Received: 3 June 2022 | Revised: 28 November 2022 | Accepted: 3 December 2022

DOI: 10.1111/jcpe.13758

ORIGINAL ARTICLE

Journal of Clinical Periodontology WILEY

RSPO4 is a potential risk gene of stages III–IV, grade C periodontitis through effects on innate immune response and oral barrier integrity

Avneesh Chopra[1] | Jiahui Song[1] | January Weiner 3rd[2] | Huseyin G. Keceli[3] | Pervin R. Dincer[4] | Raquel Cruz[5,6] | Angel Carracedo[5,6] | Juan Blanco[7] | Henrik Dommisch[1] | Arne S. Schaefer[1]

[1]Department of Periodontology, Oral Medicine and Oral Surgery, Institute for Dental and Craniofacial Sciences, Charité - University Medicine Berlin, Berlin, Germany

[2]Core Unit Bioinformatics, Berlin Institute of Health, Berlin, Germany

[3]Department of Periodontology, Faculty of Dentistry, Hacettepe University, Ankara, Turkey

[4]Faculty of Medicine, Department of Medical Biology, Hacettepe University, Ankara, Turkey

[5]Fundación Pública Galega de Medicina Xenómica, Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain

[6]CIBERER-Instituto de Salud Carlos III, Centro Singular de Investigación en Medicina Molecular y Enfermedades Crónicas (CIMUS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

[7]Grupo de Investigación en Odontología Médico-Quirúrgica (OMEQUI), Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

Correspondence
Arne S. Schaefer, Genetics of Oral Inflammatory Diseases Group, Charité - University Medicine Berlin, Berlin, Germany.
Email: arne.schaefer@charite.de

Funding information
Charité – University Medicine Berlin; the Hacettepe University Scientific Research Unit, Grant/Award Number: THD-2021-19415

Abstract

Aim: *R-spondin 4* (*RSPO4*) is a suggestive risk gene of stage III–IV, grade C periodontitis and upregulated in gingiva of mice resistant to bacteria-induced alveolar bone loss. We aimed to replicate the association, identify and characterize the putative causal variant(s) and molecular effects, and understand the downstream effects of *RSPO4* upregulation.

Materials and Methods: We performed a two-step association study for *RSPO4* with imputed genotypes of a German–Dutch (896 stage III–IV, grade C periodontitis cases, 7104 controls) and Spanish sample (441 cases and 1141 controls). We analysed the allelic effects on transcription factor binding sites with reporter gene and antibody electrophoretic mobility shift assays. We used CRISPR/dCas9 activation and RNA sequencing to pinpoint *RSPO4* as the target gene and to analyse downstream effects.

Results: *RSPO4* was associated with periodontitis (rs6056178, $p_{meta} = 4.6 \times 10^{-5}$). rs6056178 contains a GATA-binding motif. The rs6056178 T-allele abolished reporter activity ($p = .004$) and reduced GATA binding (−14.5%). CRISPRa of the associated region increased *RSPO4* expression (25.8 ± 6.5-fold, $p = .003$). *RSPO4* activation showed strongest induction of Gliomedin (439-fold) and Mucin 21 (178-fold) and of the gene set "response to interferon-alpha" (area under the curve [AUC] = 0.8, $p < 5 \times 10^{-6}$). The most repressed gene set was "extracellular matrix interactions" (AUC = 0.8, $p_{adj} = .00016$).

Conclusion: *RSPO4* is a potential periodontitis risk gene and modifies host defence and barrier integrity.

KEYWORDS
causal variant, GATA, *GLDN*, interferon alpha, *MUC21*, rs6056178

Clinical Relevance

Scientific rationale for study: For a better understanding of the aetiology of periodontal diseases, it is important to know the molecular mechanisms that contribute to resistance and/or susceptibility to periodontal bone loss and to identify the molecular genetic pathways involved.

Principal findings: We validated the association of the *R-spondin 4* (*RSPO4*) gene with stage III–IV, grade C periodontitis, identified the putative causal variant of the association, and showed that it regulates the mucin barrier, extracellular matrix interactions, and interferon-alpha signalling.

Practical implications: *RSPO4* is involved in the aetiology of severe progressive periodontitis. This knowledge contributes to the understanding of the causal factors of periodontitis better.

1 | INTRODUCTION

Periodontitis is an oral inflammatory disease that is diagnosed by progressive alveolar bone loss due to long-term inflammation caused by dysbiosis of the oral biofilm (Lamont et al., 2018). However, only a few individuals develop severe progressive forms of periodontitis and it is important to know the molecular mechanisms that contribute to resistance. In a previous study, we analysed gene expression data of periodontal biopsies from different recombinant inbred mouse lines (RILs) of the Collaborative Cross (CC) mouse resource (Churchill et al., 2004; Threadgill et al., 2011). In our work, we identified several RILs that showed either susceptibility or resistance to alveolar bone loss 6 weeks after gingival infection with the periodontal pathogenic bacteria *Porphyromonas gingivalis* and *Fusobacterium nucleatum* (Nashef et al., 2020). Mouse jaws of resistant RILs expressed four genes significantly more strongly than susceptible RILs, as revealed by RNA sequencing (RNA-Seq). These were the genes *Matrix Metalloproteinase 20* (*MMP20*; 181 fold change [FC]), *Calbindin* (*CALB1*; 60 FC), *Fibrillin* 2 (*FBN2*; 7 FC), and *R-spondin 4* (*RSPO4*; 13 FC). This suggested that increased expression of these genes in resistant RILs contributed to the prevention of alveolar bone loss in the course of bacteria-induced oral inflammation.

In complex diseases such as periodontitis, single common susceptibility variants often have moderate effects. For this reason, the probability for significant enrichment in a selected case group compared to the controls may be too low to generate statistically significant associations. Particularly, if the sample size is limited, it is difficult to discover individual variants with moderate effects at genomewide statistical significance. However, knowledge of loci which have causal implications in the aetiology of disease but did not show genomewide significant associations because of restricted statistical power from limited sample sizes is of high scientific and clinical relevance because true causal associations may provide insight into key pathways of the disease aetiology. To increase statistical power if the sample sizes are limited, gene-based tests are used to detect associations by combining the effects of several variants in the same gene. They assess whether categories of variants within a gene or set of genes occur more commonly in cases than controls. We recently performed a gene-based test with genotypes from a genomewide association study (GWAS) of stage III–IV, grade C periodontitis (de Coo et al., 2021). We identified suggestive associations of two genes, one of which was *RSPO4* ($p = 5.5 \times 10^{-5}$). *RSPO4* has a role in osteogenesis (Gaur et al., 2005) and is an activator of canonical Wnt signalling (de Lau et al., 2011; Glinka et al., 2011). The independent discoveries of *RSPO4* in two systematic genomewide expression and single nucleotide polymorphism (SNP) screens of periodontitis-resistant mouse RILs (Nashef et al., 2020) and human stage III–IV, grade C periodontitis (de Coo et al., 2021), respectively, implied a role of *RSPO4* in the aetiology of periodontitis. This makes it an interesting target for a better understanding of the molecular mechanisms that drive disease resistance and susceptibility.

In this study, we hypothesized that biological functional variants at *RSPO4* increase the risk of human stage III–IV, grade C periodontitis and that elevated *RSPO4* transcript levels contribute to resistance to alveolar bone loss. We aimed to identify the putative causal variant(s) of the association, characterize their molecular effects, and find the genes and gene sets that respond to increased *RSPO4* expression.

2 | MATERIALS AND METHODS

The study design as well as the workflow of the experiments is shown in Figure 1.

