First Insights Into the Molecular Basis Association Between Promoter Polymorphisms of the Il1b Gene and Helicobacter Pylori Infection in the Sudanese Population: Computational Approach

Abeer Babiker Idris (abeer.babiker89@gmail.com)
Department of Medical Microbiology,Faculty of Medical Laboratory Sciences, University of Khartoum https://orcid.org/0000-0001-6436-5963
Einas Babiker Idris
Department of Medical Microbiology, Rashid Complex Medical Clinics
Amany Eltayib Ataelmanan
Department of Medical Microbiology, Faculty of Medical Laboratory Sciences, University of Al-Jazirah
Ali Elbagir Ali Mohamed
Faculty of Medicine, University of Khartoum
Bashir M. Osman Arbab
Department of Internal Medicine, Modern Medical Centre
El-Amin Mohamed Ibrahim
Department of Medical Microbiology, Faculty of Medical Laboratory Sciences, University of Khartoum
Mohamed A. Hassan
Department of Bioinformatics, DetaVax Biotech

Research article

Keywords: H. pylori, IL1B, 5'- region, In silico analysis, Sudan

Posted Date: August 24th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-60531/v1

Version of Record: A version of this preprint was published on January 7th, 2021. See the published version at https://doi.org/10.1186/s12866-020-02072-3.

Abstract

Background:*Helicobacter pylori* (*H. pylori*) infects nearly half of the world's population with a variation in incidence among different geographic regions. Genetic variants in the promoter regions of *IL1B* gene can affect cytokine expression and creates a condition of hypoacidity which favors the survival and colonization of *H. pylori.*

Aim: To characterize the polymorphic sites in the 5'- region [-687_+297] of *IL1B* in *H. pylori* infection using in silico tools.

Methods: Genomic DNA was extracted from 121 gastric biopsies and genotyping of IL1B-31 polymorphism was performed using PCR-CTPP to investigate its association with the susceptibility to *H. pylori* infection in the Sudanese population. In addition, Sanger sequencing was applied to detect SNPs in the 5'-region [-687_+297] of *IL1B* in 14 *H. pylori*-infected patients; and bioinformatics analyses were used to predict whether these mutations would alter TFBSs or CEs in this region. Also, comparative analysis was conducted for this region in 11 species using ECR browser and Mulan search engine.

Results: A total of five nucleotide variations were detected in the 5'-regulatory region which led to the addition or alteration of the TFBSs and CEs. Genotyping of *IL1B*-31 C>T revealed a significant association between -31T and susceptibility to *H. pylori* infection in the studied population (P=0.0280). Comparative analysis showed conservation rates of *IL1B* upstream [−368_+10] region above 70% in chimpanzee, rhesus monkey, a domesticated dog, cow and rat.

Conclusion: In *H. pylori*-infected patients, three detected SNPs (-338, -155 and -31) located in the *IL1B* promoter were predicted to alter TFBSs and CE, thereby might affecting gene expression. These in silico predictions provide insight for further experimental *in vitro* and *in vivo* studies of the regulation of *IL1B* expression and its relationship to *H. pylori* infection. However, recognition of regulatory motifs by computer algorithms is fundamental for understanding gene expression patterns.

1. Background

Helicobacter pylori (*H. pylori*) is a Gram-negative, spiral-shaped and microaerophilic bacterium that infects nearly half of the world's population with a variation in incidence among different geographic regions.(1, 2) Epidemiological studies have indicated that the highest prevalence of *H. pylori* was found in Africa (79.1%), and the lowest prevalence in Northern America (37.1%) and Oceania (24.4%) with an overall global *H. pylori* prevalence of 44.3%, ranging from 50.8% in developing countries to 34.7% in developed countries.(3-6) The global annual recurrence rate of *H. pylori* was (4.3%) and it was found to be related to the human development index and prevalence of infection.(3) However, the clinical aspects of chronic infection with *H. pylori* vary from gastroduodenal inflammation and peptic ulceration to the most dangerous aspects, gastric carcinoma and mucosa-associated lymphoid tissue (MALT) lymphoma.(7, 8) Also, *H. pylori* may be implicated in several extra-gastric diseases such as iron deficiency anemia, idiopathic thrombocytopenic purpura, several dermatological disorders, hepatic encephalopathy, diabetes, pulmonary disease and cardiovascular.(7, 9) Indeed, the susceptibility to *H. pylori* infection and its diverse clinical presentation is determined by multiple factors including heterogeneity of *H. pylori* strains and their virulence factors, environmental factors, and the host genetic background, especially those regarding polymorphisms in certain cytokines, gene regulation and their receptor antagonist genes.(10-13) One of these cytokines is the *interleukin 1-beta* (*IL1B*) gene.

IL-1 family genes, spanning ~430 kb, cluster on chromosome 2q13-21 and consists of *IL-1A*, *IL-1B*, and *IL-1RN* genes which encode the pro-inflammatory cytokines IL-1α and IL-1β and the endogenous receptor antagonist IL-1ra, respectively.(14) IL-1β, the crucial cytokine in the gastrointestinal tract (15), has a variety of biological activities on a wide range of tissues and plays an important role in inflammatory, metabolic, physiologic, hematopoietic, and immunological processes (for a review, see references (16) and (17)). Because of the ability of IL-1β to inhibit gastric acid secretion it may have a profound effect on the natural history of *H. pylori* infection by allowing expansion of *H. pylori* colonization from the gastric antrum to the corpus.(15, 18-20) On a molar basis, IL-1β a 100-fold more potent inhibitor than PPIs, and 6000-fold more potent than H2 receptor antagonists.(21)

IL-1β is expressed at high levels in myeloid cell lineages in response to tissue injury and microbial invasion.(22, 23) Also, many different types of cells, e.g. B cells, T cells, NK cells, dendritic cells, fibroblasts and epithelial cells, express this protein in response to a broad range of stimuli and under inflammatory conditions but in much low level.(24-26) LPS-inducible IL1β expression is regulated by two regions, a proximal promoter that contains a TATA box and an upstream LPS-responsive enhancer (located between -3757 and -2729), which also known as the upstream inducible sequence.(27, 28) In monocyte, this promoter is packaged into a highly accessible chromatin structure which is characterized by the constitutive association of PU.1 and C/ EBPβ, but inducible association of RNA polymerase II.(24, 29, 30) The following multiple transcription factors, that constitutively and inducible associate with IL-1β regulatory regions have been identified: Spi-1/PU.1 (Spi1), NF-κB, C/EBPβ, AP-1, TBP, SSRP, or c-Jun and c-fos.(29)

Genetic variations in the promoter region of genes encoding cytokines were shown to correlate with individual differences in the expression of respective cytokine which may influence the intensity of the inflammatory response and susceptibility to many diseases.(31-35) *IL-1B* gene has two allelic variants (CT; dbSNP: rs16944) and (TC; dbSNP: rs1143627) which located at positions -511 and -31, respectively in the promoter region. These SNPs have been proposed to be associated with the susceptibility to *H. pylori* infection; and *H. pylori*-related gastric cancer and peptic ulcer disease (36), but it is still a contradictory topic of debate. Many studies have been published analyzing the contribution of *IL1B* promoter polymorphisms to *H. pylori* susceptibility with conflicting results explained, in part, by ethnic differences.(19, 36, 37). In the present study, genomic DNA Sanger sequencing was applied to detect SNPs in the region [-687_+297] of *IL1B* in *H. pylori*-infected patients; and bioinformatics analyses were used to study whether these mutations would alter transcription factor binding sites. Further computational analysis was also made to investigate other potential regulatory elements in this region. Finally, comparative profiling was conducted to assess the conservation of these genetic variations in 11 species. However, to our knowledge, the association between promoter polymorphisms of the *IL1B* gene and *H. pylori* infection in the Sudanese population has not been studied. It is imaginable that individual differences in *H. pylori* susceptibility or individual differences in *H. pylori*-related disease severity are linked to genetically determined differences in *IL1B* production. Therefore, studying the regulation of *IL1B* gene expression is of great significance.

2. Methods

2.1 Study Methodology

2.2 Study sitting and Study population

A prospective hospital-based study was carried out at public and private hospitals in Khartoum state. The hospitals included Ibin Sina specialized hospital, Soba teaching hospital, Modern Medical Centre and Al Faisal Specialized Hospital. A total of 121 individuals, who had been referred for endoscopy, were recruited. Out of that, 14 had gastric cancer, 27 had peptic ulceration, 64 had gastroduodenal inflammation, while 15 showed normal upper gastroduodenoscopic features. The diagnosis of gastroduodenal diseases had been made by an experienced gastroenterologist during the upper GI endoscopy procedure. While gastric cancer was diagnosed based on histology. Participant's demographic and clinical data were obtained by structured questionnaire, personal interviews, and review of case records. The selection criteria included the Sudanese population from both sexes, no antibiotic or NSAIDS uses. All the participants were informed with the objectives and purposes of the study and the written informed consents were taken. The demographic characteristic of participants is presented in Table 1.

Table 1. Demographic characteristic of participants

Variables		Total n=121	*H. pylori* (+ve) n=61	*H. pylori* (-ve) n=60	P-value
Age years ± Std. Deviation (range)		44.55±17.44 (15-89)	44.30± 16.76 (15-85)	44.82±18.24 (17-89)	0.9959
Gender	Male	72 (59.50%)	37 (51.38%)	35 (48.61%)	0.7046
	Female	49 (40.50%)	24 (48.98%)	25 (51.02%)	
Residence	Urban	54 (44.63%)	24 (44.44%)	30 (55.56%)	0.132
	Rural	67 (55.37%)	37 (55.22%)	30 (44.78%)	

2.3 DNA extraction

Gastric biopsies were collected in 400µl phosphate buffer saline (PBS). For histological examination, the biopsies were transported in formalin. DNA extraction was carried out by using innuPREP DNA Mini Kit (analytikjena AG, Germany) according to the protocol given by the manufacturer, as previously described in (38).

