

# Beyond SNPs and CNV: Pharmacogenomics of Polymorphic Tandem Repeats

Evgeny Krynetskiy

School of Pharmacy, Temple University, PA, USA

## Abstract

Polymorphic Short Tandem Repeats (STR) emerged as a separate class of genetic mutation, which together with Single Nucleotide Polymorphisms (SNPs) and Copy Number Variations (CNVs) can explain variability in response to pharmacotherapy. STR draws interest in pharmacogenomics research because of their prevalence in the human genome, and their putative functional role as regulators of gene expression. Depending on the search algorithm, there are approximately 700,000–1,000,000 STR loci with 2-6 bp long motifs in the human reference genome. STR is non-randomly distributed across Untranslated Regions (UTRs), protein-coding sequences, and introns, and is overrepresented in the promoter regions of the human genes. The functional role of STR has been demonstrated by effects on gene expression, splicing, protein sequence, and association with pathogenic effects. An intrinsic property of STR is the high rate of mutation by expansion or contraction in the number of repeat units. Variation in the length of STR plays an important role in modulating gene expression, and STR is likely to be general regulatory elements which attenuate expression of multiple genes. Elucidating the effects of STR on gene expression may in part explain variability in drug response, something that cannot be achieved by focusing analysis exclusively on SNPs or CNV. This review summarizes the role of polymorphic STR in clinical manifestations including response to pharmacotherapy.

**Keywords:** Short Tandem Repeats (STR); Pharmacogenomics; Pharmacogenetics; Gene expression; Microsatellite; Human genome

## Introduction

Polymorphic tandem repeats draw interest in pharmacogenomics research because of their prevalence in the human genome, and their putative functional role. Polymorphic Short Tandem Repeats (STR) emerged as a separate class of genetic mutations, which together with Single Nucleotide Polymorphisms (SNPs), Copy Number Variations (CNVs), and biallelic indels can explain variability in response to pharmacotherapy.

Repetitive DNA sequences may comprise over two-thirds of the human genome [1]. Depending on the length of the repeated motif, tandem repeats are categorized as microsatellites (short tandem repeats of DNA motifs 1 to 6 bp long, or STR), minisatellites (tandem repeats of moderate motifs 10-100 bp long), and macrosatellites with motifs longer than 100 bp. In humans, STR makes up to 3% of the total genomic DNA which exceeds the protein coding part of the human genome [2].

Depending on the search algorithm, there are approximately 700,000–1,000,000 STR loci with 2-6 bp long motifs in the human reference genome [3,4]. Di- and tetra-nucleotide STR constitute about 75% of STR, with the remaining loci containing tri-, penta-, and hexa-nucleotide repeats. The overall STR density in the human genome is comparable across chromosomes (mean  $\pm$  SD=13,613  $\pm$  1,887 bp/Mb), with chromosome 19 showing the highest STR density (20,351 bp/Mb) [5]. Within genes, microsatellite repeats are non-randomly distributed across protein-coding sequences, untranslated regions (UTRs), and introns. In the coding regions of the genes, repeats predominantly have either trimeric or hexameric repeat unit, likely as a result of selection against frameshift mutations [4,6]. STR containing dinucleotide repeat units are much more abundant in the regulatory or UTR regions than in other genomic regions [2].

Initially labeled as nonfunctional (junk) DNA, STR is now considered to have biological functions. Microsatellite repeats are concentrated at the start of human genes, where they are highly conserved near transcription start sites [6,7]. About 19% of human

genes contain at least one STR in their upstream regulatory region [6,8]. Recent studies demonstrated that STR in the human genome contributes to variation in gene expression [9]. Statistical analysis of tandem repeat distribution showed a sharp increase around Transcription Start Site (TSS), spanning several kilobases up- and downstream from the TSS [7].

The exact mechanism of expression modulation by STR remains a matter of discussion, and may vary for different STR motifs. For example, poly-A STR (microsatellites with A/T motif), the most frequent microsatellites in the human genome, are common elements of the promoters in the human genome [7]. Poly-A STR are hypothesized to participate in the regulation of gene activity because they disrupt nucleosome binding which could be a molecular mechanism for modulating gene expression [10]. Another example of STR overrepresented in the promoter regions of multiple genes is AC/GT dinucleotide tandem repeat [7]. The AC/GT tandem repeat sequence composed of alternating purine-pyrimidine bases facilitates formation of Z-DNA, and could prevent nucleosome binding leading to chromatin opening. Formation of H-DNA triplex structure composed of CT/AG poly-purine/poly-pyrimidine mirror repeats potentially leads to modified chromatin structure and transcriptional activation.

An intrinsic property of STR is the high rate of mutation. STR mutates by expansion or contraction in the number of repeat units, and is often described as a Variable Number of Tandem Repeats (VNTR). The frequency of STR mutations depends on the length of the repeat unit, the number of the repeats, and the match to the

**\*Corresponding author:** Evgeny Krynetskiy, School of Pharmacy, Temple University, PA 19140, USA, Tel: 215-707-4257; Fax: 215-707-5620; E-mail: [ekrynets@temple.edu](mailto:ekrynets@temple.edu)

Received May 18, 2017; Accepted June 20, 2017; Published June 27, 2017

**Citation:** Krynetskiy E (2017) Beyond SNPs and CNV: Pharmacogenomics of Polymorphic Tandem Repeats. J Pharmacogenomics Pharmacoproteomics 8: 170. doi: 10.4172/2153-0645.1000170

**Copyright:** © 2017 Krynetskiy E. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

consensus sequence (purity) of the repeat tract [11,12]. The mutation rates of STRs often lie between  $10^{-3}$ - $10^{-6}$  per cell generation) which is 10 to  $10^5$ -fold higher than the average mutation rates observed in non-repeated regions of the genome [8,12,13]. Apparently, short repeats consisting of two or three repeat units are the starting point for the microsatellite expansion, either through base substitution, or duplication of an adjacent sequence. Once the critical number of repeat units is formed (10 repeat units for A/T and 5-6 repeats for AC/GT), the locus becomes hyper-variable [14,15]. Based on this mechanism, a true microsatellite can be defined as a repeat containing a minimal number of units required for the production of indel mutations at a frequency greater than the average frequency of indel mutations within the genome. Importantly, repeats in the coding regions are significantly less variable than repeats in introns, intergenic regions, 3'UTR, and non-coding RNA [16].

There are currently two major models that describe the mechanisms by which STR expand or contract: strand-slippage replication and recombination. Strand-slippage replication, also known as slipped-strand mispairing, or DNA slippage, occurs during replication of the tandem repeats. After the newly synthesized DNA strand denatures from the template strand during the synthesis of a tandem repeat region, it may shift its position along the tandem repeat in the process of renaturation ([8], and references therein). Recombination events, including unequal crossing over and gene conversion, can also lead to contraction and expansions of STR sequences. Recombination mechanism is hypothesized to activate in response to double-strand breaks formed during DNA replication. The repair of double-strand breaks by recombination will result either in addition or loss of repeat units. The importance of DNA strand breakage in the mutation of tandem repeat tracts is supported by studies demonstrating involvement of the double-strand break repair pathway in tandem repeat expansion and contraction [17].

While the causal relation with the variable numbers of repeats in microsatellites was convincingly documented for many diseases, importance of STR for drug response remains largely unexplored. Elucidating the effects of STR on expression of the genes involved in drug metabolism, drug transport, and drug-target interaction may help explain variability in efficacy and adverse reactions to pharmacotherapy, thus complementing analysis of SNPs, CNV, and biallelic indels in individual genomes. This review summarizes the role polymorphic STR play in clinical manifestations including their potential importance for pharmacogenetic analysis.