2.1 | Genetic association study of *RSPO4*

In a genomewide gene-based test analysis, we previously identified an association of *RSPO4* (de Coo et al., 2021). The gene-based test design assessed whether categories of variants within a gene or genomic region occur more commonly in cases than controls, but it did not identify associations for specific SNPs. In this study, we aimed to identify specific SNPs within the genetic region of *RSPO4* that cause the observed gene-wide association with stage III–IV, grade C periodontitis. To generate hypotheses on putative susceptibility SNPs, we performed an explorative candidate-gene association study with a North-West European case–control sample of 896 stage III–IV, grade C periodontitis cases and 7104 controls (see Figure 1 for the experimental workflow and study design). The study sample has been described in detail earlier (Munz et al., 2017; Freitag-Wolf et al., 2021). In brief, the cases (61% female sex) were recruited across Germany and the Netherlands and in Vienna (Austria) who had >30% alveolar bone loss at two or more teeth diagnosed at 35 or less years of age. This diagnosis indicated severe, rapid disease progression corresponding to stage III–IV, grade C periodontitis (Caton et al., 2018).

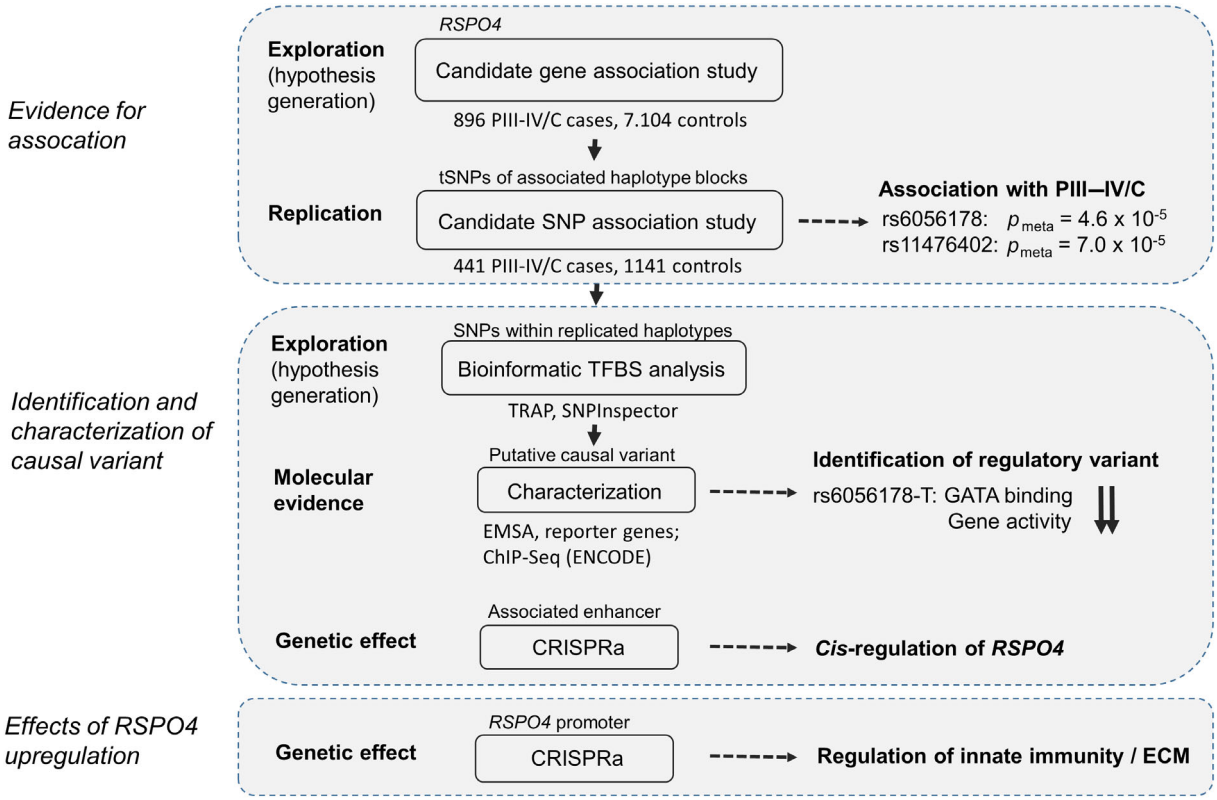

FIGURE 1 Experimental workflow

The German/Dutch control sample was recruited across north and west Germany and the Netherlands and consisted of 7104 individuals (51% female). We used the North European case–control sample, which was more than 2 times larger than the Spanish sample, because it had a higher statistical power, thereby reducing the probability of testing false positives and tests in the replication. Subsequently, we tested the haplotype tagging SNPs (tSNPs) for association, which showed significant associations in the explorative study. For this replication, we used the genotypes of the Spanish GWAS sample of de Coo et al. because another similar large case–control sample of stage III–IV, grade C periodontitis was not available. This sample has been described in detail earlier (de Coo et al., 2021). In brief, 441 Spanish cases (66% female sex) had stage III–IV, grade C disease diagnosed between 14 and 35 years of age. The controls ($n = 1141$, 51% female sex, mean age 46 years) were randomly selected from the Spanish National DNA BioBank.

Associations were tested with the additive genetic model using the software Plink (Purcell et al., 2007). Combined *p*-values were calculated from the individual GWAS summary statistics as described in Supplementary Methods. We assessed the linkage disequilibrium (LD, $r^2 > 0.8$) with genotypes of the European populations CEU, GBR, and IBS, using the online tool LDproxy (Machiela & Chanock, 2015). All persons gave their informed consent prior to their inclusion in the study. The local ethics committees at the medical faculty of the Christian-Albrechts-Universität zu Kiel, Germany (vote for cases: A 156/03, vote for controls: B 231/98), and at Santiago de Compostela (Spain; 2015-372) approved the study.

2.2 | Screening for functional periodontitis-associated variants

We bioinformatically investigated whether the effect alleles of the associated SNPs changed predicted transcription factor binding sites (TFBS) using the Gene Regulation Module of the Software SNPInspector (Genomatix, Precigen) and the Transcription factor Affinity Prediction (TRAP) Web Tool (Thomas-Chollier et al., 2011) with the methods TRAP single sequence (Roider et al., 2007; Manke et al., 2008) and sTRAP (Manke et al., 2010). LD and population-specific haplotype structure analyses were performed using the webtool LDlink (Machiela & Chanock, 2015). We analysed all seven common SNPs (MAF >0.05) that are in LD ($r^2 > 0.8$) with the two associated haplotype blocks, which were tagged by rs6056178 and rs11476402 (Table S1, Figure S1). TF binding motives were confirmed using the web interface for Position Weight Matrix (PWM) model generation and evaluation, PWMTools (Ambrosini et al., 2018).

We used the published chromatin immunoprecipitation followed by sequencing (ChIP-Seq) data from multiple cell types that were provided by ENCODE (Hoffman et al., 2013) to interrogate whether binding of the predicted TFs at the SNP sequences had been found in the chromatin context (Supplementary Material, Figure S2).

2.3 | In vitro experiments

Methods used for cell culture, luciferase reporter gene assays, CRISPR/dCas9 activation, quantitative real-time PCR, and RNA sequencing are described in detail in Supplementary Material. In brief, we performed electrophoretic mobility shift assays (EMSA: Gelshift Chemiluminescent EMSA Kit; Active Motif, monoclonal GATA1 and GATA2 antibodies; Santa Cruz Biotechnology) and reporter gene experiments (pGL4.24 plasmid and Dual-Luciferase Stop & Glo Reporter Assay System; Promega) with the osteoblast-like human bone osteosarcoma cell line SaOS-2. To activate *RSPO4* expression for subsequent RNA-Seq, we used CRISPRa, which provides an advantage to episomal over-expression plasmids because it allows gene activation in the native chromatin context. However, the CRISPRa transfection efficiency can be low and the CRISPRa plasmids can be toxic for various cell types. To this end, we used HeLa cells. HeLa cells can be transfected efficiently with the CRISPRa dCAS9-VP64 system. RNA was harvested after 48 h. RNA-Seq was performed at the Berlin Institute of Health, Core Facility Genomics, as described in Methods.