2.4 PCR amplification of specific *16S rRNA* of *H. pylori*

The specific *16S rRNA* gene of *H. pylori* was amplified by using the following primers (primers: F:5'-GCGCAATCAGCGTCAGGTAATG-3') (R:5'-GCTAAGAGAGCAGCCTATGTCC-3').(39) The PCR condition was previously described.(40)

2.5 PCR amplification and sequencing of the *IL1B* promotor region

The *IL-1B*-511 and -31, promoter polymorphisms, were amplified using the following primers: F:5'- CATCCATGAGATTGGCTAG-3' and R:5'-AGCACCTAGTTGTAAGGAAG-3'.(41) The cycling conditions were an initial denaturation at 94°C for 5min, followed by 35 cycles of 94°C for 1min, 60°C for 1min and 72°C for 1min, with a final extension at 72°C for 7min. The amplified PCR product is 800bp and was located between -687bp upstream and +297bp downstream of the *IL-1B* gene.

Out of 14 PCR products, of *H. pylori*-infected subjects, which have the clearest bands, were sent for DNA purification and Sanger dideoxy sequencing. Both DNA strands were sequenced commercially by Macrogen Inc, Korea.

2.6 Sequence analysis and SNPs detection

The sequencing results, two chromatograms for each patient (forward and reverse), were visualized, checked for quality and analyzed using the Finch TV program version 1.4.0.(42) The nucleotide Basic Local Alignment Search Tool (BLASTn; https://blast.ncbi.nlm.nih.gov/) was used to assess nucleotide sequence similarities.(43)

To determine the SNPs in the *IL-1B* promoter region, multiple sequences alignment (MSA) for tested sequences with a reference sequence (NG_008851) were performed by using BioEdit software.(44)

2.7 Bioinformatics analysis of the IL-1B promoter region in *H. pylori*-infected subjects

2.7.1 in silico prediction of the promoter

The crucial element for initiating and regulating messenger RNA transcription is the promoter sequence which is generally located in the 5' upstream region of a structural gene.(45) In this study, five computational promoter recognition tools were used: Berkeley Drosophila Genome Project (BDGP)

(http://www.fruitfly.org/),(46) Promoter 2.0 Prediction Server (http://www.cbs.dtu.dk/),(47) FPROM, TSSG, and TSSW (http://softberry.com/) which are based on neutral network and linear discriminant approach.(48)

2.7.2 In silico analysis of the predicted promoter regions

2.7.2.1 Assessment for the presence of promoter associated features

In silico predicted promoter regions were additionally assessed for the presence of promoter associated features, including promoter-associated histone marks, broad chromatin state segmentation, transcription factor ChIP-seq, and DNase I hypersensitivity clusters, using the ENCODE data (https://epd.epfl.ch/cgi-bin/get_doc?db=hgEpdNew&format=genome&entry=IL1B_1).(49-51)

2.7.2.2 Prediction of CpG Islands

A CpG island is often regarded as a marker for the initiation of gene expression. It is a segment of DNA with high GC and CpG dinucleotide contents which located in the 5' UTR (untranslated regions) of genes. In this study, MethPrimer(45, 52) and GpC finder software(http://www.softberry.com/berry.phtml?topic=cpgfinder&group=programs&subgroup=promoter) were employed to predict CpG islands in the promoter.

2.7.2.3 Prediction of Transcription Factor Binding Sites

One of the important steps in the chain of promoter analytical events is the prediction of the potentially functional transcription factor binding sites (TFBSs). Protein binding sites in a promoter represent the most important elements and the corresponding proteins are called transcription factors (TFs). In this step, the promoter region was analyzed for possible TFBSs using the following prediction software: Alggen Promo (http://alggen.lsi.upc.es/cgi-bin/promo_v3/),(53, 54) AliBaBa2 (http://www.gene-regulation.com/),(55) GPMiner (http://GPMiner.mbc.nctu.edu.tw/)(56), TF-Bind (http://tfbind.hgc.jp/),(57) and Tfsitescan (http://www.ifti.org).(58) These software were employed to predict possible TFBSs in the promoter region and corresponding TFs using binding sites from TRANSFAC®Public which provides data on eukaryotic TFs, their experimentally-proven binding sites, consensus binding sequences (positional weight matrices) and regulated genes.(59)

2.7.2.4 Prediction of composite regulatory elements

Composite regulatory element (CE) is the minimal functional unit, which can provide combinatorial transcriptional regulation of gene expression. Structurally a CE consists of two closely located DNA binding sites for distinct transcription factors. But its regulatory function is qualitatively different from regulation effects of either individual DNA binding sites. In this study, we identified the composite regulatory elements in our region by using MatrixCatch algorithm.(60)

2.7.2.5 Comparative analysis

Promoter region was analyzed for possible conservation using the ECR Browser (http://ecrbrowser.dcode.org),(61) NCBI BLASTn (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and ClustalW(https://www.genome.jp/tools-bin/clustalw). Conservation was assessed in 11 species: chimpanzee (Pan troglodytes), rhesus monkey (Macacamulatta), mouse (Mus musculus), rat (Rattusnorvegicus), dog (Canisfamiliaris), cow (Bostaurus), opossum (Monodelphisdomestica), chicken (Gallus gallus), frog (Xenopuslaevis), zebrafish (Danio rerio), fugu pufferfish (Takifugurubripes), and spotted green pufferfish (Tetraodon nigrovoridis).

Also, conservation of SNPs was evaluated and the possible conservation of TFBSs at these SNP locations was screened with Multiple-sequence local alignment and visualization (Mulan) search engine (https://mulan.dcode.org/).(62)

2.8Detection of the *IL-1B*-31 C/T polymorphism using PCR with confronting two-pair primer (PCR-CTPP)

For detection of the *IL-1B*-31 polymorphism, PCR-CTPP was applied. The primers for the C allele were (F:5'-ACT TCT GCT TTT GAA GGC C-3') and (R:5'-TAG CAC CTA GTT GTA AGG A-3'); and those for the T allele were (F:5'-AGA AGC TTC CAC CAA TAC T-3') and (R:5'-CTC CCT CGC TGT TTT TAT A-3').(63) 1μl of extracted DNA was used in a 25μl reaction mixture with a prepared Maxime PCR PreMix Kit (*i*-Taq) (iNtRON BIOTECHNOLOGY, Seongnam, Korea), 23μl of de-ionized sterile water, 0.25μl of each primer. PCR conditions were as follow: 5 min of initial denaturation at 94°C, followed by 25 cycles of 1min at 94°C, 1min at 54°C, and 1min at 72°C, and a 5min final incubation at 72°C. The PCR products were visualized by electrophoresis on a 2% agarose gel stained with ethidium bromide. Genotyping was performed as follows; 240, 155 bp for CC genotype, 240, 155, 122 bp for CT genotype, and 240, 122 bp for TT genotype.(63)

2.9 Statistical analysis

Deviations from Hardy-Weinberg equilibrium in control were examined by χ^2test. According to prevalence *of H. pylori* infection, differences in distribution by age was assessed by Mann-Whitney test, while differences in distribution by categorical variables were examined by χ^2test or *Fisher's* test. Odds ratios (ORs) were calculated and reported within the 95% confidence intervals (CIs). P<0.05 was considered to be statistically significant. The statistical analyses were performed using the GraphPad Prism 5.

3. Results

3.1 Nucleotide variations in the 5'-regulatory region of the *IL1B* gene

In Sudanese *H. pylori* infected patients, a total of five nucleotide variations were detected in the 5'-regulatory region. Among which, 3 are bimodally mutated heterozygous SNPs, and they were newly discovered; the positions of these three SNPs are -338, -155 and +38. The others two SNPs were rs16944 and

rs1143627, see Table 2 and Figures 4b, 5b, 6b and 6c for more illustration. The nucleotide sequences of the *IL1B* 5'- region [-687_+297] were deposited in the GenBank database under the following accession numbers: from MT767762 to MT767775.

Table 2. Nucleotide changes in the 5'-regulatory [-687_+297] region of the *IL1B* gene in *H. pylori*-infected patients

SNP	Event	SNP Locus	Chromosome Position	Serial Number (rs)
T>C (heterozygous)	Transition	+38	NG_008851.1:g.5039	-
C>T (heterozygous and homozygous)	Transition	-31	NG_008851.1:g.4970	rs1143627
G>C (heterozygous)	Transversion	-155	NG_008851.1:g.4844	-
A>T (heterozygous)	Transversion	-338	NG_008851.1:g.4666	-
T>C (heterozygous and homozygous)	Transition	-511	NG_008851.1:g.4490	rs16944

3.2 in silico prediction of the *IL1B* promoter regions

Five types of promoter prediction programs were employed to predict the promoter regions of *the IL1B 5'- region* [-687_+297] and the results are presented in Table 3. The Promoter 2.0 Prediction Server predicted no promoter region. Neural Network Promoter Prediction (NNPP version 2.2) predicted three promoter regions, located at -328bp, -124bp and +1bp relative to the *IL1B* translational start codon (transcript NM_000576.3), whose prediction scores were 0.97, 0.60 and 0.96, respectively. While FPROM, TSSG and TSSW programs predicted one promoter region, +1 bp, which is the only region predicted by more than one prediction program.

Table 3. Overview of the computational predicted *IL1B* promoter regions for the respective prediction programs. The most predicted region is indicated in bold.

Prediction program	Predicted promoter regions
BDGP (NNPP version 2.2)**	-368 _ -319
	-164 _ -115
	-40 _ +10
FPROM	**+1**
Promoter 2.0 Prediction Server	no promoter
TSSG***	**+1**
TSSW****	**+1**

*All positions are given in base pairs relative to the translational IL-1B start codon (transcript NM_000576.3).