## Polymorphic Tandem Repeats Modify Gene Expression

Several pieces of evidence confirm the role of STR in regulation of gene expression. First, microsatellite tandem repeats are overrepresented in the vicinity of transcription start points within the promoter regions of many genes in the human genome [6,7]. Next, polymorphic promoters STR are associated with increased variance in local gene expression and DNA methylation, suggesting functional role for STR [18]. Finally, recent studies revealed contribution of STR alleles to gene expression levels and phenotypes [9]. The functional role of STR has been demonstrated by effects on gene expression, splicing, protein sequence, and association with pathogenic effects [18]. Inherently hypervariable microsatellites modulate gene expression through several possible mechanisms, e.g. direct addition of functional DNA motifs, modification of local DNA or RNA structure, epigenetic modification of the local region, altered spacing or orientation of regulatory molecules, and alteration of nucleosome positioning [9,18].

The mechanisms by which microsatellite repeats affect gene expression are likely to be region-specific rather than site-specific because within the gene, regulatory microsatellite sequences are found in proximal or distal promoter regions, 5'-UTR, and introns.

Polymorphic microsatellites with putative regulatory functions, the corresponding genes and the effect of microsatellite length variability on gene expression are exemplified by fifteen STR in Table 1 which effects on gene expression were comprehensively documented. The corresponding genes encode proteins with different functions including receptors (ESK1, IFNAR1, GRIN2A, TLR2), growth factor (IGF1), cytokine (TNF), enzymes (HMOX1, MMP9, UGT1A1, UGT1A8/9), and transcription factors (FOXA2, FOXP3, STAT6). To find correlation between transcription level and a particular STR allele, several *in vitro*, *in vivo*, and *ex vivo* experimental techniques were used. Using transient transfection of cultured cells, the effect of STR on promoter activity can be estimated using reporter constructs where a specific promoter region drives expression of a reporter gene. Despite the simple experimental design, this method suffers certain limitations, particularly because important regulatory elements might be missing in the interrogated promoter region. Analysis of mRNA level in tissues or Peripheral Blood Mononuclear Cells (PBMC) obtained from individuals with a certain genotype often provides more relevant data on gene regulation by STR allelic variants. The analysis of the gene product (protein level, or enzymatic activity), or clinical phenotype (disease susceptibility, or drug response) is also used to find a correlation with the specific STR allele. Finally, bioinformatics analysis of multiple genomes followed by an association study can provide data on correlation between STR length and expression rate [19]. This last approach may provide a wealth of information for correlative studies though methodical problems still remain to be resolved [9,18].

Heme oxygenase 1 encoded by the *HMOX1* gene is an important stress response, anti-inflammatory, and antioxidant protein inducible in response to several types of stress and drug therapy. The GT repeat in *HMOX1* gene is one of the most extensively characterized examples of regulatory polymorphic microsatellites, and detailed mechanistic study of *HMOX1* gene expression helped to elucidate the role of polymorphic STR in its promoter. *HMOX1* gene contains a microsatellite sequence in the promoter region with a variable number of 10-43 GT repeats [20]. The effect of polymorphic STR in the *HMOX1* promoter on the gene expression was extensively studied using reporter constructs, gene expression experiments in patients' samples, and enzymatic activity. Transfection of rat aortic smooth muscle cells with reporter constructs [21], *HMOX1* mRNA level quantification in three malignant melanoma cell lines [22], and analysis of HO-1 protein in six urothelial cancer cell lines indicated that shorter GT repeats ( $n < 25$ ) were associated with higher level of basal expression of HO-1, and with higher induction of this protein [23,24]. Inducible promoter activity of 5'-flanking regions in the *HMOX1* gene was estimated by transfection of A549 and Hep3B cancer cells containing different number of GT repeats, with reporter constructs [23]. Exposure to  $H_2O_2$  induced the transcription of the reporter constructs with short (GT16 and GT20) but not long (GT29 or GT38) microsatellites. Similarly, umbilical endothelial cells (HUVEC) with short GT repeats produced more HO-1 upon induction with  $H_2O_2$  [25]. In contrast to these findings, baseline levels of *HMOX1* mRNA were found lower (and protein level higher) in PBMC from healthy subjects [26], while both carriers and non-carriers of L allele ( $n > 32$ ) showed similar *HMOX1* mRNA expression. Importantly, upon induction with heat or hemin, non-L carriers manifested 1.9-fold higher increase in mRNA level in response to heat, and after hemin stimulation the median *HMOX1* mRNA in non-L carriers was 3.9

S. No.	Motif	Gene ID	Gene region	Position	Chr	Reporter construct	mRNA expression	Protein activity	Longer alleles associated with	Reference
1	(GT) <sub>n</sub> , n=12-19	<i>FOXP3</i>	Promoter	Exon 0 (-150 – -177)	X	Yes	Yes	n/a	Decreased expression in three cell lines (HeLa, COS-7, and Jurkat T)	[35,36,38]
2	(GT) <sub>n</sub> , n=12-42	<i>GRIN2A</i>	Promoter	Upstream of TSS (-721 – -679)	16	Yes	n/a	Yes	Decreased expression in cell culture	[33,34]
3	(GT) <sub>n</sub> , n=10-43	<i>HMXO1</i>	Promoter	Upstream of TSS (-172)	22	Yes	Yes	Yes	Decreased expression	[23,24,27-29,31]
4	(CA) <sub>n</sub> , n=13-26	<i>EGFR</i>	Intron 1	Upstream of enhancer 2 (+1,788 to +2,318)	7	Yes	Yes	Yes	Decreased expression	[40,42,43,77]
5	(TA) <sub>n</sub> , n=5-8	<i>UGT1A1</i>	Promoter TATA box	Upstream of TSS (-23-38)	2	yes	n/a	yes	Decreased expression	[87-89]
6	(GT) <sub>n</sub> , n=12-24	<i>STAT6</i>	5'-UTR	Downstream from TSS (+94) (rs71802646); three more STR upstream of TSS (-2514); (-801); (-734);	12	yes	n/a	n/a	Decreased expression in HMC-1, BEAS-2B, and Jurkat cells	[90]
7	(CGC) <sub>n</sub> , n=5 ≥ 200	<i>FMR1</i>	5'-UTR	Downstream from TSS (+85)	X	Yes	Yes	Yes	Decreased expression	[69-72]
8	(AC) <sub>n</sub> , n=14-27; (GT) <sub>n</sub> , n=12-19	<i>COL1A2</i>	Promoter Intron 1	AC repeat (-1457-1374) GT repeat (+1413 – +1480)	7	Yes	n/a	n/a	Expression non-linearly depends on both repeats	[39]
9	(CCT) <sub>n</sub> , n=13-19	<i>FOXA2</i>	Intron 1	(-415) upstream of TSS in alternative exon 2	20	Yes	Yes	n/a	Non-linear (allele with n=14 shows max expression)	[46,71]
10	(GT) <sub>n</sub> , n=15-27	<i>VWF</i>	Promoter	Upstream of TSS (-2144 -2105)	12	Yes	n/a	n/a	Increased expression	[67,91]
11	(CA) <sub>n</sub> , n=12-28	<i>MMP9</i>	Promoter	Upstream of TSS (-90)	20	Yes	n/a	n/a	Increased expression	[92-94]
12	(GT) <sub>n</sub> , n=10-17	<i>IFNG</i>	Intron 1	Downstream from TSS (+994 bp) (rs3138557)	12	Yes	n/a	Yes	Increased expression in cultured cells (HepG2, Jurkat); decreased expression <i>ex vivo</i>	[95-97]
13	(GT) <sub>n</sub> , n=12-28	<i>TLR2</i>	Intron 2	(-100 upstream of the start codon)	4	Yes	Yes	n/a	Increased expression in K562 cells; decreased expression <i>ex vivo</i>	[45,98]
14	(T) <sub>n</sub> , n=9-19	<i>UTG1A8</i> <i>UGT1A9</i>	Promoter	Upstream of TSS (-120)	2	Yes	n/a	n/a	Increased expression in Caco 2 cells	[99]
15	(GAG) <sub>n</sub> , n=4-10	<i>GCLC</i>	5' UTR	Upstream of the start codon (-10)	6	Yes	n/a	yes	Increased expression	[100,101]

**Table 1:** Polymorphic microsatellite sequences in the regulatory regions of the human genes, and their effect on gene expression.