3 | RESULTS

3.1 | *RSPO4* is associated with stage III–IV, grade C periodontitis in different European populations

We explored the associations of common SNPs at *RSPO4* with imputed genotypes of our GWAS stage III–IV, grade C periodontitis case–control sample of North-West European descent. We identified SNP associations that showed moderate significant associations, with $p > 5 \times 10^{-4}$, spanning a 3.5-kbp (kilo base pair)-long intergenic region downstream of *RSPO4* (chr20:931170–934617, hg19). The associated SNPs showed strong linkage ($r^2 > 0.8$) separated on three independent haplotype blocks (Figure S1). For replication, we selected a tSNP of each haplotype block and tested these three SNPs for association in our second stage III–IV, grade C periodontitis case–control sample of Spanish descent (Table 1). In this sample, the associations of two haplotype blocks replicated. These blocks are located near the 3′ end of *RSPO4* (chr20:934003–934617). tSNPs rs6056178 and rs11476402 were associated with $p = 1.8 \times 10^{-4}$ and $p = 5.8 \times 10^{-5}$, respectively. In the meta-analysis of both samples, these SNPs were associated with $p = 4.6 \times 10^{-5}$ and $p = 7.0 \times 10^{-5}$, respectively.

3.2 | rs6056178 is a biological functional variant that modulates GATA binding

To identify putative functional SNPs, we bioinformatically interrogated the the common SNPs at these two haplotype blocks to find TFs that possibly bind at the SNP sequences and to find whether potential TFBS were affected by the different alleles. The strongest allele-specific differences were found at rs6056178 for GATA2 binding (difference $\log_p = 2.1$) with significant predicted binding at the common G-allele ($p < .0025$) and no significant binding at the alternative T-allele (Table S1). The highest PWM similarity score was found for the GATA1 motif with 97.5% (Figure 2).

rs6056178 was located within a reported TF ChIP-seq cluster that demonstrated GATA2 and GATA3 binding (reported by ENCODE, Figure S2). Within the DNA sequence of this ChIP peak, we found a single GATA binding site (Supplementary Material).

We validated the predicted GATA1 and GATA2 binding at rs6056178 in vitro. To this end, we conducted GATA1 and GATA2 antibody EMSAs with allele-specific DNA probes and nuclear protein extract from SaOS-2 cells. We observed allele-specific binding of GATA1 and GATA2 antibodies at the DNA probes (Figure 2). The effect T-allele (alternative rare allele) abrogated antibody binding by 16.5% and 14.5%, respectively, compared to the reference common G-allele.

Subsequently, we quantified putative effects of the two rs6056178 alleles on the activity of the luciferase reporter gene in SaOS-2 cells. The DNA sequence with the rs6056178 common G-allele cloned upstream of the minimal promoter of the reporter gene significantly increased gene activity compared to the control reporter plasmid, which lacked the putative regulatory sequence at rs6056178 (1.9 ± 0.2-fold, $p = .004$). In contrast, the rare T-allele showed no regulatory effect on luciferase activity (Figure 3). This indicated that the regulatory element at rs6056178 acts as an enhancer and that the common G-allele is functional. The difference between both alleles was significant with $p = .01$, indicating a reduced functional effect of the disease-associated rare T-allele.

3.3 | *RSPO4* is the target gene of the association

To validate cis-regulation of the disease-associated chromatin elements tagged by rs114691134, rs6056178, and rs11476402 on the expression of *RSPO4*, we performed CRISPRa in HeLa cells. We chose this cell type because SaOS-2 cells showed poor survival after transfection of the CRISPRa system. A CRISPRa-gRNA that targeted the *RSPO4* promoter increased the expression of *RSPO4* 6878 ± 394-fold ($p \leq .0001$) relative to the unspecific control sgRNA (Figure 3). CRISPRa of rs114691134 (gRNA1 annealing 50 bp downstream to this SNP) increased the expression of *RSPO4* 11.6 ± 3.3-fold compared to the scrambled unspecific control sgRNA ($p = .005$). The second gRNA, located 1 kb upstream and downstream of the TF-ChIP-seq clusters tagged by rs114691134 and rs6056178, respectively, increased *RSPO4* expression 25.8 ± 6.5-fold ($p = .003$). These results showed that this associated chromatin element strongly regulates *RSPO4* expression in cis, implying *RSPO4* as a target gene of the association. Two other gRNAs, which targeted the chromatin at rs6056178 (95–155 bp distance) and rs11476402 (0–54 bp distance), showed no effect (data not shown).

TABLE 1 *RSPO4* associated with stages III–IV, grade C periodontitis in different European populations.

Haplotype block	tSNP	Exploration *p* (GER/NL)	OR (95% CI)	Replication *p* (Spain)	OR (95% CI)	*p* (meta)
1	rs114035391	.0014	1.3 (1.12–1.58)	n.s.	n.s.	-
2	rs6056178	.045	1.2 (1.0–1.42)	1.8e−4	1.39 (1.17–1.65)	4.6×10^{-5}
3	rs11476402	.025	1.14 (1.02–1.27)	5.8e−5	1.40 (1.19–1.65)	7.0×10^{-5}

Abbreviations: CI, confidence interval; GER, Germany; NL, The Netherlands; n.s., not significant; OR, odds ratio; *RSPO4*, *R-spondin 4*; tSNP, tagging single nucleotide polymorphism.