**Neural Network Promoter Prediction, Threshold 0.50

***Threshold for LDF- 4.00

****Thresholds for TATA+ promoters - 0.45, for TATA-/enhancers - 3.70

3.3 In silico analysis of predicted *IL1B* promoter regions

ENCODE data showed high level of DNase I hypersensitivity, promoter associated histone modifications and transcription factor occupancy patterns at -124 and +1 bp promoter regions. While no Nuclease hypersensitivity around -328 bp region.

3.3.1 Prediction of CpG island

MethPrimer and CpGFinder software were used to predict the presence of CpG islands. The search parameter values for the software were CpG island length >100bp, CG% > 50%, and Obs/Exp > 0.6. However, both software predicted no CpG islands presence in the predicted promoter regions, see Figure 2. Also, ENCODE data confirmed no presence of the CpG island in the predicted promoters.

3.3.2 Conservancy of the predicted promoters

The ECR Browser revealed mammalian conservation for the [−368_+10] region in chimpanzee (Pan troglodytes - pan-Tro2) (99.2%), rhesus monkey (Macaca mulatta - rheMac2) (93.4%), domesticated dog (Canis lupus familiaris - canFam2) (73%), cow (Bos Taurus - bosTau3) (71.1%), rat (Rattus norvegicus - rn4) (70.5%) and mouse (Mus musculus - mm9) (69.3%). But the region was not conserved in opossum (Monodelphis domestica), chicken (Gallus gallus), frog (Xenopus laevis), zebrafish (Danio rerio), fugu pufferfish (Takifugu rubripes), and spotted green pufferfish (Tetraodon nigrovoridis), see Figure 3 for more illustration.

3.3.3 Prediction of TFBSs

In this study, five programs were used to predict TFBSs and to insure proper analysis we only selected factors that are predicted by three out of the four programs or the factors predicted by two programs but verified in the literature. The five prediction programs reported multiple putative TFBSs within the [−368_+10] region, see Table 4 and Figure 4c, 5a and 6a. However, screening of this region, by using the NCBI SNP databases (dbSNPs), revealed the presence of 9 SNPs upstream of the *IL1B* core promoter region which are shown in Table 4. The ECR Browser and NCBI BLASTn showed the conservation of these SNPs in chimpanzee, rhesus monkey, cow and dog. Mulan revealed multiple TFBSs to be located at rs749558279, rs140623868 and -338A>T. The overview of conserved TFBSs predicted by Mulan to be conserved (100%) between human, chimpanzee, rhesus monkey, cow and domesticated dog is summarized in Table 6.

3.3.4 Prediction of composite regulatory elements

MatrixCatch was used to find known regulatory elements (both single sites and pairs) which verified experimentally. Also it found novel regulatory elements by computational comparison but without experimental verification on functionality. These elements found by using similarity to known ones in a library of CE models.(59) The summary of predicted CEs by MatrixCatch is presented in Table 5.

Table 4. Summary of the in silico predicted TFBSs for [−368_+10] region.

Transcription factor	Position*		SNP in binding site	Software				
	Start	End		AliBaba2.1	Alggen Promo	Tfsitescan	TF-Binding	GPMiner
GR	-345	-327	rs140623868 NG_008851.1:g.4666 A>T#	X	X	X	X	
NF-kappaB	-297	-288		X	X	X	X	
c-Myb	-269	-260	rs769027934		X	X	X	
Oct-1	-233	-225		X		X	X	
GATA-1	-227	-214	rs749558279	X	X	X	X	X
GR	-207	-202			X	X	X	
GATA-1	-164	-155	rs866837107 NG_008851.1:g.4844 G>C#	X	X	X		
Sp1	-149	-132	rs74579367	X		X		
AP-2	-148	-137				X	X	
MAZ	-147	-137			X	X		X
NF-Y	-126	-114			X	X	X	
NF-AT	-116	-105			X			X
Spi-1/PU.1	-115	-97		X		X		
STAT1	-108	-101			X		X	X
IRF	-106	-99			X	X		
C/EBPalpha	-94	-83			X		X	X
NF-kappaB	-70	-61	rs4986962	X			X	
HSTF/HSF	-69	-60		X			X	
C/EBP	-67	-58		X			X	
Spi-1/PU.1	-50	-39		X		X		
TBP	-31	-25	rs1143627#		X		X	X
c-Myb	-14	-5			X		X	X
NF-1	-13	-4		X	X		X	
GATA-1	-7	6	rs529869449		X	X	X	X

* Positions are given relative to the translational *IL1B* start codon (transcript NM_000576.3)

SNPs observed in this study

Table 5. Summary of the in silico predicted CE for [−368_+10] region.

CE	MatrixName for a 1st element	S_1	Distance in between	MatrixName for a 2nd element	S_2	Position*	orientation	CS	P-Value	Sequence
CE00047	V$POU1F1_Q6	0.935	-1	V$OCT1_04	0.949	-239	-	-0.004	7.72e-06	ATGCATATTTGCATGGTGATA(
CE00058	V$NFKB_Q6_01	0.937	5	V$HMGIY_Q6	0.995	-300	+	-0.46	1.22e-05	ACGTGGGAAAAT
CE00158	V$OCT_C	0.907	10	V$AP1_01	0.869	-236	-	-0.232	1.88e-04	CATATTTGCATGGTGATACATT
CE00058	V$NFKB_Q6_01	0.791	6	V$HMGIY_Q6	0.995	-301	+	-0.181	1.05e-03	AACGTGGGAAAAT
CE00186	V$ETS_Q6	0.936	13	V$CEBPA_01	0.942	-107	-	-1.049	6.55e-03	CTTTCCTTTaactTGATTGTGA
CE00249	V$IRF_Q6	0.924	5	V$PU1_Q6	0.839	-111	+	-0.314	1.01e-02	TCCCCTTTCCTTT
CE00078	V$GR_Q6	0.914	39	V$CEBPB_02	0.799	-377	+	0.049	2.37e-02	GAAGAAAAGTATGTGCATGTa
CE00186	V$ETS_Q6	0.896	20	V$CEBPA_01	0.942	-120	-	-0.793	2.73e-02	TTTTCCCCTttcctttaactTGATT
CE00135	V$ETS_Q6	0.781	25	V$MYB_Q5_01	0.974	-292	-	0.002	4.31e-02	AAATCCAGTattttaatgtggacatC

CE - Composite regulatory element, $S_{1,2}$ - PWM scores for the first and second elements, CS - composite score

*Beginning of the element relative to TSS

Table 6. Conserved TFBSs predicted by Mulan (https://mulan.dcode.org/) to be conserved (100%) between human, chimpanzee, rhesus monkey, cow and domesticated dog at the [−368_+10] region. The most important transcription factor Sp1 for *IL1B* transcription was not included in this list.

Transcription Factor	Binding sequence	Position*		Conserved SNPs
		Start	End	
HNF4_01	TCTTCCACTTTGTCCCACA	-344	-326	rs140623868 [-335]
DR1_Q3	TCCACTTTGTCCC	-341	-329	-338A>T#
PPAR_DR1_Q2	TCCACTTTGTCCC	-341	-329	
NFKAPPAB65_01	GGAAAATCCA	-295	-286	
OCT1_01	TGCATATTTGCATGGTGAT	-236	-220	rs749558279 [-224]
OCT1_Q5_01	ATATTTGCATG	-235	-225	
OCT_Q6	ATATTTGCATG	-235	-225	
OCT1_B	TATTTGCATG	-234	-225	
AP1_Q2_01	CAAAATGTGTCA	-212	-201	
ZTA_Q2	TGTGTCATAGTTT	-207	-196	
CEBPDELTA_Q6	GATTGTGAAATC	-93	-82	
BARBIE_01	ACTTCTGCTTTTGAA	-49	-35	

*All positions are given in base pairs relative to the translational IL1B start codon (transcript NM_000576.3).

Observed in this study.

3.4 Allele and genotype frequencies of *IL1B*-31 and susceptibility to *H. pylori* infection

The molecular detection of *H. pylori* among participants was 50.41% and no significant differences were noted for *H. pylori* infection rates by age, gender, and residence, see Table1. Sixty-one patients and sixty uninfected controls were successfully genotyped for the *IL1B*-31 C/T polymorphism (Figure 4a). The frequency of allele T and *IL1B*-31 T/C + T/T genotypes were significantly higher in *H. pylori*-infected individuals compared to uninfected controls (40.16% versus 23.33%, P=0.0058 and 54.1% versus 33.33%, P=0.0280, resp.). The allelic and genotype distributions of *IL1B*-31 C/T polymorphism followed those expected in HWE for control population (P=0.08). (Table 7)

Table 7. Allele and genotype frequencies of *IL1B*-31 polymorphism among *H. pylori* infected and uninfected subjects, and their contributions to *H. pylori* infection

	H. pylori (+ve) n=61	*H. pylori* (-ve) n=60	OR (95% CI)*
Allele frequency			
IL1B-31-C	73(59.84%)	92(76.67%)	0.4534 (0.2598-0.7913)
IL1B-31-T	49(40.16%)	28(23.33%)	
P-value	**0.0058**		
Genotype frequency			
C/C	28(45.90%)	40(66.67%)	1 (reference)
T/C	17(27.87%)	12(20%)	0.4941 (0.2043 - 1.195)
T/T	16(26.23%)	8(13.33%)	0.3500 (0.1318 - 0.9295)
T/C + T/T	33(54.1%)	20(33.33%)	0.4242 (0.2032 - 0.8858)
C/C * C/T+T/T P-value	**0.0280**		

*OR, odds ratio; 95%CI, 95% confidence interval.