fold higher than in L carriers ( $p=0.0028$ ) [27]. Studies of transcription in cell lines and blood cells collected from patients [27-29], protein synthesis and enzymatic activity [23,28,30], and clinical effects [27,30] demonstrated significant association with the number of GT-repeats in this STR. HO-1 mRNA expression and HO activity were significantly higher in lymphoblastoid cell lines with SS genotype compared with those with LL genotype [28]. Inducible regulation of *HMOX1* gene is under control of a complex regulatory mechanism [31], and the phenotypic manifestation of STR polymorphism may therefore be obscured [26]. The level of HO-1 protein determined by flow cytometry in PBMC from healthy individual carriers of S alleles (number of repeats  $\leq 25$ ) and L alleles (number of repeats  $> 25$ ) was significantly different: following LPS stimulation, monocytes from individuals with SS genotype showed a significantly higher HO-1 expression compared to LL homozygous individuals [30]. In a separate study, PBMC isolated from patients with suspected coronary atherosclerosis were treated with hemin. Assessment of *HMOX1* mRNA revealed significantly higher hemin-stimulated mRNA expression in SS genotypes compared to SL and LL genotypes (S allele, number of repeats  $< 26$ , and L allele, number of repeats  $> 26$ ) [29]. Analysis of the microsatellite effect is often

complicated by the presence of alternative TSS or splicing. A novel exon 1a was found in the *HMOX1* gene placing a (GT) microsatellite in intronic position within the 5'-untranslated region [31]. The quantitative outcome of alternative splicing within the 5'-untranslated region was affected by (GT)<sub>n</sub> microsatellite polymorphism.

The similar relation between the number of GT repeats in the promoter region and gene expression activity was detected in *GRIN2A* gene (Table 1) encoding the NR2A subunit of NMDA receptor expressed in neurons [32]. Reporter constructs with *GRIN2A* promoter coupled with *luc* gene demonstrated increased luc expression with shortening STR in *GRIN2A* promoter [33,34]. The promoter activity of the construct with 25-42 GT-repeats was 50-61% lower than that with no GT repeats. The receptor binding assay in postmortem brains indicated reduction of *GRIN2A* expression in the carriers of longer GT repeats [33].

The *FOXP3* gene is located on the X chromosome and mediates functional activity of T regulatory cells (Table 1). An association between *FOXP3* microsatellite polymorphism in a region with promoter/enhancer activity, and gene expression has been found in

S. No.	Gene	Protein function	Motif	Associated condition	Potential drug interactions	Ref
1	<i>ESR1</i>	Receptor	(GT) <sub>n</sub> , n=11-18 (TA) <sub>n</sub> , n=10-27	Breast cancer; lone atrial fibrillation; postpartum depression; harm avoidance score; osteoporosis	Anastrozole, atorvastatin, cisplatin, conjugated estrogens, dexamethasone, exemestane, glibenclamide, leflunomide, letrozole, medroxyprogesterone, methamphetamine, raloxifene, tamoxifen	[102-109]
2	<i>IFNAR1</i>	Receptor	(GT) <sub>n</sub> , n=5-17	Depressive symptoms during IFG-alpha therapy for hepatitis C; response to interferon therapy of hepatitis C	Glatiramer acetate	[107-109]
3	<i>IGF1</i>	Insulin-like growth factor	(CA) <sub>n</sub> , n=10-24	Endometrial cancer; age at natural menopause; disease onset in HNPCC; colorectal cancer	No known drug association	[110-114]
4	<i>TIGR/MYOC</i>	Cytoskeletal function	(GT) <sub>n</sub> , n=13-14 and n=15-16	Juvenile-onset primary open-angle glaucoma; glaucoma	No known drug association	[115,116]
5	<i>TNF</i>	Cytokine	(GT) <sub>n</sub> , n=7-18	Myocardial infarction; systemic lupus erythematosus; dengue; gastric and hepatocellular cancer	Adalimumab, atorvastatin, carbamazepine, clozapine, cyclosporine, etanercept, ethambutol, infliximab, isoniazid, lansoprazole, Mycophenolate mofetil, omeprazole, pyrazinamide, rabeprazole, rifampin, rituximab, sirolimus, sorafenib, stavudine, tumor necrosis factor alpha (TNF-alpha) inhibitors	[117-121]
6	<i>COL1A2</i>	The fibrillar collagen	(AC) <sub>n</sub> , n=14-27; (GT) <sub>n</sub> , n=12-19	Bone mineral density; systemic sclerosis	Daunorubicin, doxorubicin	[39,74,122-124]
7	<i>EGFR</i>	Target	(CA) <sub>n</sub> , n=14-26	Thymoma aggressiveness; response to TKI therapy in NSCLC patients; response to 5-FU therapy	Afatinib, alkylating agents, carboplatin, cetuximab, docetaxel, erlotinib, fluorouracil, gefitinib, geldanamycin, gemcitabine, irinotecan, leucovorin, paclitaxel, panitumumab, tegafur, topoisomerase I inhibitors	[40,42,78,125-128]
8	<i>FMR1</i>	Cognitive development	(CGC) <sub>n</sub> , n=5 ≥ 200	Premature ovarian failure; primary ovarian insufficiency and tremor-ataxia syndrome	No drug interaction	[68,71,72,129,130]
9	<i>FOXA2</i>	Forkhead box A2	(CCT) <sub>n</sub> , n=13-19	Type 2 diabetes; CYP3A4 expression	Drugs metabolized by CYP3A4	[46,131]
10	<i>FOXP3</i>	Transcription factor	(GT) <sub>n</sub> , N=12-19	Type 1 diabetes; survival of renal transplant patients; graft versus host disease	Tacrolimus	[35-38]
11	<i>GCLC</i>	Target	(GAG) <sub>n</sub> , n=4-10	Schizophrenia; Type 1 diabetes	Sulphamethoxazole	[100,101,132-134]
12	<i>GRIN2A</i>	NMDA receptor	(GT) <sub>n</sub> , n=12-42	Schizophrenia, alcoholism; hippocampal and amygdala volumes; concussion recovery	Methylphenidate	[33,34,60-64]
13	<i>HMXO1</i>	Stress response anti-inflammatory, anti-oxidant, anti-proliferative	(GT) <sub>n</sub> , n=10-43	Melanoma; inhibitory Ab to F8 in severe hemophilia A; Type 2 diabetes mellitus; coronary atherosclerosis; emphysema; severe malaria; pulmonary disease, cardiovascular disease, renal transplantation, obstetrics, neurological disease, hematological/serological disorders	Aspirin, statins, mimetic peptides, probucol, losartan, paclitaxel, rapamycin, cyclosporin, curcumin, resveratrol	[22-29,53,56,82,135]
14	<i>IFNG</i>	Interferon-gamma	(GT) <sub>n</sub> , n=10-17	Generalized vitiligo; malaria; response to immunosuppressive treatment; sporadic breast cancer	Infliximab, adalimumab, etanercept	[96,97,109,136]
15	<i>MMP9</i>	Matrix metalloproteinase	(CA) <sub>n</sub> , n=17-25	Diabetic end-stage renal disease; diabetic nephropathy; bladder cancer invasiveness; multiple sclerosis; age-related macular degeneration; carotid atherosclerosis	Hydralazine, nifedipine, methyl dopa	[92-94,137]