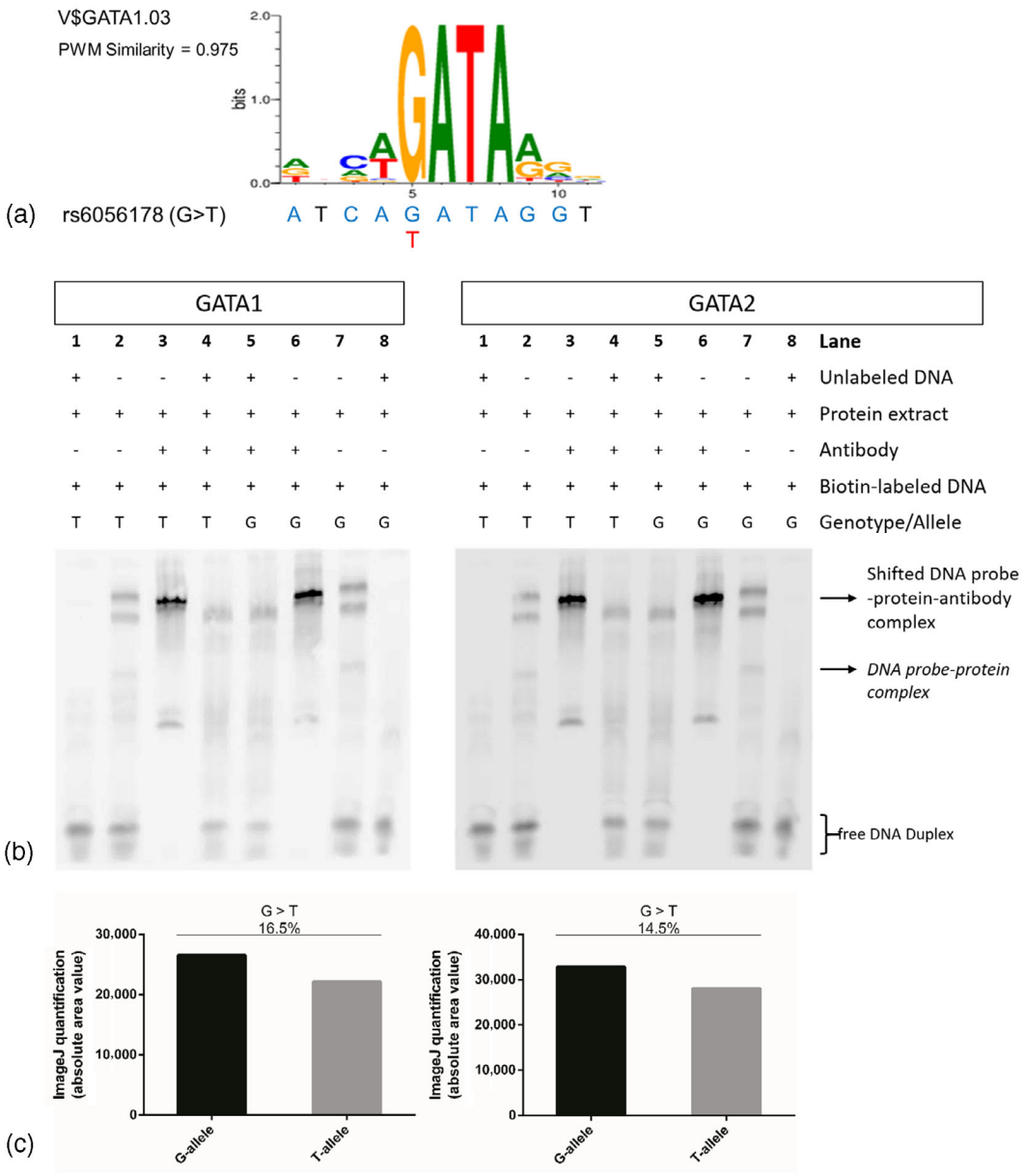

FIGURE 2 rs6056178 has allele-specific effects on GATA binding. (a) The DNA sequence at the common G-allele of rs6056178 shares a matrix similarity of 97.5% with the GATA transcription factor (TF) binding motif. The rare T-allele impairs the GATA motif. (b) GATA1 and GATA2 EMSAs performed with rs6056178 allele-specific oligonucleotide probes and nuclear protein extract from SaOs-2 cells. Binding of GATA antibodies to allele-specific probes is shown in lanes 1 and 4. Unlabeled DNA was added to verify that the band shift was antibody-specific in lanes 2 and 3. (c) Absolute value area of the antibody-specific bands. In the background of the rare T-allele of rs6056178, GATA1 and GATA2 binding to the allele-specific oligonucleotide probes were reduced 16.5% and 14.9%, respectively, compared to the common G-allele.

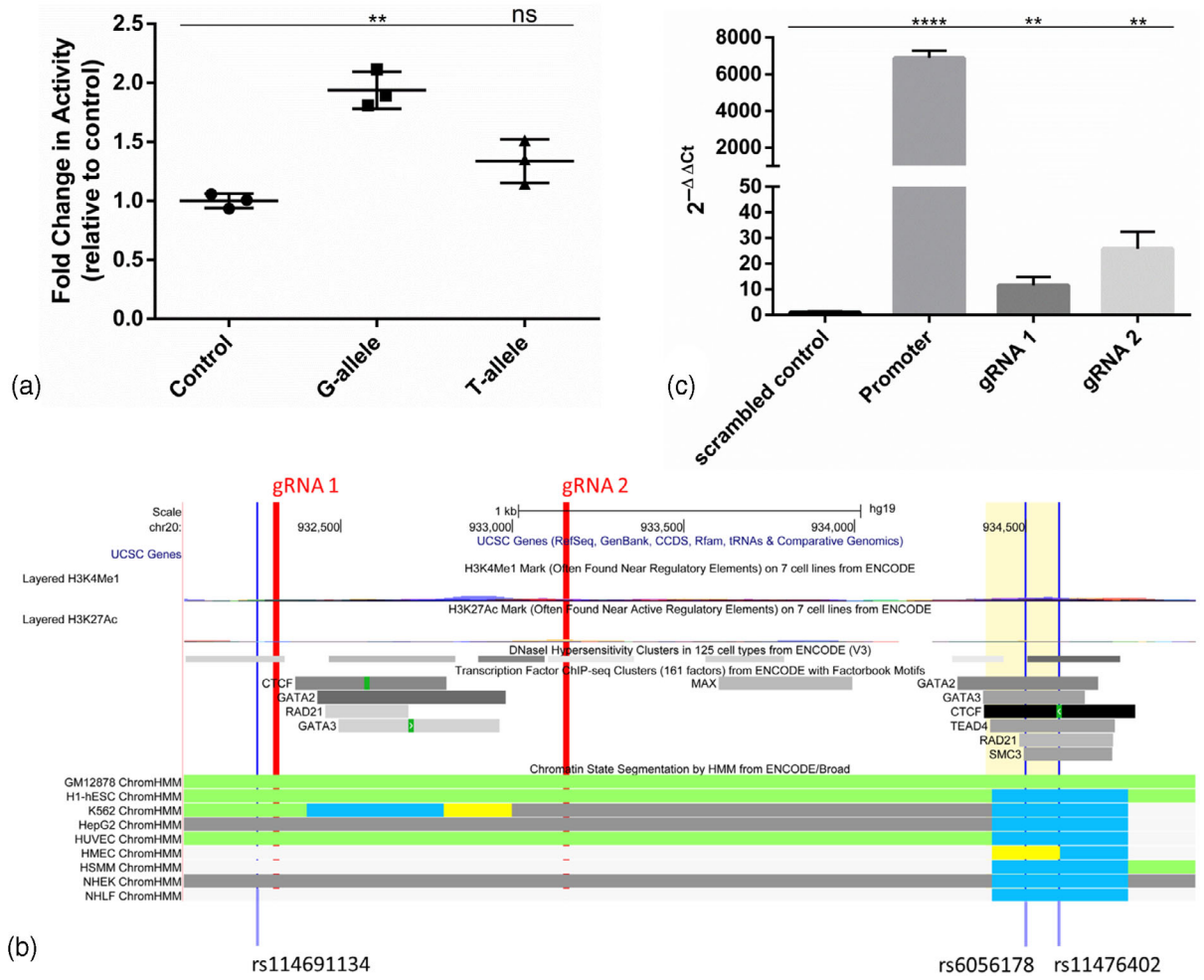

FIGURE 3 rs6056178-G activates gene expression and is located at a cis-regulatory element. (a) rs6056178 has allele-specific effects on luciferase activity. The 220-bp DNA sequence, spanning 110 bp upstream and downstream of the individual alleles of single nucleotide polymorphism (SNP) rs6056178 showed allele-specific enhancer activity in HeLa cells. In the background of the common rs6056178 G-allele, luciferase activity was increased (1.9 ± 0.2-fold, $p = .004$). The reporter sequence containing the rare T-allele showed no significant increase in luciferase activity. The difference between both alleles was significant with $p = .01$, indicating a reduced functional effect of the disease-associated rare T-allele. (b) The putative causal SNP rs6056178 locates to ENCODE DNAse I hypersensitivity sites, transcription factor binding sites experimentally confirmed by ChIP-Seq for various cell types and to human methylation marks (HMM) indicating a weak enhancer (yellow) in HMEC cells (HMEC, human epithelial cells isolated from female breast tissue). The indicated SNPs are in strong linkage disequilibrium with $r^2 > 0.8$. GATA2 and GATA3 binding at the sequence of rs6056178 was shown by ChIP-experiments (HMM colour codes: green = transcribed, grey = repressed, blue = insulator). (c) CRISPRa of the periodontitis-associated genomic region at rs6056178 showed a cis-regulatory effect on *R-spondin 4* (*RSPO4*) expression. CRISPRa of gRNA1, located 50 bp downstream of rs114691134, increased *RSPO4* expression 11.6 ± 3.3-fold compared to the unspecific control sgRNA ($p = .005$). gRNA2, 1 kb distance of the ChIP-Seq TF clusters tagged by rs6056178 and rs114691134, increased *RSPO4* expression 25.8 ± 6.5-fold ($p = .003$). gRNA targeting the *RSPO4* promoter increased *RSPO4* expression 6878 ± 394-fold ($p \leq .0001$), relative to the unspecific control sgRNA (error bars indicate 95% confidence interval, $^{**}p < .01$. $^{****}p < .0001$, transcript levels determined by quantitative real-time PCR).