4. Discussion

Genetic variants in the promoter regions of *IL1B* gene can affect cytokine expression and creates a condition of hypoacidity which favors the survival and colonization of *H. pylori*.(15, 36) In the present study we functionally analyzed SNPs in the *IL1B* 5'-region [-687_+297] of Sudanese patients infected with *H. pylori* and developed divergent clinical outcomes. We observed three novel mutations (-338, -155 and +38) and interestingly, two of them (-338 and -155) were located at promoter regions predicted by NNPP algorithm. Thus, these mutations might play a role in regulating the expression of *IL1B*. However, in this study, the computational analysis predicted three promoter regions at -328, -124 and +1. Nuclease hypersensitivity and histone modifications are characteristic for cis-regulatory regions such as promoters. The ENCODE data showing these hallmarks to be present in the putative promoter region at the +1 bp region. Moreover, the upstream region around -124 bp showed some of these characteristics, although to a lesser degree. While, region around -328 bp showed only histone marks.(49-51) Also, In silico comparative analysis showed the [−368_+10] region to be mammalian conserved, with conservation rates above 70% in chimpanzee, rhesus monkey, a domesticated dog, cow and rat. This conservation might indicate a possible regulatory role for this region (Figure 3). But the region was not conserved in opossum, chicken, frog, zebrafish, fugu pufferfish, and spotted green pufferfish; it is possible that the regulation of *IL1B* in these species is controlled by a different mechanism or pathway.

Regulation of gene transcription depends on the interaction between TFs and TFBSs. Any changes in these sites may develop significant effects on the binding of TFs to regulatory sequences and then the expression products of genes.(45, 48, 64) In this study, an in silico-based prediction analysis using different algorithms indicated that the transcription factors NF, C/EBP, Spi-1/PU.1, NF-kappaB, AP-1, TBP, IRFs and STAT, c-Myb and GATA-1 are involved in the regulation of *IL1B* gene expression and have the potential to bind in the polymorphic regions (Table 5). This indication is in agreement with the results of previous studies.(26, 65, 66) The two novel SNPs located in the promoter regions led to the addition or alteration of the TFBSs. As illustrated in Table 8, −338 (A>T) polymorphism resulted in the alteration of GR to PU.1 and the −155(G>C) polymorphism led to an addition of a C/EBPbeta. T allele in position -31 instead of C allele resulted in an addition of RSRFC4 protein, which is partially in agreement with experiments that assessing allele-specific oligonucleotides for -31. These experiments reported that there were one or more TFs resulting in a fivefold increase in DNA binding on the *IL1B*-31T oligonucleotide after LPS stimulation. These TFs may be unable to interact with the C-bearing *IL1B*-31 allele to form the transcription initiation complex.(67)

Table 8. Variations in transcription factors before and after nucleotide changes (-338, -155 and -31) in the *IL1B* gene by using AliBaba2.1 software.

SNP Site	Base Group	TF	TFBSs Sequence	TF Position
-338	A	GCR1	CTTCC**A**CTTT	-343 _ -334
	T	GR	C**A**CTTTGTCC	-339 _ -330
		PU.1	CTTCC**T**CTTT	-343 _ -334
		GCR1		
-155	G	Zen-1	ATTATCTCA**G**	-164 _ -155
	C	GATA-1	TTATCTCA**G**T	-163 _ -154
		C/EBPbeta	ATTATCTCA**C**	-164 _ -155
		Zen-1	TTATCTCA**C**T	-163 _ -154
		GATA-1		
-31	C	-	-	-
	T	RSRFC4	GC**T**ATAAAAA	-33 _ -24

The extensively studied SNPs in relation to *H. pylori* infection (-31 and -511) were also detected in our patients. We observed a significant association between -31T and susceptibility to *H. pylori* infection in the studied population (P=0.0280). This result is in concordance with a number of studies conducted in different ethnic groups that showed an association between *IL1B*-31T and *H. pylori* infection.(67-72) Also, there are a some studies found a negative association,(33, 72, 73) and this variation could be due to differences in genetic backgrounds of the studied population, method of genotyping and sample size.(36, 74) Interestingly, we found that the T-511C SNP was not located in the in silico-predicted promoter regions, hence it could not affect the expression of *IL1B*. While -31 which involves a TATA-box could directly affect the induction of *IL1B*. These findings are in agreement with the result obtained by Al-Omer *et al.* that the -31 polymorphism was markedly affecting DNA-protein interactions *in vitro* while -511 does not alter *in vitro* protein secretion and its effect may be mediated by linkage disequilibrium (LD) with -31.(67) Also, R Kimura *et al.* found that the expression of the -31T allele was 2.2 times of the -31C allele and this higher transcription efficiency may correspond to the fact that C-31T is located in a TATA box.(75) In contrast, other observations of IL-1β production have suggested that there was no significant association between the known allelisms in the *IL-1B* gene and IL-1β induction *in vitro* and that the -31C and was the higher expressing allele *in vivo*.(75-77). However, the production of IL1β is affected by several factors besides gene polymorphisms such as epigenetic conditions and other genetic backgrounds. And to exclude the influence of trans-acting factors which are able to confound the effects of the polymorphisms, the allele-specific transcript quantification coupled with haplotype analysis (75, 78) is recommended to identify the cis-acting effect of T-511C polymorphism and our novel detected polymorphisms (-338 and -155) on the *IL1B* transcription and susceptibility to multifactorial diseases including *H. pylori* infection.

However, recognition of regulatory motifs by computer algorithms is fundamental for understanding gene expression patterns, cell specificity and development.(48) Identifying SNPs that might be a genetic modifiers in *IL1B* gene may be valuable in preventive, diagnostic, and therapeutic strategies against the incidence and progression of *H. pylori* infection. This study revealed three nucleotide variations in the *IL1B* 5'-region which possibly leading to modification of transcriptional regulation in *H. pylori* infection, however, this conclusion requires further *in vitro* and *in vivo* validation in subsequent studies.

Conclusion

In *H. pylori*-infected patients, three detected SNPs located in the *IL1B* promoter were predicted to alter CEs and TFBSs, thereby might affecting gene expression. This computational analysis provide insight for further experimental *in vitro* and *in vivo* studies of the regulation of *IL1B* expression and its relationship to *H. pylori* infection. However, recognition of regulatory motifs by computer algorithms is fundamental for understanding gene expression patterns.

Abbreviations

16S rRNA, 16Svedberg ribosomal RNA; AP-1, activator protein 1; BDGP, Berkeley Drosophila Genome Project; BLASTn, The nucleotide Basic Local Alignment Search Tool; bosTau3, Bos Taurus (cow) full genome as provided by UCSC (Jan 2008); bp, base pair; C/ EBPβ, CCAAT-enhancer binding protein beta; C/EBPα, CCAAT-enhancer binding protein alpha; canFam2, Canis lupus familiaris (domesticated dog) full genome as provided by UCSC (Jan 2008); CE, Composite regulatory element; ChIP-seq, Chromatin Immunoprecipitation Sequencing; CI, confidence interval; CS, composite score; CTPP, Confronting two-pair primer; DR1, down-regulator of transcription 1; ECR, evolutionary conserved region; ENCODE, encyclopedia of DNA elements; FPROM, human promoter prediction; GATA-1, globin transcription factor 1; GR, glucocorticoid receptor; *H. pylori, Helicobacter pylori*; hg18, Homo sapiens (Human) full genome as provided by UCSC (Jan 2008); HNF4, hepatocyte nuclear factor 4; HSF, heat-shock factor; HSTF, heat shock transcription factor; HWE, Hardy-Weinberg equilibrium; *IL-1, interleukin 1*; *IL1A, interleukin 1-alpha*; *IL1B, interleukin 1-beta*; *IL-1RN, interleukin 1 receptor antagonist*; IRF, interferon-regulatory factor; LD, linkage disequilibrium; LDF, linear discriminant factor; LPS, Lipopolysaccharide; MALT, mucosa-associated lymphoid tissue lymphoma; MAZ, Myc-associated zinc finger protein; mm9, Mus musculus (mouse) full genome as provided by UCSC (Jan 2008); MSA, multiple sequence alignment; Mulan, Multiple-sequence local alignment and visualization; NCBI, National Center for Biotechnology Information; NF-AT, nuclear factor of activated T cells; NF-Y, nuclear transcription factor Y; NF-κB, nuclear factor Kappa β, NK, natural killer cell; NNPP, Neural Network Promoter Prediction; NSAIDS, Non-steroidal anti-inflammatory drugs; Oct-1, Octamer transcription factor; OR, odds ratio; pan-Tro2, Pan troglodytes (chimpanzee) full genome as provided by UCSC (Jan 2008); PBS, phosphate buffer saline ; PCR, polymerase chain reaction; PPAR, peroxisome proliferator-activated receptor; PWM, position weight matrices; rheMac2, Macaca mulatta (rhesus monkey) full genome as provided by UCSC (Jan 2008); rn4, Rattus norvegicus (rat) full genome as provided by UCSC (Jan 2008); SNP, single nucleotide

polymorphism; Spi-1, SV40 promoter-1; SRF, serum response factor; SSRP, Structure specific recognition protein 1; STAT1, Signal transducer and activator of transcription 1; Std, standard; T cell, T lymphocyte cell; T, thymidine; TBP, TATA- binding protein; TF, transcription factor; TFBS, transcription factor binding site; TSS, transcription start site; TSSG, recognition of human PolII promoter region and start of transcription; TSSW, recognition of human PolII promoter region and start of transcription; UTR, unstranslated region; ZTA, the Epstein-Barr virus bZIP transcription factor;

Declarations

Ethics approval and consent to participate The study was approved by the Khartoum ministry of health research department, University of Khartoum, faculty of Medical Laboratory Sciences review board, and Research Ethics Committees of hospitals. Written informed consent was taken from participants before they enrolled. Consent for publication Not applicable Competing interests The authors declare that there are no conflicts of interest. Availability of data and materials The data regarding IL1B-31C>T genotypes and alleles distributions among participants; and the in silico results of the software that used to support the findings of this study are available from the corresponding author on reasonable request. Funding The authors received no specific funding for this work. Authors' contributions MAH and EMI supervised the methodology. ABI, AEA, AEAM and OBA collected the samples. ABI and AEA extracted the DNA. ABI amplified the 16S rRNA and IL1B genes. ABI and EBI analyzed the data. ABI performed the bioinformatics and statistical analysis for the data. ABI wrote the manuscript. MAH edited and revised the manuscript. All authors have read and approved the manuscript. Acknowledgements We gratefully acknowledge the participants and the staff of the gastroscopic unit in Ibin Sina specialized hospital, Soba teaching hospital, Modern Medical Centre, Al-Shorta hospital, and Al Faisal Specialized Hospital. We would also like to show our gratitude to the Department of Medical Microbiology, Faculty of Medical Laboratory Sciences, the University of Khartoum for their cooperation.