16	STAT6	Transcription factor	(GT)n, n=12-24	Bronchial asthma, atopic dermatitis, food-related anaphylaxis asthma; eosinophil cell count;	No known drug interactions	[90,138-140]
17	TLR2	Receptor	(GT)n, n=12-28	Colorectal cancer; leprosy; TB; acute pancreatitis; spontaneous bacterial peritonitis; rheumatoid arthritis	TNF alpha inhibitors	[45,141-146]
18	UGT1A1	DME	(TA)n, n=5-8	Gilbert's syndrome, Hyperbilirubinemia transient familial neonatal, Crigler-Najjar syndrome	Acetaminophen, antivirals for treatment of HIV infections, atazanavir, belinostat, bevacizumab, bilirubin, cisplatin, deferasirox, dolutegravir, fluorouracil, gepirone hydrochloride, irinotecan, leucovorin, nilotinib, olanzapine, oxaliplatin, pazopanib, peginterferon alfa-2b, raloxifene, raltegravir, ribavirin, ritonavir, SN-38, sorafenib	[87,147,148]
19	UTG1A8 UGT1A9	DME	(T)n, n=9-19	Pericholangitis	UGT1A8: ABT-751, allopurinol, anthracyclines and related substances, atazanavir, cyclosporine, febuxostat, irinotecan, lamivudine, mycophenolate mofetil, mycophenolic acid, sirolimus, tacrolimus, tipifarnib, valproic acid, zidovudine UGT1A9: acetaminophen, allopurinol, anthracyclines and related substances, aspirin, atazanavir, cisplatin, entacapone, febuxostat, irinotecan, labetalol, lamivudine, microsatellite, mycophenolate mofetil, mycophenolic acid, oxcarbazepine, propofol, raltegravir, ritonavir, simvastatin, SN-38, sorafenib, sulfapyrazone, tipifarnib, tolcapone, valproic acid, zidovudine	[99]
20	VWF	An antihemophilic factor carrier	(GT)n, N=15-24	Circulating level of VWF; Cortisol-dependent increase of VWF	No known drug interactions	[65,91]

**Table 2:** Association of polymorphic microsatellite sequences in the regulatory regions of the human genes with clinical phenotypes, and potential effect on drug response.

several cell lines (HeLa, COS-7, and Jurkat T) transfected with reporter constructs [35,36]. In addition, an association between the number of GT repeat in *FOXP3* gene and several conditions including type 1 diabetes, development of severe acute graft versus host disease in patients transplanted from donors harboring short alleles, and renal allograft survival was detected [35-37]. Type 1 diabetes-associated allele (GT)15 in *FOXP3* promoter demonstrated a higher activity in reporter gene experiments with three cell lines supporting a direct effect of STR on *FOXP3* expression [35,36]. On the other hand, there was no significant difference in *FOXP3* expression level between the groups of asthmatic patients stratified by (GT)n genotype [38]. This latter finding may indicate that this microsatellite polymorphism does not directly regulate the expression of the chromosomal gene.

Transcription driven by *COL1A2* promoter (Table 1) is enhanced by the presence of two dinucleotide repeats located in the 5'-flanking region of the gene, and the first intron of the gene, at the distance about 1.4 kB [39]. When tested in the reporter construct-transfected human skin fibroblasts, these repeats were found essential for the transcriptional stimulation of the *COL1A2* gene. Importantly, the stimulating effect was due to the presence of both repeats, rather than any one of them. Moreover, the transcriptional activity was modulated by combination of the alleles differed in the number of dinucleotide repeats.

In addition to microsatellites in the promoter region, AC/GT

repeats with recognized gene modulating activity were detected in intronic sequences (Table 1). An important example of a functional microsatellite residing outside the promoter was found in the gene coding for Epidermal Growth Factor Receptor (EGFR). The human *EGFR* gene is located on chromosome 7, and regulated by a single promoter and two enhancer regions. The length of a highly polymorphic CA/TG microsatellite in the intron 1 of *EGFR* correlates with expression of EGFR both *in vitro* and *in vivo* [40,41]. The longer CA21 allele exhibits 80% lower transcription than CA16 allele [42]. Cells with shorter CA repeats manifested higher transcription level of EGFR mRNA, higher level of EGFR protein, and were more sensitive to tyrosine kinase inhibitor erlotinib in 12 head and neck cancer cell lines [43]. Importantly, the length of CA repeats modulates *EGFR* transcription in NSCLC patients, and affects protein expression and response to therapy [43,44].

Finally, GT microsatellite repeat in the second intron of *TLR2* gene was demonstrated to affect promoter activity. Expression of TLR2 protein in PBMC cells in the carriers of short alleles (n≤16) was higher than in non-carriers of short alleles [45]. CCT trinucleotide repeat in the alternative promoter (212 bp upstream of TSS.1) modulates expression of FoxA2 transcription factor, with the highest expression from the allele containing 14 repeats [46]. Transfection experiments with polymorphic repeats in the reporter constructs confirmed the data on the transcriptional activity of this allele in HepG2 cells.

In many cases, the shorter GT repeats reveal increased promoter activity, as demonstrated by promoter assays. Table 1 summarizes examples of microsatellite polymorphisms in the regulatory regions of the human genes, and available data on promoter activity. For instance, reporter constructs with shorter AC/GT repeats demonstrated increased activity with *COL2A1*, *FOXP3*, *GRIN2A*, *HMOX1*, *MMP9*, *STAT6* promoters (Table 1). This correlation holds true also for GT repeats in intronic sequences of *EGFR*, *TLR2*, as well as trinucleotide repeats of *FOXA2* (intron 1), and *FMR1* (5'-UTR) genes. On the other hand, such correlation was not found for GT repeat-containing promoters in *VWF*, *IFNG* genes, mononucleotide (T)<sub>n</sub> STR in *UGT1A8/9*, and trinucleotide STR in the 5'-UTR of *GCLC* gene.

## STR and Clinical Phenotype

Because of high variability, microsatellite loci are often used in forensics, population genetics, and genetic genealogy [47]. Significant associations were demonstrated between microsatellite variants and genetic diseases including some neurological conditions such as Huntington's disease, Parkinson's disease, autism, amyotrophic lateral sclerosis, and certain types of ataxia [48,49]. Such relationship supports the premise about phenotypic manifestation of microsatellite polymorphisms. Multiple studies evidenced the clinical impact of GT<sub>n</sub> microsatellite polymorphism in pulmonary disease, cardiovascular disease, renal transplantation, obstetrics, neurological disease, and hematological/serological disorders [21,22,50-59]. Clinical manifestations of microsatellite polymorphisms are exemplified by twenty STR with clearly defined phenotypes shown in Table 2.

The functional role of GT repeat polymorphism in *HMOX1* promoter in human disease was overviewed by Exner et al. [50]. Daenen et al. performed a systematic review and a meta-analysis on the association of GT microsatellite polymorphism in the *HMOX1* promoter and cardiovascular disease with the cutoff of the short allele set to 25-27 repeats [59]. The results from 41 selected studies revealed that the proportion of the short SS genotype was lower in the Cardiovascular patient (CVD) group compared with non-CVD group (13.3% vs. 18.9%,  $P < 0.0001$ ). The odds ratio in LL vs SS genotype was 1.769 (95% Confidence Interval [CI], 1.594-1.963). Another systematic review and meta-analysis of 5 studies of the association between the microsatellite polymorphism in the *HMOX1* promoter and type 2 diabetes contained data on 1751 cases and 2902 controls. The odds ratio for type 2 diabetes in persons with LL genotype was significantly increased compared with the SS genotype (OR=1.25, 95% CI: 1.04, 1.50;  $P=0.02$ ). Statistical analysis showed that carriers of longer GT repeats ( $\geq 25$ -27 repeats) in the *HMOX1* promoter had higher risk of type 2 diabetes [53]. Analysis of GT repeat distribution in 942 children with sickle cell disease demonstrated that children with two short alleles ( $\leq 25$  repeats) had lower rate of hospitalization for acute chest syndrome (incidence rate ratio 0.28, 95% CI, 0.10-0.81) [54]. The GT repeat polymorphism in the *HMOX1* promoter was associated with severe disease and death in Gambian children with malaria [27].

Polymorphism of microsatellite sequences was also associated with psychiatric disorders, and concussion recovery rate. Clinical studies indicated that the longer alleles of *GRIN2A* were overrepresented in schizophrenics, and the score of symptom severity correlated with repeat length. This study was later expanded to 672 schizophrenics vs. 686 controls, and confirmed the significant association between GT-repeat polymorphism and disease [60]. Similarly, association of the GT-repeat polymorphism with schizophrenia was demonstrated in 122 Chinese sib-pair families [61]. Clinical studies demonstrated association of GT-

polymorphism in *GRIN2A* promoter and hippocampal and amygdala volumes [62], alcoholism [63], and D-serine level in schizophrenics [34]. In a study of 87 athletes suffering with a concussion, homozygous carriers of the longer alleles were six times more likely to experience longer recovery, compared with homozygous carriers of short ( $< 25$  repeats) alleles [32,64].