3.4 | *RSPO4* activation regulates distinct genes and gene sets

Because *RSPO4* was significantly upregulated in mouse RILs that were resistant to alveolar bone loss caused by oral bacterial infection, we were interested in knowing the genes and gene sets in human cells that showed differential regulation after *RSPO4* upregulation. First, we tested whether HeLa cells provided an adequate model to interrogate the effects of *RSPO4* gene activation by showing that *RSPO4* activation significantly enriched the expression of a custom set of predicted functional partners of *RSPO4* ($p = .0045$), which were indicated by the STRING database (Szklarczyk et al., 2021) (Figure S3). Next, we contrasted *RSPO4* CRISPRa cells to scrambled sgRNA as controls and performed gene set enrichment analyses (GSEA) with the co-expression gene set tmod and the Molecular Signatures Database (MSigDB) gene sets reactome, hallmark, KEGG, and GO (Table S1). The most significant enriched gene sets of both tmod and GO were "type I interferon response" (FDR = 0.0006, area under the curve

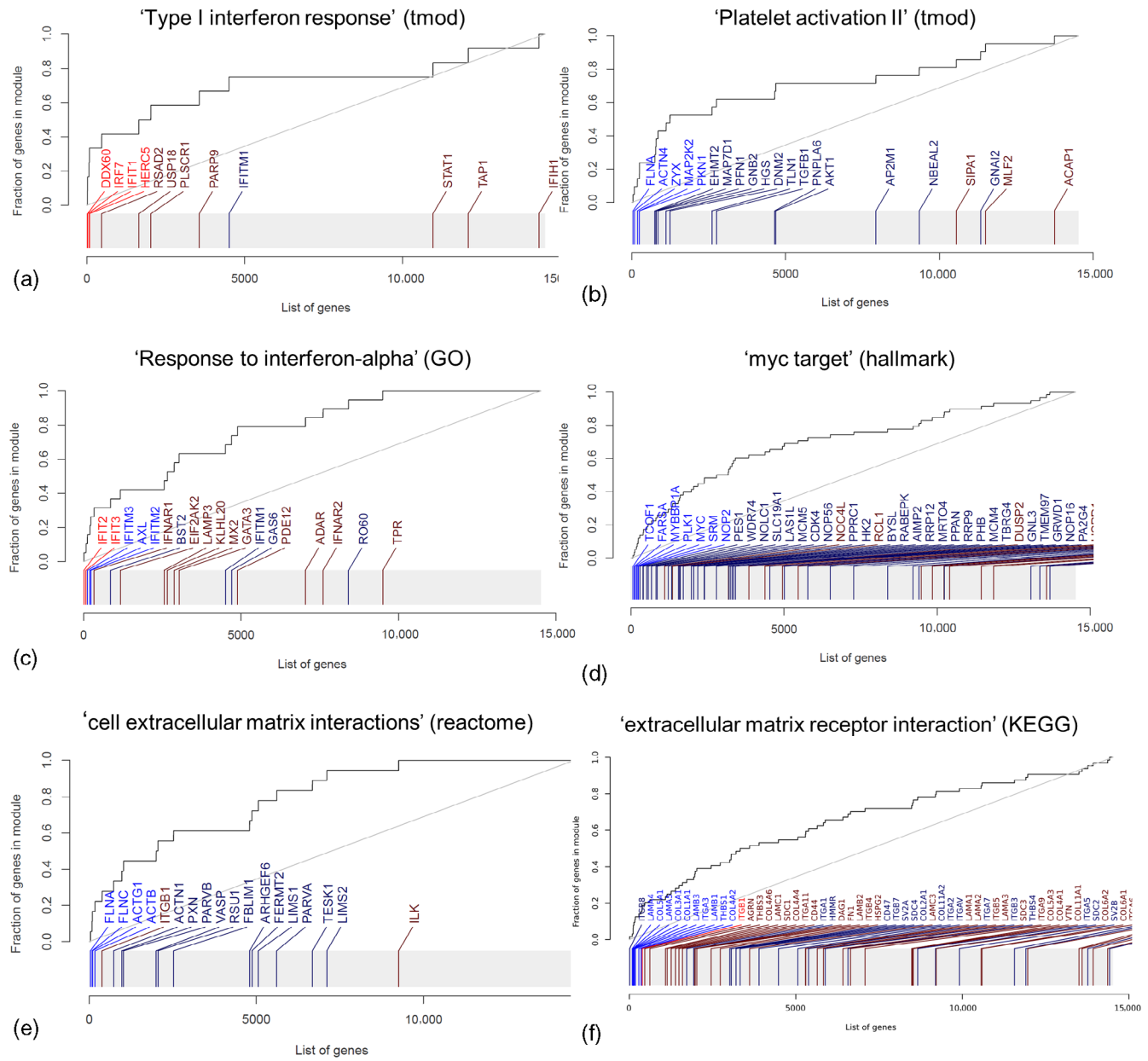

FIGURE 4 Gene set enrichment after *R-spondin 4* (*RSPO4*) upregulation in HeLa cells. Evidence plots (receiver operator characteristic curves) for the top gene sets from tmod, GO, hallmark, reactome, and KEGG. (a) tmod: "type I interferon response," FDR = 0.0006, AUC = 0.71. (b) tmod: "platelet activation II," FDR = 6×10^{-6}, AUC = 0.72. (c) GO: "response to interferon-alpha," FDR = 0.0002, AUC = 0.78. (d) Hallmark; "myc target," FDR = 4.5×10^{-7}, AUC = 0.7. (e) Reactome: "cell extracellular matrix interactions," FDR = 0.0002, AUC = 0.8. (f) KEGG: "extracellular matrix receptor interaction," FDR = 3.4×10^{-6}, AUC = 0.65. The grey rug plot underneath each curve corresponds to genes sorted by *p*-value, with the genes belonging to the corresponding gene sets highlighted in red (upregulated genes) or blue (downregulated genes). Bright red or bright blue indicates that the genes are significantly regulated. The area under the curve (AUC) corresponds to the effect size of the enrichment, with 0.5 being no enrichment and 1.0 being maximal possible enrichment. Data show three biological replicates for each experiment. CRISPR-activated cells with a gRNA that targeted the promoter were used for RNA-seq.

[AUC] = 0.71) and "response to interferon-alpha" (FDR = 0.0002, AUC = 0.78), respectively (Figure 4). tmod gene sets were derived from clustering expression profiles from human blood collected for various immune conditions, and we noted that the tmod gene set "platelet activation II" showed strong enrichment (FDR = 6×10^{-6}, AUC = 0.72). The top regulated hallmark gene set was "myc target" (FDR = 4.5×10^{-7}, AUC = 0.7). The most significant enriched gene sets of reactome and KEGG were "cell extracellular matrix interactions" (FDR = 0.0002, AUC = 0.8) and "extracellular matrix receptor interaction" (FDR = 3.4×10^{-6}, AUC = 0.65), respectively.

