin-mediated immune-responsive protein networks revealed in macrophages through affinity purification spacer-arm controlled cross-linking (AP-SPACC) proteomics. *Mol. Omics* **19**(1), 48–59 (2023).
85. Boros, F. & Vécsei, L. Progress in the development of kynurenine and quinoline-3-carboxamide-derived drugs. *Expert Opin. Investig. Drugs* **29**(11), 1223–1247 (2020).
86. Björk, P. *et al.* Identification of human S100A9 as a novel target for treatment of autoimmune disease via binding to quinoline-3-carboxamides. *PLoS Biol.* **7**(4), e97 (2009).

87. Raymond, E. *et al.* Mechanisms of action of tasquinimod on the tumour microenvironment. *Cancer Chemother. Pharmacol.* **73**(1), 1–8 (2014).
88. Deronic, A. *et al.* The quinoline-3-carboxamide paquinimod (ABR-215757) reduces leukocyte recruitment during sterile inflammation: Leukocyte- and context-specific effects. *Int. Immunopharmacol.* **18**(2), 290–297 (2014).
89. Deronic, A. *et al.* The anti-tumor effect of the quinoline-3-carboxamide tasquinimod: Blockade of recruitment of CD11b(+) Ly6C(hi) cells to tumor tissue reduces tumor growth. *BMC Cancer* **16**, 440 (2016).
90. Bengtsson, A. A. *et al.* Pharmacokinetics, tolerability, and preliminary efficacy of paquinimod (ABR-215757), a new quinoline-3-carboxamide derivative: Studies in lupus-prone mice and a multicenter, randomized, double-blind, placebo-controlled, repeat-dose, dose-ranging study in patients with systemic lupus erythematosus. *Arthritis Rheum.* **64**(5), 1579–1588 (2012).
91. Hesselstrand, R. *et al.* An open-label study to evaluate biomarkers and safety in systemic sclerosis patients treated with paquinimod. *Arthritis Res. Ther.* **23**(1), 1–12 (2021).
92. Fizazi, K. *et al.* A randomized, double-blind, placebo-controlled phase II study of maintenance therapy with tasquinimod in patients with metastatic castration-resistant prostate cancer responsive to or stabilized during first-line docetaxel chemotherapy. *Ann. Oncol.* **28**(11), 2741–2746 (2017).
93. Vogl, D. T. *et al.* A phase 1 study of tasquinimod in patients with relapsed or refractory multiple myeloma. *Blood* **136**, 17–18 (2020).
94. Gleitz, H. *et al.* S205: Inhibiting the alarmin-driven hematopoiesis-stromal cells crosstalk in primary myelofibrosis ameliorates bone marrow fibrosis. *HemaSphere* **7**, e36177f7 (2023).
95. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**(1), 139–140 (2010).
96. Quirk, T. J. & Quirk, T. J. One-way analysis of variance (ANOVA). *Excel 2007 for Educational and Psychological Statistics: A Guide to Solving Practical Problems*, 163–179 (2012).
97. Abdi, H. & Williams, L. J. Tukey's honestly significant difference (HSD) test. *Encycl. Res. Des.* **3**(1), 1–5 (2010).
98. Willse, J. T. & Willse, M. J. T. *Package 'CTT'* (2014).
99. Bland, J. M. & Altman, D. G. Multiple significance tests: The Bonferroni method. *BMJ* **310**(6973), 170 (1995).
100. Zarnegarnia, Y. *et al.* Application of fuzzy clustering in analysis of included proteins in esophagus, stomach and colon cancers based on similarity of Gene Ontology annotation. *Koomesh* **12**(1), 14–21 (2010).
101. Kanehisa, M. *et al.* KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* **40**(D1), D109–D114 (2012).
102. Bindea, G. *et al.* ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**(8), 1091–1093 (2009).
103. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**(11), 2498–2504 (2003).
104. Thissen, D., Steinberg, L. & Kuang, D. Quick and easy implementation of the Benjamini–Hochberg procedure for controlling the false positive rate in multiple comparisons. *J. Educ. Behav. Stat.* **27**(1), 77–83 (2002).
105. Remo, A. *et al.* Systems biology analysis reveals NFAT5 as a novel biomarker and master regulator of inflammatory breast cancer. *J. Transl. Med.* **13**(1), 138 (2015).
106. Floratos, A. *et al.* geWorkbench: An open source platform for integrative genomics. *Bioinformatics* **26**(14), 1779–1780 (2010).
107. Margolin, A. A. *et al.* ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinform.* **7**, S7 (2006).
108. Breiman, L. Random forests. *Mach. Learn.* **45**(1), 5–32 (2001).
109. Wright, M. N. & Ziegler, A. ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Stat. Softw.* **77**(1), 1–17 (2015).
110. Hoo, Z. H., Candlish, J. & Teare, D. What is an ROC curve?. *Emerg. Med. J.* **34**(6), 357–359 (2017).
111. Law, V. *et al.* DrugBank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Res.* **42**(D1), D1091–D1097 (2014).
112. Cotto, K. C. *et al.* DGIdb 3.0: A redesign and expansion of the drug–gene interaction database. *Nucleic Acids Res.* **46**(D1), D1068–D1073 (2018).

## Acknowledgements

We would like to thank The Five Senses Health Institute for providing us with research facilities.

## Author contributions

All authors contributed to the study conception and design. N.Kh. and M.J. conceived and directed the study. N.Kh. designed the data analysis framework and performed the data analyses and evaluated the results. M.J. discussed the results. N.Kh. wrote the first draft of manuscript. Z.A. had an effective role in writing the discussion section. M.F. assisted with the revision of the manuscript. All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-60205-4>.

**Correspondence** and requests for materials should be addressed to N.K. or M.J.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024