The GT repeat element is a part of *VWF* promoter coding for the von Willebrand factor, an essential plasma glycoprotein which mediates platelet adhesion and aggregation at the sites of vascular injury. Plasma concentration of VWF protein is under control of several factors including genetic polymorphism, and the dinucleotide tandem repeat (GT)<sub>n</sub> was hypothesized to influence the VWF level. An association of variable number of tandem GT repeats with the level of VWF was demonstrated in sixty-nine Cushing's syndrome patients [65]. Short repeats ( $n=15$ -19) were found more frequently in a group with high VWF induced by glucocorticoid excess, while long repeats ( $n=20$ -24) were predominant in a group with normal VWF. Risk of cortisol-induced increase of VWF was three times higher for allele's with 15-19 GT repeats (GT15-19) than for 20-24 repeats (GT20-24), and 13-fold higher for non-carriers of (GT20-24) alleles compared to non-carriers of (GT15-19) alleles. In a larger group of samples (1115 European male and female healthy controls), VWF:Ag values were lower in homozygous carriers of short allele ( $< 20$  repeats) when compared to heterozygous carriers of short and long alleles, or homozygous carriers of long ( $\geq 20$  repeats) allele. It should be mentioned that in a separate study in a group of 394 healthy individuals, the number of GT repeats did not correlate with VWF level [66]. These results were in contrast to greater VWF promoter activity under shear stress conditions when long GT repeats were present [67]. In this study, the bovine aortic endothelial cells (BAEC) were transfected with reporter construct prepared with different combinations of SNPs and GT repeats in the VWF promoter governing expression of a reporter gene. The upstream SNP haplotype did not affect promoter activation after shear stress. Rather, the promoters were more active when contained 23 repeats vs. 17 repeats. Importantly, the absence of GT repeats did not change the basal VWF promoter activity, but they were essential for the shear stress induced promoter activation [67]. One possible explanation for this discrepancy is different mechanisms for VWF upregulation after shear stress and cortisol induction [65].

Increased level of *FMR1* transcription was observed among premutation (61-200 CGG repeats) carriers of expanded CGG repeat in the 5'-untranslated region of the *FMR1* gene. The premutation form is highly unstable when transmitted from parent to child. Increase in the number of CGG repeats leads to the increase of *FMR1* transcript [68,69]. Interestingly, FMR protein levels decrease, likely due to poor initiation of translation at the downstream initiation codon [70,71]. A large study in 238 individuals confirmed a significant linear relationship between transcript level and CGG repeat size within the premutation size alleles [72]. Expansion of the repeat number beyond 200 results in hypermethylation and silencing of the gene which is phenotypically manifested as Fragile X Syndrome (FXS) [73].

An association between the microsatellite GT polymorphism in the intron 1 of *COL1A2* gene and bone mineral density was found in Chinese population after analysis of 388 nuclear families with a total of 1220 individuals [74]. In addition, an association was found between the number of GT repeats in *FOXP3* gene and several conditions including type 1 diabetes, development of severe acute graft versus host disease in patients transplanted from donors harboring short alleles, and renal allograft survival [35-37]. Importantly, no difference in the mean expression of

*FOXP3* mRNA was detected between asthmatic (n=49) and healthy (n=7) subjects grouped by GT repeat genotypes [38].

## Pharmacogenomics of STR

Phenotypic manifestation of variability in the length of STR remains poorly understood despite the fact that their association with neuromuscular and neurodegenerative diseases, several complex disorders, and several types of cancer has been convincingly demonstrated [6]. While genome-wide association studies of complex phenotypes are complicated, pharmacogenomic analysis of drug therapy provides a simpler model for deducing macro scale characteristics of a human organism based on his/her genome. Drug response is often determined by a variant of one or a few genes, for example those involved in drug metabolism, drug transport, or drug-target interaction [75-129]. Pharmacogenomic datasets include big populations of patients who have been medically characterized, have been treated with a standardized chemical stimulus (drug therapy), have records on the medical outcome, and for whom the genome sequence is often available.

Information about the effect of genetic variants on drug response is summarized in the PharmGKB database (<https://www.pharmgkb.org>). This database contains manually curated information about genetic variant-drug pairs based on individual PubMed publications. Because PharmGKB database collects information about all kinds of genetic variants (mostly SNPs), it is impossible at this time to focus this analysis on microsatellite length variations. Table 2 lists clinically used medications that may give observable clinical effects of STR polymorphism. The examples shown in the Table 2 substantiate future studies of relationship between microsatellite variants and drug response, and can be used as a starting point for a search of STR important for pharmacogenetic analysis.

Variable *UGT1A1* expression due to microsatellite polymorphism in the TATA box of its promoter is extensively studied, and *UGT1A1*\*28 allele is an important pharmacogenetic polymorphism [76-148]. The strong effect of additional TA repeats upstream the TSS is likely explained by the deviation of the TATA box from the canonical sequence. Expression analysis of (TA)<sub>n</sub> sequence in the *UGT1A1*\*28 allele provides mechanistic insights into the effect of polymorphic STR on transcription.

Reduced expression of EGFR mRNA associated with elongation of CA/TG STR located in intron 1 upstream of enhancer 2 of the *EGFR* gene is another well-documented example of STR effect on gene expression. Longer alleles for the CA SSR I repeat were associated with significantly lower EGFR expression, and predicted poor outcome of chemotherapy [77]. Evaluation of CA repeats in 62 EGFR somatic mutation-positive patients with advanced NSCLC treated with erlotinib demonstrated a significantly higher median progression free survival (HR=0.39, 0.22-0.70; p=0.002) and overall survival (HR = 0.43, 0.23-0.78); p=0.006) in patients harboring short CA repeat alleles (n≤16 repeats in any allele) compared to those with long alleles (n>16 in both alleles) [78]. The length of CA repeats was also a significant predictor for clinical outcome in 84 advanced NSCLC patients treated with gefitinib. The response rate of short CA repeat genotype was significantly higher (88.5% vs. 48.3%, p<0.001), and a combination of shorter CA repeat genotype with rs2293347GG had pronounced clinical benefit. More than 90% of patients with rs2293347GG and short CA repeat genotypes respond to gefitinib therapy, vs. 26% response in patients who carried longer CA repeats along with at least one rs2293347A allele [79]. Finally, the EGFR intron 1 CA repeat polymorphism was associated with survival

of 38 advanced gastric cancer patients treated with cetuximab (a monoclonal antibody targeting EGFR). Among 38 patients, twenty-one had short repeats (sum of both alleles ≤37), and 17 patients carried longer alleles (sum ≥38). The first category had longer progression-free survival (HR=0.42, 0.19-0.96; p=0.040) and overall survival (HR=0.40, 0.16-0.99; p=0.048) compared to the second category of patients. EGFR expression in the tumor tissues was higher in patients with short CA repeats [80].

In patients with colorectal cancer, 84% of patients with a sum of alleles of <35 developed an acneiform rash, compared with 33% of those with the sum of alleles of ≥35 (p=0.04) [43]. Association between the length of the CA SSR I and the response of locally advanced rectal cancer following adjuvant or neoadjuvant chemo radiation therapy was related to the additive effect of the EGFR R497K polymorphism and the length of CA SSR I. In addition, carriers of <20 CA repeats were more likely to show disease progression than were patients with ≥20 repeats (P< 0.05) when treated with 5-FU/oxaliplatin chemotherapy. These results suggest that the short CA SSR I alleles and, consequently, higher EGFR expression may predict for worse outcome after conventional therapy [81].