At the individual gene level, we found that the strongest upregulated gene was gliomedin (*GLDN*) with a fold change (FC) = 439 ($p < 5 \times 10^{-21}$) (Figure 5, Table 2). The second most upregulated gene was *MUCIN21* (*MUC21*), with 178 FC ($p = 1.4 \times 10^{-6}$). We also found

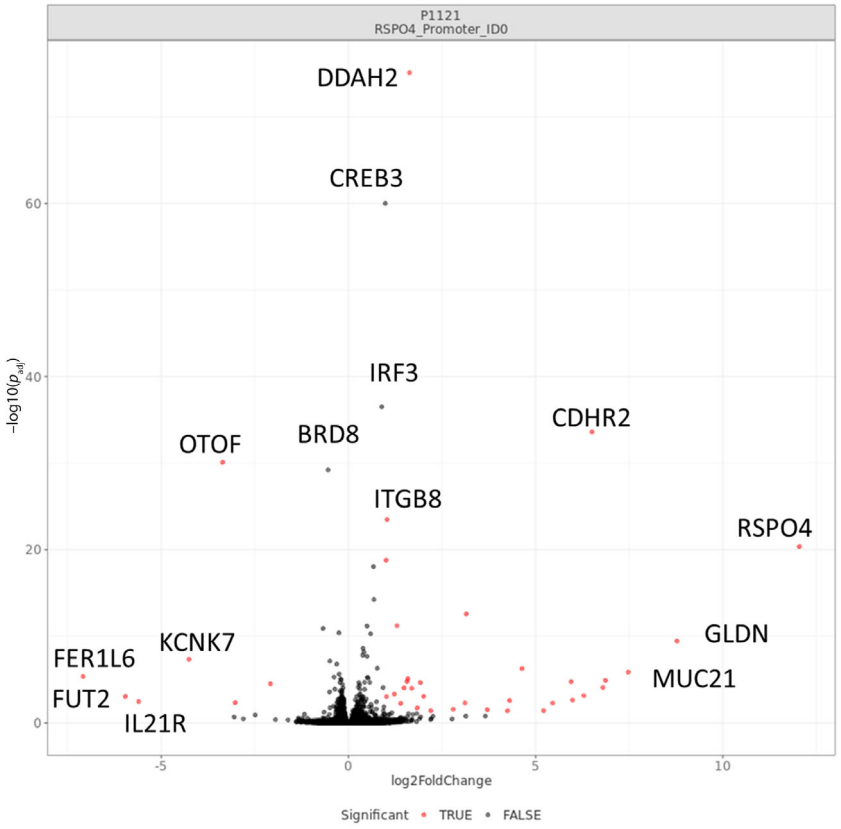

FIGURE 5 Volcano plot of differentially regulated genes after *R-spondin 4* (*RSPO4*) upregulation in HeLa cells.

TABLE 2 Top differentially expressed genes in HeLa cells after *RSPO4* gene activation

	Gene	$\log_2$Fold change	lfcSE	p_{adj} ($p < 0.005$)
Top upregulated (>6 lfc)	***RSPO4***	**12.046**	**1.181**	**4.2e−21**
	GLDN	8.777	1.209	3.5e−10
	MUC21	7.475	1.251	1.4e−6
	CDHR2	6.504	0.505	2.4e−34
	ANKRD62	6.287	1.342	.00070
	VWA2	5.986	1.375	.0023
Top downregulated (>2 lfc)	*FER1L6*	−7.088	1.228	4.3e−6
	FUT2	−5.961	1.287	.00086
	IL21R	−5.603	1.315	.0033
	KCNK7	−4.26	0.651	4.4e−8
	FER1L2	−3.362	0.275	7.7e−31
	CCDC187	−3.027	0.727	.0045
	PIK3AP1	−2.082	0.388	3e−5

Note: Bold lettering indicates that *RSPO4* was targeted by CRISPRa, resulting in strongest upregulation of all genes. *RSPO4* values were given here as an experimental control.
Abbreviation: lfc, logarithmic fold change.

several significantly downregulated genes (Table 2), but their effects were less strong compared to those of the upregulated genes. The strongest downregulation was shown by the gene *Fer-1 Like Family Member 6* (*FER1L6*, chr.8; FC $= -181$, $p = 4.3 \times 10^{-6}$). The second most downregulated gene was *Fucosyltransferase 2* (*FUT2*, FC $= -64$, $p = .0009$). The third most downregulated gene was *interleukin 21 receptor* (*IL21R*; FC $= -49$, $p = .003$). Notably, the fifth most downregulated gene was *FER1L2* (Otoferlin, *OTOF*, chr.2), a paralogue of *FER1L6*.

4 | DISCUSSION

In this study, we provided evidence for association of the gene *RSPO4*, an activator of canonical Wnt signalling, with stage III–IV, grade C periodontitis in different European populations, suggesting the involvement of this gene in the aetiology of severe rapid forms of periodontitis. We proved that *RSPO4* is the target gene of this association. *RSPO4* overexpression led most notably to the upregulation of *GLDN* and *MUC21*. *GLDN* is a reported risk gene of bone mineral

density (BMD) (Mullin et al., 2018), a trait that has been related to periodontitis susceptibility (Gil-Montoya et al., 2021). While the function of *GLDN* is largely unknown, an important paralogue is *OLFML2B* (Olfactomedin Like 2B). Gene ontology (GO) annotations related to this gene include extracellular matrix (ECM) binding. *MUC21* was reported in a genomewide expression profiling study to be one of the most upregulated genes in gingival wound healing (Wang & Tatakis, 2017) and the most upregulated mucin gene in gingival epithelial cells following infection with the protozoan parasite *Entamoeba gingivalis* (Bao et al., 2020). We observed that the effects on downregulation were weaker compared to the effects on upregulation, implying that *RSPO4* rather is an activating gene. *FER1L6* showed the strongest downregulation, and *FER1L2* (Otoferlin, *OTOF*, chr.2), an important paralogue of *FER1L6* encoded on a different chromosome, was the fifth most downregulated gene. While the function of *FER1L6* is largely unknown, *FER1L2* was reported to be involved in the transport of intracellular compartments to the cell membrane and subsequent vesicle membrane fusion (Heidrych et al., 2009). The second most downregulated gene was *FUT2*, which has a role in the fucosylation of host glycoproteins and glycolipids that mediate interaction with the host microbiota influencing its composition. *FUT2* is a risk gene of multiple early-onset diseases including inflammatory bowel diseases (Liu et al., 2015), childhood asthma (Ahluwalia et al., 2020), childhood diarrhoea (Bustamante et al., 2016), and psoriasis (Tsoi et al., 2017), to name a few. The third most downregulated gene was the interleukin 21 receptor (*IL21R*; log2FC $= -5.6$, $p = .003$), which transduces the growth-promoting signal of IL21. This gene is important for the proliferation and differentiation of T cells, B cells, and natural killer (NK) cells (provided by RefSeq, Jul 2010). Taken together, these findings indicate a novel, broad role of *RSPO4* in host barrier defence.

Interestingly, the literature has reported a genomewide association of *RSPO4* with response to antiviral therapy (Leger et al., 2014). In response to viral infection, human cells release interferons, a group of immune signalling proteins. Notably, the most significant enriched gene sets of tmod and GO are related to interferon response. In addition, reactome and KEGG suggested enrichment of gene sets that are involved in ECM interactions, mechanisms that involve cell proliferation and differentiation, cell migration, and cell adhesion. This implies broader roles of *RSPO4* during inflammation at the environment–body interface, which may encompass innate immune and ECM functions.

We found rs6056178 downstream of the coding region of *RSPO4* as a biological functional and putative causal variant that affects GATA binding. Additional evidence for this result was the observation of GATA2 and 3 binding to the chromatin at rs6056178, which was determined by ChIP-Seq experiments performed by ENCODE. We demonstrated allele-specific GATA1 and 2 binding at rs6056178 by EMSA. GATA TFs largely share the same motif ([A/T]GATA[A/G]) and GATA1 and 2 are theoretically able to bind at the motif found within the SNP sequence. Thus, we cannot discriminate which GATA TF binds in vivo. In the mature organism, GATA1 primarily regulates the expression of genes that mediate the development of red blood cells and platelets. GATA2 regulates many genes that are critical for self-renewal, maintenance, and functionality of blood-forming, lymphatic system-forming, and other tissue-forming stem cells. Loss of function causes haematological, immunological, or other disorders that progress to opportunistic infections, lung failure, or virus-induced cancers. GATA3 is critical for inflammatory and humoral immune responses and the functioning of the endothelium.