References

1. Moodley Y, Linz B, Bond RP, Nieuwoudt M, Soodyall H, Schlebusch CM, et al. Age of the association between Helicobacter pylori and man. PLoS pathogens. 2012;8(5):e1002693-e.
2. Suerbaum S, Michetti P. Helicobacter pylori infection. New England J Med. 2002;347:1175–86.
3. Sjomina O, Pavlova J, Niv Y, Leja M. Epidemiology of Helicobacter pylori infection. Helicobacter. 2018;23(S1):e12514.
4. Zamani M, Ebrahimtabar F, Zamani V, Miller WH, Alizadeh-Navaei R, Shokri-Shirvani J, et al. Systematic review with meta-analysis: the worldwide prevalence of Helicobacter pylori infection. Alimentary Pharmacology & Therapeutics. 2018;47(7):868-76.
5. Goh K-L, Chan W-K, Shiota S, Yamaoka Y. Epidemiology of Helicobacter pylori infection and public health implications. Helicobacter. 2011;16 Suppl 1(0 1):1-9.
6. Fock KM, Ang TL. Epidemiology of Helicobacter pylori infection and gastric cancer in Asia. J Gastroenterol Hepatol. 2010;25(3):479-86.
7. Chmiela M, Karwowska Z, Gonciarz W, Allushi B, Staczek P. Host pathogen interactions in Helicobacter pylori related gastric cancer. World journal of gastroenterology. 2017;23(9):1521-40.
8. Kodaman N, Pazos A, Schneider BG, Piazuelo MB, Mera R, Sobota RS, et al. Human and Helicobacter pylori coevolution shapes the risk of gastric disease. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(4):1455-60.
9. Testerman TL, Morris J. Beyond the stomach: An updated view of Helicobacter pylori pathogenesis, diagnosis, and treatment. World journal of gastroenterology. 2014;20(36):12781-808.
10. Cadamuro ACT, Rossi AFT, Maniezzo NM, Silva AE. Helicobacter pylori infection: Host immune response, implications on gene expression and microRNAs. World Journal of Gastroenterology : WJG. 2014;20(6):1424-37.
11. Forman D, Burley VJ. Gastric cancer: global pattern of the disease and an overview of environmental risk factors. Best practice & research Clinical gastroenterology. 2006;20(4):633-49.
12. de Sablet T, Piazuelo MB, Shaffer CL, Schneider BG, Asim M, Chaturvedi R, et al. Phylogeographic origin of Helicobacter pylori is a determinant of gastric cancer risk. Gut. 2011;60(9):1189-95.
13. Chang W-L, Yeh Y-C, Sheu B-S. The impacts of H. pylori virulence factors on the development of gastroduodenal diseases. J Biomed Sci. 2018;25(1):68-.
14. Nicklin MJH, Weith A, Duff GW. A Physical Map of the Region Encompassing the Human Interleukin-1α, Interleukin-1β, and Interleukin-1 Receptor Antagonist Genes. Genomics. 1994;19(2):382-4.
15. El-Omar EM. The importance of interleukin 1beta in Helicobacter pylori associated disease. Gut. 2001;48(6):743-7.
16. Dinarello CA. Interleukin-1 and interleukin-1 antagonism. Blood. 1991;77(8):1627-52.
17. Ren K, Torres R. Role of interleukin-1beta during pain and inflammation. Brain Res Rev. 2009;60(1):57-64.
18. Rad R, Dossumbekova A, Neu B, Lang R, Bauer S, Saur D, et al. Cytokine gene polymorphisms influence mucosal cytokine expression, gastric inflammation, and host specific colonisation during Helicobacter pylori infection. Gut. 2004;53(8):1082-9.
19. Sun X, Cai H, Li Z, Li S, Yin W, Dong G, et al. Association between IL-1b polymorphisms and gastritis risk. Medicines. 2017;96:5.
20. Motamedi Rad N, Rezaeishahmirzadi M, Shakeri S, Abbaszadegan MR, Shekari M. Association of IL-1B+3954 and IL-1RN Polymorphisms in Chronic Gastritis and Peptic Ulcer. Iranian journal of public health. 2018;47(9):1364-70.
21. Robert A, Olafsson AS, Lancaster C, Zhang WR. Interleukin-1 is cytoprotective, antisecretory, stimulates PGE2 synthesis by the stomach, and retards gastric emptying. Life sciences. 1991;48(2):123-34.
22. Pulugulla SH, Packard TA, Galloway NLK, Grimmett ZW, Doitsh G, Adamik J, et al. Distinct mechanisms regulate IL1B gene transcription in lymphoid CD4 T cells and monocytes. Cytokine. 2018;111:373-81.

23. Adamik J, Wang KZQ, Unlu S, Su A-JA, Tannahill GM, Galson DL, et al. Distinct mechanisms for induction and tolerance regulate the immediate early genes encoding interleukin 1β and tumor necrosis factor α. PloS one. 2013;8(8):e70622-e.
24. Marecki S, Riendeau CJ, Liang MD, Fenton MJ. PU.1 and Multiple IFN Regulatory Factor Proteins Synergize to Mediate Transcriptional Activation of the Human IL-1β Gene. The Journal of Immunology. 2001;166(11):6829-38.
25. Dinarello C. Biologic basis for interleukin-1 in disease. Blood. 1996;87(6):2095-147.
26. Kominato Y, Galson D, Waterman WR, Webb AC, Auron PE. Monocyte expression of the human prointerleukin 1 beta gene (IL1B) is dependent on promoter sequences which bind the hematopoietic transcription factor Spi-1/PU.1. Molecular and Cellular Biology. 1995;15(1):59-68.
27. Auron PE, Webb AC. Interleukin-1: a gene expression system regulated at multiple levels. Eur Cytokine Netw. 1994;5(6):573-92.
28. Listman JA, Race JE, Walker-Kopp N, Unlu S, Auron PE. Inhibition of IL-1beta transcription by peptides derived from the hCMV IE2 transactivator. Mol Immunol. 2008;45(9):2667-77.
29. Zhang Y, Saccani S, Shin H, Nikolajczyk BS. Dynamic protein associations define two phases of IL-1beta transcriptional activation. Journal of immunology (Baltimore, Md : 1950). 2008;181(1):503-12.
30. Shirakawa F, Saito K, Bonagura CA, Galson DL, Fenton MJ, Webb AC, et al. The human prointerleukin 1 beta gene requires DNA sequences both proximal and distal to the transcription start site for tissue-specific induction. Molecular and cellular biology. 1993;13(3):1332-44.
31. Ahmed AB, Zidi S, Sghaier I, Ghazouani E, Mezlini A, Almawi W, et al. Common variants in IL-1RN, IL-1beta and TNF-alpha and the risk of ovarian cancer: a case control study. Central-European journal of immunology. 2017;42(2):150-5.
32. Figueiredo CA, Marques CR, Costa Rdos S, da Silva HB, Alcantara-Neves NM. Cytokines, cytokine gene polymorphisms and Helicobacter pylori infection: friend or foe? World journal of gastroenterology. 2014;20(18):5235-43.
33. Lu W, Pan K, Zhang L, Lin D, Miao X, You W. Genetic polymorphisms of interleukin (IL)-1B, IL-1RN, IL-8, IL-10 and tumor necrosis factor {alpha} and risk of gastric cancer in a Chinese population. Carcinogenesis. 2005;26(3):631-6.
34. Singh H, Samani D, Nema V, Ghate MV, Gangakhedkar RR. IL-1RN and IL-1beta Polymorphism and ARV-Associated Hepatotoxicity. Mediators of inflammation. 2018;2018:4398150.
35. Shehjar F, Afroze D, Misgar RA, Malik SA, Laway BA. Association of polymorphic variants of IL-1beta and IL-1RN genes in the development of Graves' disease in Kashmiri population (North India). Human immunology. 2018;79(4):228-32.
36. Ma J, Wu D, Hu X, Li J, Cao M, Dong W. Associations between cytokine gene polymorphisms and susceptibility to Helicobacter pylori infection and Helicobacter pylori related gastric cancer, peptic ulcer disease: A meta-analysis. PloS one. 2017;12(4):e0176463.
37. Figueiredo CA, Marques CR, Costa RdS, da Silva HBF, Alcantara-Neves NM. Cytokines, cytokine gene polymorphisms and Helicobacter pylori infection: friend or foe? World journal of gastroenterology. 2014;20(18):5235-43.
38. Idris A, Ataelmanan A, Eltaher S, Idris E, Arbab B, Idris A, et al. Independently Carriage of IL-1RN*2 Allele Associated with Increased Risk of Gastric Cancer in The Sudanese Population. 2019.
39. Carbone M, Maugeri TL, Gugliandolo C, La Camera E, Biondo C, Fera MT. Occurrence of Helicobacter pylori DNA in the coastal environment of southern Italy (Straits of Messina). J Appl Microbiol. 2005;98(3):768-74.
40. Idris AB, Ataelmanan AE, Eltaher SM, Idris EB, Osman Arbab BM, Ibn Idris A, et al. Independently Carriage of IL-1RN*2 Allele Associated with Increased Risk of Gastric Cancer in The Sudanese Population. medRxiv. 2019:19013573.
41. Furuta T, El-Omar EM, Xiao F, Shirai N, Takashima M, Sugimura H. Interleukin 1beta polymorphisms increase risk of hypochlorhydria and atrophic gastritis and reduce risk of duodenal ulcer recurrence in Japan. Gastroenterology. 2002;123(1):92-105.
42. FinchTV. 1.4.0 ed. USA: Geospiza, Inc.; Seattle, WA,; 2012.
43. Altschul SF, Madden TL, Schäffer AA, J Zhang ZZ, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST. A new generation of protein database search programmes. Nucleic Acids Res. 1997;25(17):3389–402.
44. Hall TA. BioEdit: a user-friendly biologicalsequence alignment editor and analysis programfor Windows 95/98/NT. Nucl Acids Symp Ser 1999;41:95-8.
45. Xin S, Wang X, Dai G, Zhang J, An T, Zou W, et al. Bioinformatics Analysis of SNPs in IL-6 Gene Promoter of Jinghai Yellow Chickens. Genes 2018;9:446.
46. Reese MG. Application of a time-delay neural network to promoter annotation in the Drosophila melanogaster genome. Computers & chemistry. 2001;26(1):51-6.
47. Knudsen S. Promoter2.0: for the recognition of PolII promoter sequences. Bioinformatics (Oxford, England). 1999;15(5):356-61.
48. Solovyev VV, Shahmuradov IA, Salamov AA. Identification of promoter regions and regulatory sites. Methods in molecular biology (Clifton, NJ). 2010;674:57-83.
49. A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS biology. 2011;9(4):e1001046.
50. Sloan CA, Chan ET, Davidson JM, Malladi VS, Strattan JS, Hitz BC, et al. ENCODE data at the ENCODE portal. Nucleic Acids Res. 2016;44(D1):D726-D32.
51. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.
52. Li L-C, Dahiya R. MethPrimer: designing primers for methylation PCRs. Bioinformatics (Oxford, England). 2002;18(11):1427-31.
53. Farré D, Roset R, Huerta M, Adsuara JE, Roselló L, Albà MM, et al. Identification of patterns in biological sequences at the ALGGEN server: PROMO and MALGEN. Nucleic Acids Res. 2003;31(13):3651-3.
54. Messeguer X, Escudero R, Farre D, Nunez O, Martinez J, Alba MM. PROMO: detection of known transcription regulatory elements using species-tailored searches. Bioinformatics (Oxford, England). 2002;18(2):333-4.
55. Grabe N. AliBaba2: context specific identification of transcription factor binding sites. In Silico Biol. 2002;2(1):S1-15.