N-acetylcysteine (NAC) is used to improve the lung function of patients with COPD, and to reduce the risk of re-hospitalization. To explore a relationship between the effectiveness of oral NAC and the *HMOX1* promoter polymorphism in COPD patients, a total of 386 patients were genotyped, and were allocated to standard therapy plus NAC [57]. The non-carriers of L allele (>32 GT repeats) manifested improvements in forced expiratory volume in 1 second (FEV1) from 1.44 ± 0.37 to 1.58 ± 0.38 (P=0.04), and FEV1% predicted (from 56.6 ± 19.2 to 59.7 ± 17.2, P=0.03). The number of yearly COPD exacerbations in non-carriers of L allele was lower when compared with carriers of L allele (1.5 ± 0.66 vs. 2.1 ± 0.53, P<0.01). The improvement of the outcome of 6-min walking distance test was higher in non-carriers than in the carriers of L allele [57].

Variation in *HMOX1* expression due to genetic polymorphism is hypothesized to affect drug response, and therefore is a pharmacogenetic factor. Considering significant interest in potential modulators of HO-1 activity [82-86], detailed analysis of this genetic polymorphism is warranted. It is quite conceivable that genetic variants with different levels of HO-1 expression may change the therapeutic effect of several drug inducers of HO-1 including aspirin, statins, mimetic peptides, probucol, losartan, paclitaxel, rapamycin, cyclosporine [82].

## Conclusion

Variations in the STR length play an important role in modulating gene expression, and STR are likely to be general regulatory elements which attenuate expression of multiple genes. Moreover, regulatory STR manifests significant polymorphism because of their high intrinsic mutation rate. Because many genes with regulatory STR manifest variability in the expression level, it is now possible to assess effects of STR on gene expression by mining the existing databases. Several technical problems still remain to be addressed, such as relatively short reads generated by next-generation sequencing technologies, “stuttering” of DNA polymerase on STR sequences, insufficient accuracy of alignment through the monotonous repeats, and duplicating and compressing the sequencing data. Despite these hurdles, correlative analysis of gene expression versus STR allelic variants is quite possible. The STR catalogs have been generated using various approaches, and became useful tools to elucidate the role of STR in genome variability and evolution. Several important mechanistic questions remained unanswered: first, does the



length of microsatellites have a general effect on gene expression, or is it relevant only for certain types of promoters? Does this effect depend on the microsatellite orientation (e.g., CA repeat versus TG repeat)? What is the role of the adjacent sequences surrounding the microsatellite? Why does length variation in some STR enhance gene expression while in the others, it has an opposite effect? Finally, what protein factors facilitate the regulatory functions of STR?

Several facts substantiate further analysis of STR as potentially important candidates for pharmacogenetic analysis. First, these genetic elements are widely spread across the human genome where they are located at the beginning of the genes. Second, a significant proportion of microsatellites manifest genetic polymorphism, mostly the variable number of tandem repeats. The regulatory functions of repeats at the level of transcription, translation, biological activity, and clinical manifestation were convincingly demonstrated for multiple genes. Finally, many studies demonstrated an association between the number of repeats and the clinical effect, e.g. drug response. The number of the human genes with CA/TG repeats ( $n \geq 25$ ) within 1 kb region upstream from transcription start site exceeds 700, including ABC and SLC transporters, drug metabolizing enzymes, and drug targets. Elucidating the effects of STR on gene expression may help explain variability in drug response, something that is not achieved by focusing exclusively on SNPs or CNV. This will be the next step toward deducing the macro characteristics of an organism from its genome, a problem brilliantly solved by myriads of fertilized egg cells.

## References

- de Koning AP, Gu W, Castoe TA, Batzer MA, Pollock DD (2011) Repetitive elements may comprise over two-thirds of the human genome. *PLoS Genet* 7: e1002384.
- Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. (2001) Initial sequencing and analysis of the human genome. *Nature* 409: 860-921.
- Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 5: 435-445.
- Willems T, Gymrek M, Highnam G, 1000 Genomes Project Consortium, Mittelman D, et al. (2014) The landscape of human STR variation. *Genome Res* 24: 1894-1904.
- Subramanian S, Mishra RK, Singh L (2003) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biol* 4: R13.
- Bolton KA, Ross JP, Grice DM, Bowden NA, Holliday EG, et al. (2013) STARRRT: a table of short tandem repeats in regulatory regions of the human genome. *BMC Genomics* 14: 795-2164-14-795.
- Sawaya S, Bagshaw A, Buschiazzi E, Kumar P, Chowdhury S, et al. (2013) Microsatellite tandem repeats are abundant in human promoters and are associated with regulatory elements. *PLoS ONE* 8: e54710.
- Gemayel R, Vines MD, Legendre M, Verstrepen KJ (2010) Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annu Rev Genet* 44: 445-477.
- Gymrek M, Willems T, Guilmatre A, Zeng H, Markus B, et al. (2016) Abundant contribution of short tandem repeats to gene expression variation in humans. *Nat Genet* 48: 22-29.
- Segal E, Widom J (2009) Poly (dA:dT) tracts: major determinants of nucleosome organization. *Curr Opin Struct Biol* 19: 65-71.
- Naslund K, Saetre P, von Salome J, Bergstrom TF, Jareborg N, et al. (2005) Genome-wide prediction of human VNTRs. *Genomics* 85: 24-35.
- Legendre M, Pochet N, Pak T, Verstrepen KJ (2007) Sequence-based estimation of minisatellite and microsatellite repeat variability. *Genome Res* 17: 1787-1796.
- Verstrepen KJ, Jansen A, Lewitter F, Fink GR (2005) Intragenic tandem repeats generate functional variability. *Nat Genet* 37: 986-990.
- Messier W, Li SH, Stewart CB (1996) The birth of microsatellites. *Nature* 381: 483.
- Kelkar YD, Strubczewski N, Hile SE, Chiaromonte F, Eckert KA, et al. (2010) What is a microsatellite: a computational and experimental definition based upon repeat mutational behavior at A/T and GT/AC repeats. *Genome Biol Evol* 2: 620-635.
- Duitama J, Zablotskaya A, Gemayel R, Jansen A, Belet S, et al. (2014) Large-scale analysis of tandem repeat variability in the human genome. *Nucleic Acids Res* 42: 5728-5741.
- Paques F, Leung WY, Haber JE (1998) Expansions and contractions in a tandem repeat induced by double-strand break repair. *Mol Cell Biol* 18: 2045-2054.
- Quilez J, Guilmatre A, Garg P, Highnam G, Gymrek M, et al. (2016) Polymorphic tandem repeats within gene promoters act as modifiers of gene expression and DNA methylation in humans. *Nucleic Acids Res* 44: 3750-3762.
- Krynetskiy E (2014) Pharmacogenomics of Simple Repeats: How Do You Solve a Problem like VNSR? *Journal of Pharmacogenomics and Pharmacoproteomics* 5: 1000e139.
- Kimpara T, Takeda A, Watanabe K, Itoyama Y, Ikawa S, et al. (1997) Microsatellite polymorphism in the human heme oxygenase-1 gene promoter and its application in association studies with Alzheimer and Parkinson disease. *Hum Genet* 100: 145-147.
- Chen YH, Lin SJ, Lin MW, Tsai HL, Kuo SS, et al. (2002) Microsatellite polymorphism in promoter of heme oxygenase-1 gene is associated with susceptibility to coronary artery disease in type 2 diabetic patients. *Hum Genet* 111: 1-8.
- Okamoto I, Krogler J, Endler G, Kaufmann S, Mustafa S, et al. (2006) A microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with risk for melanoma. *Int J Cancer* 119: 1312-1315.
- Yamada N, Yamaya M, Okinaga S, Nakayama K, Sekizawa K, et al. (2000) Microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with susceptibility to emphysema. *Am J Hum Genet* 66: 187-195.
- Miyake M, Ishii M, Kawashima K, Kodama T, Sugano K, et al. (2009) siRNA-mediated knockdown of the heme synthesis and degradation pathways: modulation of treatment effect of 5-aminolevulinic acid-based photodynamic therapy in urothelial cancer cell lines. *Photochem Photobiol* 85: 1020-1027.
- Taha H, Skrzypek K, Guevara I, Nigisch A, Mustafa S, et al. (2010) Role of heme oxygenase-1 in human endothelial cells: lesson from the promoter allelic variants. *Arterioscler Thromb Vasc Biol* 30: 1634-1641.
- Doberer D, Haschemi A, Andreas M, Zapf TC, Clive B, et al. (2010) Haem arginate infusion stimulates haem oxygenase-1 expression in healthy subjects. *Br J Pharmacol* 161: 1751-1762.
- Walther M, De Caul A, Aka P, Njie M, Amambua-Ngwa A, et al. (2012) HMOX1 gene promoter alleles and high HO-1 levels are associated with severe malaria in Gambian children. *PLoS Pathog* 8: e1002579.
- Hirai H, Kubo H, Yamaya M, Nakayama K, Numasaki M, et al. (2003) Microsatellite polymorphism in heme oxygenase-1 gene promoter is associated with susceptibility to oxidant-induced apoptosis in lymphoblastoid cell lines. *Blood* 102: 1619-1621.
- Brydun A, Watari Y, Yamamoto Y, Okuhara K, Teragawa H, et al. (2007) Reduced expression of heme oxygenase-1 in patients with coronary atherosclerosis. *Hypertens Res* 30: 341-348.
- Rueda B, Oliver J, Robledo G, Lopez-Nevot MA, Balsa A, et al. (2007) HO-1 promoter polymorphism associated with rheumatoid arthritis. *Arthritis Rheum* 56: 3953-3958.
- Kramer M, Sponholz C, Slaba M, Wissuwa B, Claus RA, et al. (2013) Alternative 5' untranslated regions are involved in expression regulation of human heme oxygenase-1. *PLoS One* 8: e77224.
- McDevitt J, Krynetskiy E (2017) Genetic findings in sport-related concussions: potential for individualized medicine? *Concussion-Future Medicine* 10.2217/cnc-2016-0020.
- Itokawa M, Yamada K, Yoshitsugu K, Toyota T, Suga T, et al. (2003) A microsatellite repeat in the promoter of the N-methyl-D-aspartate receptor 2A subunit (GRIN2A) gene suppresses transcriptional activity and correlates with chronic outcome in schizophrenia. *Pharmacogenetics* 13: 271-278.