A limitation of this study is that we did not analyse the effects of *RSPO4* upregulation in gingival cells. However, we showed that *RSPO4* activation in HeLa significantly enriched the expression of predicted functional partners of *RSPO4*, implying that HeLa cells are a valid cell model. Another limitation is that we validated the cis effects of the associated region on *RSPO4* but not on other genes. It is possible that the associated regulatory elements also affect other genes both in cis and trans. For instance, the nearest gene to the association second to *RSPO4* is angiopoietin 4 (*ANGPT4*) ~40 kb downstream, with an agonistic function on angiogenesis (provided by RefSeq, Jul 2008) and a paralogue of *ANGPT1*, which was reported in the GWAS of de Coo et al. as a suggestive risk gene of stage III–IV, grade C periodontitis. Notably, a direct link of *RSPO4* with angiopoietin-like 4 (*ANGPTL4*) has been described (Kirsch et al., 2017). *RSPO4* acts as a ligand for the LGR6 receptor (Szenker-Ravi et al., 2018). Upon binding, LGR6 associates with phosphorylated co-receptor lipoprotein receptor-related protein 6 (LRP6), triggering the canonical Wnt-signalling pathway. *ANGPTL4* as a Wnt-signalling antagonist also binds to co-receptor LRP6, leading to the attenuation of Wnt/b-catenin signalling.

5 | CONCLUSION

In conclusion, we showed that rs6056178 is a functional and putative causal variant of severe progressive forms of periodontitis and affects GATA binding. *RSPO4* is a target gene of this association and its upregulation has effects on the expression of genes that involve host defence and barrier integrity, such as *MUC21* and *FUT2*.

ACKNOWLEDGEMENTS

This work was funded by an intramural research grant of the medical faculty, Charité – University Medicine Berlin, to Arne S. Schaefer and a research grant of the Hacettepe University Scientific Research Unit (THD-2021-19415) to H. G. Keceli and P. R. Dincer. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article. All authors gave final approval and agree to be accountable for all aspects of the work.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Gene Expression Omnibus at https://www.ncbi.nlm.nih.gov/geo. GEO accession number: GSE222437 and GEO link: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi.

ORCID

Huseyin G. Keceli https://orcid.org/0000-0001-6695-2133
Juan Blanco https://orcid.org/0000-0001-9251-513X
Henrik Dommisch https://orcid.org/0000-0003-1043-2651
Arne S. Schaefer https://orcid.org/0000-0001-5816-6765

REFERENCES

Ahluwalia, T. S., Eliasen, A. U., Sevelsted, A., Pedersen, C. T., Stokholm, J., Chawes, B., Bork-Jensen, J., Grarup, N., Pedersen, O., Hansen, T., Linneberg, A., Sharma, A., Weiss, S. T., Evans, M. D., Jackson, D. J., Morin, A., Krogfelt, K. A., Schjørring, S., Mortensen, P. B., … Bonnelykke, K. (2020). FUT2-ABO epistasis increases the risk of early childhood asthma and *Streptococcus pneumoniae* respiratory illnesses. *Nature Communications*, *11*(1), 6398. https://doi.org/10.1038/s41467-020-19814-6

Ambrosini, G., Groux, R., & Bucher, P. (2018). PWMScan: A fast tool for scanning entire genomes with a position-specific weight matrix. *Bioinformatics*, *34*(14), 2483–2484. https://doi.org/10.1093/bioinformatics/bty127

Bao, X., Wiehe, R., Dommisch, H., & Schaefer, A. S. (2020). *Entamoeba gingivalis* causes oral inflammation and tissue destruction. *Journal of Dental Research*, *99*(5), 561–567. https://doi.org/10.1177/0022034520901738

Bustamante, M., Standl, M., Bassat, Q., Vilor-Tejedor, N., Medina-Gomez, C., Bonilla, C., Ahluwalia, T. S., Bacelis, J., Bradfield, J. P., Tiesler, C. M. T., Rivadeneira, F., Ring, S., Vissing, N. H., Fink, N. R., Jugessur, A., Mentch, F. D., Ballester, F., Kriebel, J., Kiefte-de Jong, J. C., … Sunyer, J. (2016). A genome-wide association meta-analysis of diarrhoeal disease in young children identifies FUT2 locus and provides plausible biological pathways. *Human Molecular Genetics*, *25*(18), 4127–4142. https://doi.org/10.1093/hmg/ddw264

Caton, J. G., Armitage, G., Berglundh, T., Chapple, I. L. C., Jepsen, S., Kornman, K. S., Mealey, B. L., Papapanou, P. N., Sanz, M., & Tonetti, M. S. (2018). A new classification scheme for periodontal and peri-implant diseases and conditions—Introduction and key changes from the 1999 classification. *Journal of Clinical Periodontology*, *45*-(Suppl 20), S1–S8. https://doi.org/10.1111/jcpe.12935

Churchill, G. A., Airey, D. C., Allayee, H., Angel, J. M., Attie, A. D., Beatty, J., Beavis, W. D., Belknap, J. K., Bennett, B., Berrettini, W., Bleich, A., Bogue, M., Broman, K. W., Buck, K. J., Buckler, E., Burmeister, M., Chesler, E. J., Cheverud, J. M., Clapcote, S., & Complex Trait Consortium. (2004). The Collaborative Cross, a community resource for the genetic analysis of complex traits. *Nature Genetics*, *36*(11), 1133–1137. https://doi.org/10.1038/ng1104-1133

de Coo, A., Cruz, R., Quintela, I., Herrera, D., Sanz, M., Diz, P., Grandío, S. R., Vallcorba, N., Ramos, I., Oteo, A., Serrano, C., Esmatges, A., Enrile, F., Mateos, L., García, R., Álvarez-Novoa, P., Noguerol, B., Zabalegui, I., Blanco-Moreno, J., … Blanco, J. (2021). Genome-wide association study of stage III/IV grade C periodontitis (former aggressive periodontitis) in a Spanish population. *Journal of Clinical Periodontology*, *48*(7), 896–906. https://doi.org/10.1111/jcpe.13460

de Lau, W., Barker, N., Low, T. Y., Koo, B. K., Li, V. S., Teunissen, H., Kujala, P., Haegebarth, A., Peters, P. J., van de Wetering, M., Stange, D. E., van Es, J. E., Guardavaccaro, D., Schasfoort, R. B. M., Mohri, Y., Nishimori, K., Mohammed, S., Heck, A. J. R., & Clevers, H. (2011). Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. *Nature*, *476*(7360), 293–297. https://doi.org/10.1038/nature10337

Freitag-Wolf, S., Munz, M., Junge, O., Graetz, C., Jockel-Schneider, Y., Staufenbiel, I., Bruckmann, C., Lieb, W., Franke, A., Loos, B. G., Jepsen, S., Dommisch, H., & Schaefer, A. S. (2021). Sex-specific genetic factors affect the risk of early-onset periodontitis in Europeans. *Journal of Clinical Periodontology*, *48*(11), 1404–1413. https://doi.org/10.1111/jcpe.13538

Gaur, T., Lengner, C. J., Hovhannisyan, H., Bhat, R. A., Bodine, P. V., Komm, B. S., Javed, A., van Wijnen, A. J., Stein, J. L., Stein, G. S., & Lian, J. B. (2005). Canonical WNT signaling promotes osteogenesis by directly stimulating Runx2 gene expression. *The Journal of Biological Chemistry*, *280*(39), 33132–33140. https://doi.org/10.1074/jbc.M500608200

Gil-Montoya, J. A., Garrido-Martinez, M., Barrios-Rodriguez, R., Ramos-Garcia, P., Lenouvel, D., Montes-Castillo, C., & Martinez-Ramirez, M. J. (2021). Association between low bone mineral density and periodontitis in generally healthy perimenopausal women. *Journal of Periodontology*, *92*(1), 95–103. https://doi.org/10.1002/JPER.20-0029