56. Lee T-Y, Chang W-C, Hsu J, Chang T-H, Shien D-M. GPMiner: An integrated system for mining combinatorial cis-regulatory elements in mammalian gene group. BMC genomics. 2012;13 Suppl 1:S3.
57. Tsunoda T, Takagi T. Estimating transcription factor bindability on DNA. Bioinformatics (Oxford, England). 1999;15(7-8):622-30.
58. Ghosh D. Object-oriented transcription factors database (ooTFD). Nucleic Acids Res. 2000;28(1):308-10.
59. Kel-Margoulis OV, Kel AE, Reuter I, Deineko IV, Wingender E. TRANSCompel®: a database on composite regulatory elements in eukaryotic genes. Nucleic Acids Res. 2002;30(2):332–4.
60. Deyneko IV, Kel AE, Kel-Margoulis OV, Deineko EV, Wingender E, Weiss S. MatrixCatch - a novel tool for the recognition of composite regulatory elements in promoters. BMC Bioinformatics. 2013;14(1):241.
61. Ovcharenko I, Nobrega MA, Loots GG, Stubbs L. ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. Nucleic Acids Res. 2004;32(Web Server issue):W280-W6.
62. Ovcharenko I, Loots GG, Giardine BM, Hou M, Ma J, Hardison RC, et al. Mulan: multiple-sequence local alignment and visualization for studying function and evolution. Genome Res. 2005;15(1):184-94.
63. Hamajima N, Saito T, Matsuo K, Kozaki K, Takahashi T, Tajima K. Polymerase chain reaction with confronting two-pair primers for polymorphism genotyping. Japanese journal of cancer research : Gann. 2000;91(9):865-8.
64. Wang J, Hannenhalli S. A mammalian promoter model links cis elements to genetic networks. Biochemical and biophysical research communications. 2006;347(1):166-77.
65. Zhang G, Zhou B, Li S, Yue J, Yang H, Wen Y, et al. Allele-Specific Induction of IL-1β Expression by C/EBPβ and PU.1 Contributes to Increased Tuberculosis Susceptibility. PLOS Pathogens. 2014;10(10):e1004426.
66. Lind H, Haugen A, Zienolddiny S. Differential binding of proteins to the IL1B -31 T/C polymorphism in lung epithelial cells. Cytokine. 2007;38(1):43-8.
67. El-Omar EM, Carrington M, Chow W-H, McColl KEL, Bream JH, Young HA, et al. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. Nature. 2000;404(6776):398-402.
68. Al-Moundhri MS, Al-Nabhani M, Al-Bahrani B, Burney IA, Al-Madhani A, Ganguly SS, et al. Interleukin-1beta gene (IL-1B) and interleukin 1 receptor antagonist gene (IL-1RN) polymorphisms and gastric cancer risk in an Omani Arab population. Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association. 2006;9(4):284-90.
69. Hartland S, Newton JL, Griffin SM, Donaldson PT. A functional polymorphism in the interleukin-1 receptor-1 gene is associated with increased risk of Helicobacter pylori infection but not with gastric cancer. Digestive diseases and sciences. 2004;49(9):1545-50.
70. Zhao Y, Wang JW, Tanaka T, Hosono A, Ando R, Tokudome S, et al. Association between TNF-alpha and IL-1beta genotypes vs Helicobacter pylori infection in Indonesia. World journal of gastroenterology. 2013;19(46):8758-63.
71. Drici Ael M, Moulessehoul S, Tifrit A, Diaf M, Turki DK, Bachir M, et al. Effect of IL-1beta and IL-1RN polymorphisms in carcinogenesis of the gastric mucosa in patients infected with Helicobacter pylori in Algeria. The Libyan journal of medicine. 2016;11:31576.
72. Kumar S, Kumar A, Dixit VK. Evidences showing association of interleukin-1B polymorphisms with increased risk of gastric cancer in an Indian population. Biochemical and biophysical research communications. 2009;387(3):456-60.
73. Ruzzo A, Graziano F, Pizzagalli F, Santini D, Battistelli V, Panunzi S, et al. Interleukin 1B gene (IL-1B) and interleukin 1 receptor antagonist gene (IL-1RN) polymorphisms in Helicobacter pylori-negative gastric cancer of intestinal and diffuse histotype. Annals of oncology : official journal of the European Society for Medical Oncology. 2005;16(6):887-92.
74. Xue H, Lin B, Ni P, Xu H, Huang G. Interleukin-1B and interleukin-1 RN polymorphisms and gastric carcinoma risk: a meta-analysis. Journal of gastroenterology and hepatology. 2010;25(10):1604-17.
75. Kimura R, Nishioka T, Soemantri A, Ishida T. Cis-acting effect of the IL1B C−31T polymorphism on IL-1β mRNA expression. Genes & Immunity. 2004;5(7):572-5.
76. Santtila S, Savinainen K, Hurme M. Presence of the IL-1RA allele 2 (IL1RN*2) is associated with enhanced IL-1beta production in vitro. Scand J Immunol. 1998;47(3):195-8.
77. Hwang I-R, Kodama T, Kikuchi S, Sakai K, Peterson LE, Graham DY, et al. Effect of interleukin 1 polymorphisms on gastric mucosal interleukin 1β production in Helicobacter pylori infection. Gastroenterology. 2002;123(6):1793-803.
78. Kaijzel EL, Bayley JP, van Krugten MV, Smith L, van de Linde P, Bakker AM, et al. Allele-specific quantification of tumor necrosis factor α (TNF) transcription and the role of promoter polymorphisms in rheumatoid arthritis patients and healthy individuals. Genes & Immunity. 2001;2(3):135-44.