34. Liu R, Dang W, Du Y, Zhou Q, Liu Z, et al. (2015) Correlation of functional GRIN2A gene promoter polymorphisms with schizophrenia and serum D-serine levels. *Gene* 568: 25-30.
35. Bassuny WM, Ihara K, Sasaki Y, Kuromaru R, Kohno H, et al. (2003) A functional polymorphism in the promoter/enhancer region of the FOXP3/Scurfin gene associated with type 1 diabetes. *Immunogenetics* 55: 149-156.
36. Noriega V, Martinez-Laperche C, Buces E, Pion M, Sanchez-Hernandez N, et al. (2015) The genotype of the donor for the (GT)<sub>n</sub> polymorphism in the promoter/enhancer of FOXP3 is associated with the development of severe acute GVHD but does not affect the GVL effect after myeloablative HLA-identical allogeneic stem cell transplantation. *PLoS ONE* 10: e0140454.
37. Engela AU, Boer K, Roodnat JI, Peeters AM, Eilers PH, et al. (2013) Genetic variants of FOXP3 influence graft survival in kidney transplant patients. *Hum Immunol* 74: 751-757.
38. Karagiannidis C, Akdis M, Holopainen P, Woolley NJ, Hense G, et al. (2004) Glucocorticoids upregulate FOXP3 expression and regulatory T cells in asthma. *J Allergy Clin Immunol* 114: 1425-1433.
39. Akai J, Kimura A, Hata RI (1999) Transcriptional regulation of the human type I collagen alpha2 (COL1A2) gene by the combination of two dinucleotide repeats. *Gene* 239: 65-73.
40. Buerger H, Gebhardt F, Schmidt H, Beckmann A, Huttmacher K, et al. (2000) Length and loss of heterozygosity of an intron 1 polymorphic sequence of egfr is related to cytogenetic alterations and epithelial growth factor receptor expression. *Cancer Res* 60: 854-857.
41. Buerger H, Packeisen J, Boecker A, Tidow N, Kersting C, et al. (2004) Allelic length of a CA dinucleotide repeat in the egfr gene correlates with the frequency of amplifications of this sequence--first results of an inter-ethnic breast cancer study. *J Pathol* 203: 545-550.
42. Gebhardt F, Zanker KS, Brandt B (1999) Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1. *J Biol Chem* 274: 13176-13180.
43. Amador ML, Oppenheimer D, Perea S, Maitra A, Cusatis G, et al. (2004) An epidermal growth factor receptor intron 1 polymorphism mediates response to epidermal growth factor receptor inhibitors. *Cancer Res* 64: 9139-9143.
44. Nie Q, Yang XN, An SJ, Zhang XC, Yang JJ, et al. (2011) CYP1A1\*2A polymorphism as a predictor of clinical outcome in advanced lung cancer patients treated with EGFR-TKI and its combined effects with EGFR intron 1 (CA)<sub>n</sub> polymorphism. *Eur J Cancer* 47: 1962-1970.
45. Suryadevara NC, Neela VS, Devalraju KP, Jain S, SivaSai KS, et al. (2013) Influence of Intron II microsatellite polymorphism in human toll-like receptor 2 gene in leprosy. *Hum Immunol* 74: 1034-1040.
46. Lamba V, Panetta JC, Strom S, Schuetz EG (2010) Genetic predictors of interindividual variability in hepatic CYP3A4 expression. *J Pharmacol Exp Ther* 332: 1088-1099.
47. Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y (2013) Identifying personal genomes by surname inference. *Science* 339: 321-324.
48. Gatchel JR, Zoghbi HY (2005) Diseases of unstable repeat expansion: mechanisms and common principles. *Nat Rev Genet* 6: 743-755.
49. Fondon JW, Hammock EA, Hannan AJ, King DG (2008) Simple sequence repeats: genetic modulators of brain function and behavior. *Trends Neurosci* 31: 328-334.
50. Exner M, Minar E, Wagner O, Schillinger M (2004) The role of heme oxygenase-1 promoter polymorphisms in human disease. *Free Radic Biol Med* 37: 1097-1104.
51. Dick IM, Devine A, Prince RL (2005) Association of an aromatase TTTA repeat polymorphism with circulating estrogen, bone structure, and biochemistry in older women. *Am J Physiol Endocrinol Metab* 288: E989-95.
52. Song F, Li X, Zhang M, Yao P, Yang N, et al. (2009) Association between heme oxygenase-1 gene promoter polymorphisms and type 2 diabetes in a Chinese population. *Am J Epidemiol* 170: 747-756.
53. Bao W, Song F, Li X, Rong S, Yang W, et al. (2010) Association between heme oxygenase-1 gene promoter polymorphisms and type 2 diabetes mellitus: a HuGE review and meta-analysis. *Am J Epidemiol* 172: 631-636.
54. Bean CJ, Boulet SL, Ellingsen D, Pyle ME, Barron-Casella EA, et al. (2012) Heme oxygenase-1 gene promoter polymorphism is associated with reduced incidence of acute chest syndrome among children with sickle cell disease. *Blood* 120: 3822-3828.
55. Slachtova L, Kaminska D, Chval M, Kralik L, sMartasek P, et al. (2013) Stress perception and (GT)<sub>n</sub> repeat polymorphism in haem oxygenase 1 promoter are both risk factors in development of eating disorders. *Folia Biol (Praha)* 59: 233-239.
56. Repesse Y, Peyron I, Dimitrov JD, Dasgupta S, Moshai EF, et al. (2013) Development of inhibitory antibodies to therapeutic factor VIII in severe hemophilia A is associated with microsatellite polymorphisms in the HMOX1 promoter. *Haematologica* 98: 1650-1655.
57. Zhang JQ, Zhang JQ, Fang LZ, Liu L, Fu WP, et al. (2015) Effect of oral N-acetylcysteine on COPD patients with microsatellite polymorphism in the heme oxygenase-1 gene promoter. *Drug Des Devel Ther* 9: 6379-6387.
58. Zhang ZY, Guan J, Li H, Zhou ZQ, Zhou GW (2016) Heme Oxygenase-1 Promoter Polymorphism Protects Liver Allograft. *Indian J Surg* 78: 14-19.
59. Daenen KE, Martens P, Bammens B (2016) Association of HO-1 (GT)<sub>n</sub> Promoter Polymorphism and Cardiovascular Disease: A Reanalysis of the Literature. *Can J Cardiol* 32: 160-168.
60. Iwayama-Shigeno Y, Yamada K, Itokawa M, Toyota T, Meerabux JM, et al. (2005) Extended analyses support the association of a functional (GT)<sub>n</sub> polymorphism in the GRIN2A promoter with Japanese schizophrenia. *Neurosci Lett* 378: 102-105.
61. Tang J, Chen X, Xu X, Wu R, Zhao J, et al. (2006) Significant linkage and association between a functional (GT)<sub>n</sub> polymorphism in promoter of the N-methyl-D-aspartate receptor subunit gene (GRIN2A) and schizophrenia. *Neurosci Lett* 409: 80-82.
62. Inoue H, Yamasue H, Tochigi M, Suga M, Iwayama Y, et al. (2010) Functional (GT)<sub>n</sub> polymorphisms in promoter region of N-methyl-d-aspartate receptor 2A subunit (GRIN2A) gene affect hippocampal and amygdala volumes. *Genes Brain Behav* 9: 269-275.
63. Domart MC, Benyamina A, Lemoine A, Bourgain C, Blecha L, et al. (2012) Association between a polymorphism in the promoter of a glutamate receptor subunit gene (GRIN2A) and alcoholism. *Addict Biol* 17: 783-785.
64. McDevitt J, Tierney RT, Phillips J, Gaughan JP, Torg JS, et al. (2015) Association between GRIN2A promoter polymorphism and recovery from concussion. *Brain Inj* 29: 1674-1681.
65. Daidone V, Pontara E, Romualdi C, Cattini MG, Scaroni C, et al. (2010) Microsatellite (GT)<sub>n</sub> is part of the von Willebrand factor (VWF) promoter region that influences the glucocorticoid-induced increase in VWF in Cushing's syndrome. *Thromb Res* 125: e275-80.
66. Daidone V, Cattini MG, Pontara E, Sartorello F, Gallinaro L, et al. (2009) Microsatellite (GT)<sub>n</sub> repeats and SNPs in the von Willebrand factor gene promoter do not influence circulating von Willebrand factor levels under normal conditions. *Thromb Haemost* 101: 298-304.
67. Hough C, Cameron CL, Notley CR, Brown C, O'Brien L, et al. (2008) Influence of a GT repeat element on shear stress responsiveness of the VWF gene promoter. *J Thromb Haemost* 6: 1183-1190.
68. Tassone F, Hagerman RJ, Loesch DZ, Lachiewicz A, Taylor AK, et al. (2000) Fragile X males with unmethylated, full mutation trinucleotide repeat expansions have elevated levels of FMR1 messenger RNA. *Am J Med Genet* 94: 232-236.
69. Kenneson A, Zhang F, Hagedorn CH, Warren ST (2001) Reduced FMRP and increased FMR1 transcription is proportionally associated with CGG repeat number in intermediate-length and premutation carriers. *Hum Mol Genet* 10: 1449-1454.
70. Feng Y, Zhang F, Lokey LK, Chastain JL, Lakkis L, et al. (1995) Translational suppression by trinucleotide repeat expansion at FMR1. *Science* 268: 731-734.
71. Primerano B, Tassone F, Hagerman RJ, Hagerman P, Amaldi F, et al. (2002) Reduced FMR1 mRNA translation efficiency in fragile X patients with premutations. *RNA* 8: 1482-1488.
72. Allen EG, He W, Yadav-Shah M, Sherman SL (2004) A study of the distributional characteristics of FMR1 transcript levels in 238 individuals. *Hum Genet* 114: 439-447.
73. Crawford DC, Acuna JM, Sherman SL (2001) FMR1 and the fragile X syndrome: human genome epidemiology review. *Genet Med* 3: 359-371.
74. Lei SF, Deng FY, Dvornyk V, Liu MY, Xiao SM, et al. (2005) The (GT)