Glinka, A., Dolde, C., Kirsch, N., Huang, Y. L., Kazanskaya, O., Ingelfinger, D., Boutros, M., Cruciat, C.-M., & Niehrs, C. (2011). LGR4 and LGR5 are R-spondin receptors mediating Wnt/beta-catenin and Wnt/PCP signalling. *EMBO Reports*, *12*(10), 1055–1061. https://doi.org/10.1038/embor.2011.175

Heidrych, P., Zimmermann, U., Kuhn, S., Franz, C., Engel, J., Duncker, S. V., Hirt, B., Pusch, C. M., Ruth, P., Pfister, M., Marcotti, W., Blin, N., & Knipper, M. (2009). Otoferlin interacts with myosin VI: Implications for maintenance of the basolateral synaptic structure of the inner hair cell. *Human Molecular Genetics*, *18*(15), 2779–2790. https://doi.org/10.1093/hmg/ddp213

Hoffman, M. M., Ernst, J., Wilder, S. P., Kundaje, A., Harris, R. S., Libbrecht, M., Giardine, B., Ellenbogen, P. M., Bilmes, J. A., Birney, E., Hardison, R. C., Dunham, I., Kellis, M., & Noble, W. S. (2013). Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Research*, *41*(2), 827–841. https://doi.org/10.1093/nar/gks1284

Kirsch, N., Chang, L. S., Koch, S., Glinka, A., Dolde, C., Colozza, G., Benitez, M. D. J., De Robertis, E. M., & Niehrs, C. (2017). Angiopoietin-like 4 is a Wnt signaling antagonist that promotes LRP6 turnover. *Developmental Cell*, *43*(1), 71–82.e76. https://doi.org/10.1016/j.devcel.2017.09.011

Lamont, R. J., Koo, H., & Hajishengallis, G. (2018). The oral microbiota: Dynamic communities and host interactions. *Nature Reviews. Microbiology*, *16*(12), 745–759. https://doi.org/10.1038/s41579-018-0089-x

Leger, P. D., Johnson, D. H., Robbins, G. K., Shafer, R. W., Clifford, D. B., Li, J., McLaren, P. J., & Haas, D. W. (2014). Genome-wide association study of peripheral neuropathy with D-drug-containing regimens in AIDS Clinical Trials Group protocol 384. *Journal of Neurovirology*, *20*(3), 304–308. https://doi.org/10.1007/s13365-014-0235-9

Liu, J. Z., van Sommeren, S., Huang, H., Ng, S. C., Alberts, R., Takahashi, A., Ripke, S., Lee, J. C., Jostins, L., Shah, T., Abedian, S., Cheon, J. H., Cho, J., Dayani, N. E., Franke, L., Fuyuno, Y., Hart, A., Juyal, R. C., Juyal, G., … Weersma, R. K. (2015). Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nature Genetics*, *47*(9), 979–986. https://doi.org/10.1038/ng.3359

Machiela, M. J., & Chanock, S. J. (2015). LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*, *31*(21), 3555–3557. https://doi.org/10.1093/bioinformatics/btv402

Manke, T., Heinig, M., & Vingron, M. (2010). Quantifying the effect of sequence variation on regulatory interactions. *Human Mutation*, *31*(4), 477–483. https://doi.org/10.1002/humu.21209

Manke, T., Roider, H. G., & Vingron, M. (2008). Statistical modeling of transcription factor binding affinities predicts regulatory interactions. *PLoS Computational Biology*, *4*(3), e1000039. https://doi.org/10.1371/journal.pcbi.1000039

Mullin, B. H., Zhu, K., Xu, J., Brown, S. J., Mullin, S., Tickner, J., Pavlos, N. J., Dudbridge, F., Walsh, J. P., & Wilson, S. G. (2018). Expression quantitative trait locus study of bone mineral density GWAS variants in human osteoclasts. *Journal of Bone and Mineral Research*, *33*(6), 1044–1051. https://doi.org/10.1002/jbmr.3412

Munz, M., Willenborg, C., Richter, G. M., Jockel-Schneider, Y., Graetz, C., Staufenbiel, I., Wellmann, J., Berger, K., Krone, B., Hoffmann, P., van

der Velde, N., Uitterlinden, A. G., de Groot, L. C. P. G. M., Sawalha, A. H., Direskeneli, H., Saruhan-Direskeneli, G., Guzeldemir-Akcakanat, E., Keceli, H. G., Laudes, M., … Schaefer, A. S. (2017). A genome-wide association study identifies nucleotide variants at SIGLEC5 and DEFA1A3 as risk loci for periodontitis. *Human Molecular Genetics*, *26*, 2577–2588. https://doi.org/10.1093/hmg/ddx151

Nashef, A., Matthias, M., Weiss, E., Loos, B. G., Jepsen, S., van der Velde, N., Uitterlinden, A. G., Wellmann, J., Berger, K., Hoffmann, P., Laudes, M., Lieb, W., Franke, A., Dommisch, H., Schäfer, A., Houri-Haddad, Y., & Iraqi, F. A. (2020). Translation of mouse model to human gives insights into periodontitis etiology. *Scientific Reports*, *10*(1), 4892. https://doi.org/10.1038/s41598-020-61819-0

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575.

Roider, H. G., Kanhere, A., Manke, T., & Vingron, M. (2007). Predicting transcription factor affinities to DNA from a biophysical model. *Bioinformatics*, *23*(2), 134–141. https://doi.org/10.1093/bioinformatics/btl565

Szenker-Ravi, E., Altunoglu, U., Leushacke, M., Bosso-Lefevre, C., Khatoo, M., Thi Tran, H., Naert, T., Noelanders, R., Hajamohideen, A., Beneteau, C., de Sousa, S. B., Karaman, B., Latypova, X., Başaran, S., Yücel, E. B., Tan, T. T., Vlaminck, L., Nayak, S. S., Shukla, A., … Reversade, B. (2018). RSPO2 inhibition of RNF43 and ZNRF3 governs limb development independently of LGR4/5/6. *Nature*, *557*(7706), 564–569. https://doi.org/10.1038/s41586-018-0118-y

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & von Mering, C. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, *49*(D1), D605–D612. https://doi.org/10.1093/nar/gkaa1074

Thomas-Chollier, M., Hufton, A., Heinig, M., O'Keeffe, S., Masri, N. E., Roider, H. G., Manke, T., & Vingron, M. (2011). Transcription factor binding predictions using TRAP for the analysis of ChIP-seq data and regulatory SNPs. *Nature Protocols*, *6*(12), 1860–1869. https://doi.org/10.1038/nprot.2011.409

Threadgill, D. W., Miller, D. R., Churchill, G. A., & de Villena, F. P. (2011). The collaborative cross: A recombinant inbred mouse population for the systems genetic era. *ILAR Journal*, *52*(1), 24–31. https://doi.org/10.1093/ilar.52.1.24

Tsoi, L. C., Stuart, P. E., Tian, C., Gudjonsson, J. E., Das, S., Zawistowski, M., Ellinghaus, E., Barker, J. N., Chandran, V., Dand, N., Duffin, K. C., Enerbäck, C., Esko, T., Franke, A., Gladman, D. D., Hoffmann, P., Kingo, K., Kõks, S., Krueger, G. G., … Elder, J. T. (2017). Large scale meta-analysis characterizes genetic architecture for common psoriasis associated variants. *Nature Communications*, *8*, 15382. https://doi.org/10.1038/ncomms15382

Wang, Y., & Tatakis, D. N. (2017). Human gingiva transcriptome during wound healing. *Journal of Clinical Periodontology*, *44*(4), 394–402. https://doi.org/10.1111/jcpe.12669

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Chopra, A., Song, J., Weiner, J. 3rd, Keceli, H. G., Dincer, P. R., Cruz, R., Carracedo, A., Blanco, J., Dommisch, H., & Schaefer, A. S. (2023). *RSPO4* is a potential risk gene of stages III–IV, grade C periodontitis through effects on innate immune response and oral barrier integrity. *Journal of Clinical Periodontology*, *50*(4), 476–486. https://doi.org/10.1111/jcpe.13758