Figures

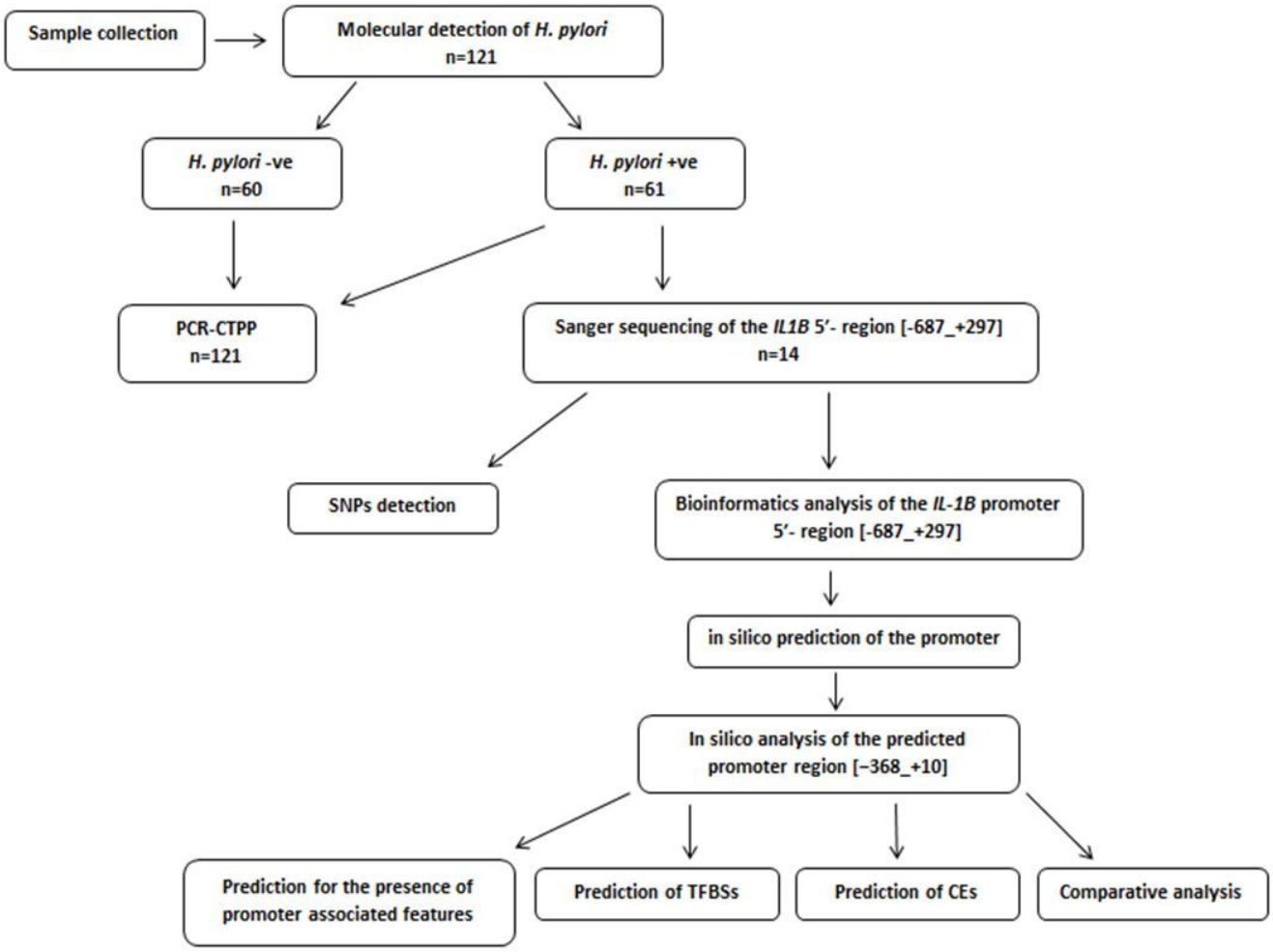

Figure 1

Schematic representation of the methodology.

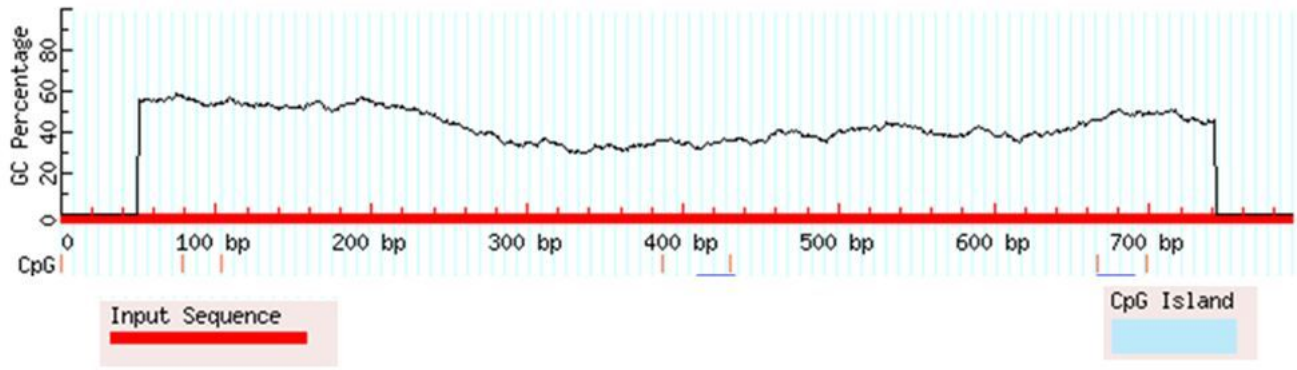

Figure 2

MethPrimer software prediction of no CpG islands in the predicted promoter regions which are located at -328 bp, -124 bp and +1 bp.

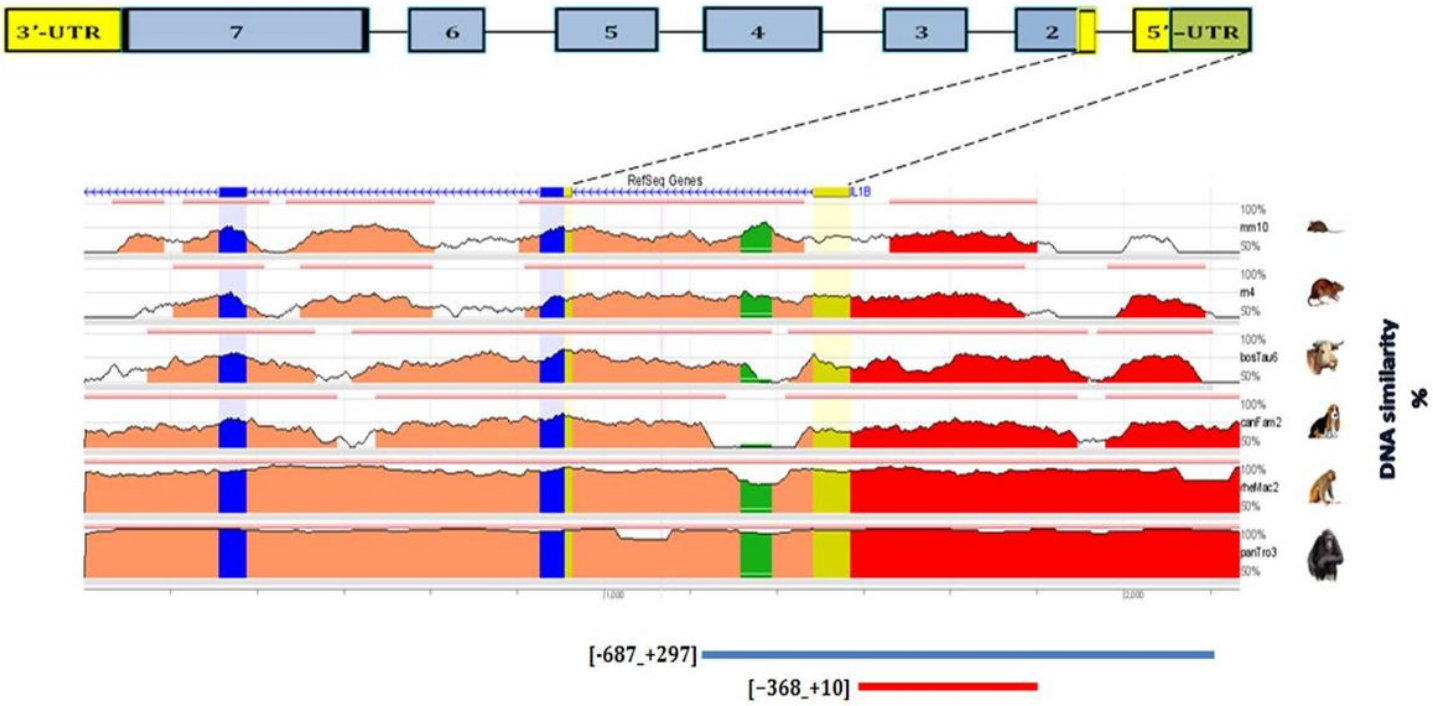

Figure 3

Shows conservation of the IL1B 5'-upstream region compared to the human sequence in region [−687_+297] (hg19 chr2:113594193-113594801). The height of the conservation plot at each position represents the number of nucleotides conserved in a windows of 100 nucleotides centered on that position. The pink rectangles at the top of the plot represent the evolutionary conserved regions, which defined as regions of 100 nucleotides with at least 70% identity. Blue boxes represent IL1B exon , while yellow indicates the IL1B 5'-UTR. Intragenic positions are highlighted in red, or in green when corresponding to transposable elements and simple repeats. Overview from the ECR Browser.(61)

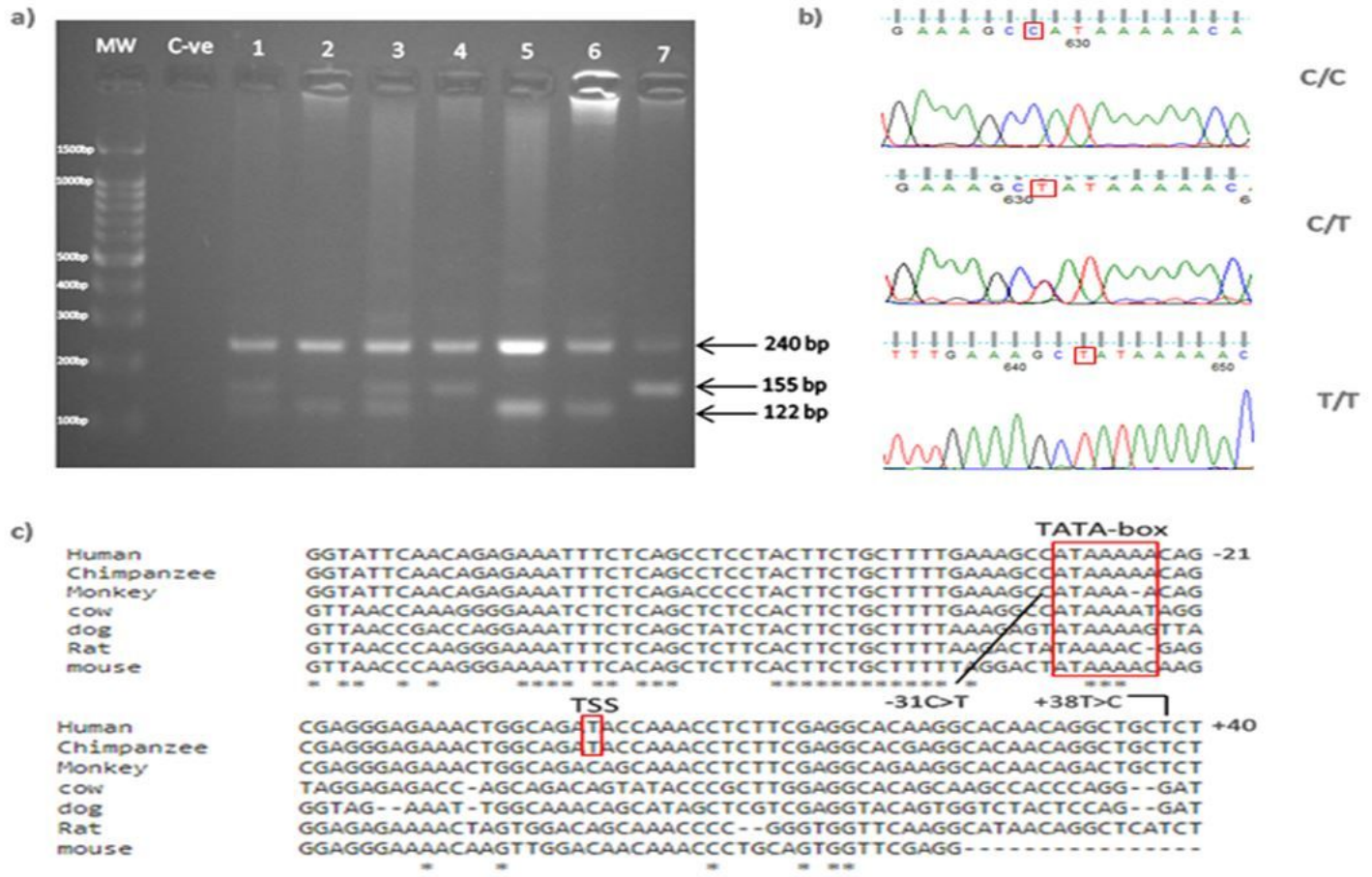

Figure 4

4a) Illustrates PCR-CTPP products analyzed on a 2% agarose gel stained with ethidium bromide. Three genotypes can be seen. 1and 3 show 240bp, 155bp and 122bp which indicate a heterozygous genotype. 2, 5 and 6 show 240bp and 122bp which indicate a homozygous T genotype. While 4 and 7 show 240bp and 155bp which indicate a homozygous C genotype. 4b) shows the sequencing result of the polymorphism by using Finch TV software. 4c) Illustrates the conservation of (TC; dbSNP: rs1143627) at position -31 among different species.

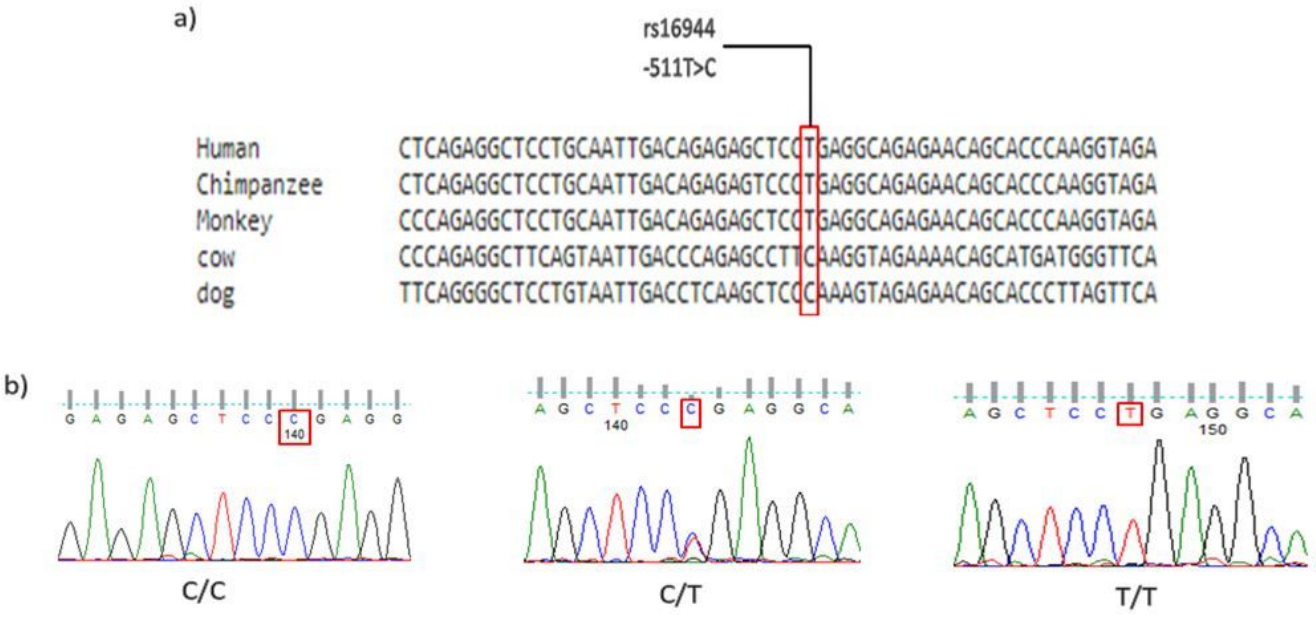

Figure 5

5a) Conservation of (CT; dbSNP: rs16944) at position -511 among various species. 5b) shows the chromatogram results of the polymorphism by using Finch TV software.

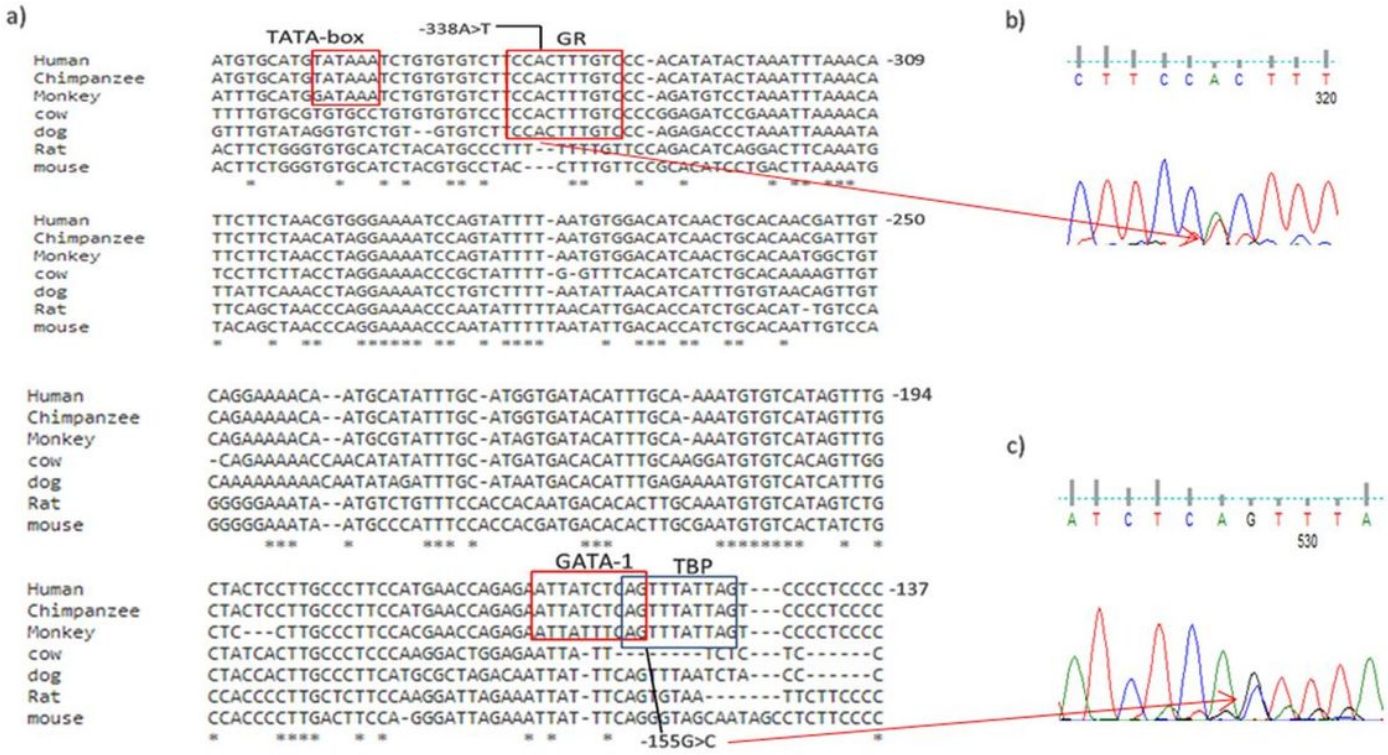

Figure 6

6a) Overview on the mammalian conservation of -338A>T and -155G>C SNPs . The nucleotides are enumerated at each line on the right side. The in silico predicted TATA- and TFBSs are marked in boxes. 6b) and 6c) Shows the chromatogram results of the polymorphisms by using Finch TV software.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- originalfiguresofIL1Bmanuscript.docx
- Questionair.docx