- n polymorphism and haplotype of the COL1A2 gene, but not the (AAAG)<sub>n</sub> polymorphism of the PTHR1 gene, are associated with bone mineral density in Chinese. *Hum Genet* 116: 200-207.
75. Krynetskiy E (2012) At the Origin of a Never-ending Story. *Journal of Pharmacogenomics and Pharmacoproteomics* 3: 1000e107.
  76. Biason P, Masier S, Toffoli G (2008) UGT1A1\*28 and other UGT1A polymorphisms as determinants of irinotecan toxicity. *J Chemother* 20: 158-165.
  77. Brandt B, Meyer-Staeckling S, Schmidt H, Agelopoulos K, Buerger H (2006) Mechanisms of egfr gene transcription modulation: relationship to cancer risk and therapy response. *Clin Cancer Res* 12: 7252-7260.
  78. Winther Larsen A, Nissen PH, Meldgaard P, Weber B, Sorensen BS (2014) EGFR CA repeat polymorphism predict clinical outcome in EGFR mutation positive NSCLC patients treated with erlotinib. *Lung Cancer* 85: 435-441.
  79. Ma F, Sun T, Shi Y, Yu D, Tan W, et al. (2009) Polymorphisms of EGFR predict clinical outcome in advanced non-small-cell lung cancer patients treated with Gefitinib. *Lung Cancer* 66: 114-119.
  80. Han SW, Oh DY, Im SA, Park SR, Lee KW, et al. (2010) Epidermal growth factor receptor intron 1 CA dinucleotide repeat polymorphism and survival of advanced gastric cancer patients treated with cetuximab plus modified FOLFOX6. *Cancer Sci* 101: 793-799.
  81. Zhang W, Park DJ, Lu B, Yang DY, Gordon M, et al. (2005) Epidermal growth factor receptor gene polymorphisms predict pelvic recurrence in patients with rectal cancer treated with chemoradiation. *Clin Cancer Res* 11: 600-605.
  82. Abraham NG, Kappas A (2008) Pharmacological and clinical aspects of heme oxygenase. *Pharmacol Rev* 60: 79-127.
  83. Pittala V, Salerno L, Romeo G, Modica MN, Siracusa MA (2013) A focus on heme oxygenase-1 (HO-1) inhibitors. *Curr Med Chem* 20: 3711-3732.
  84. Liao G, Li R, Chen X, Zhang W, Du S, et al. (2016) Sodium valproate prevents radiation-induced injury in hippocampal neurons via activation of the Nrf2/HO-1 pathway. *Neuroscience* 331: 40-51.
  85. Alam A, Mukhopadhyay ND, Ning Y, Reshko LB, Cardnell RJ, et al. (2015) A Preliminary Study on Racial Differences in HMOX1, NFE2L2, and TGFbeta1 Gene Polymorphisms and Radiation-Induced Late Normal Tissue Toxicity. *Int J Radiat Oncol Biol Phys* 93: 436-443.
  86. Chen WJ, Chen YH, Lai YJ, Hsu YJ, Yeh YH, et al. (2016) GT-repeat length polymorphism in heme oxygenase-1 promoter determines the effect of cilostazol on vascular smooth muscle cells. *Int J Cardiol* 222: 407-415.
  87. Beutler E, Gelbart T, Demina A (1998) Racial variability in the UDP-glucuronosyltransferase 1 (UGT1A1) promoter: a balanced polymorphism for regulation of bilirubin metabolism? *Proc Natl Acad Sci U S A* 95: 8170-8174.
  88. Rajmakers MT, Jansen PL, Steegers EA, Peters WH (2000) Association of human liver bilirubin UDP-glucuronyltransferase activity with a polymorphism in the promoter region of the UGT1A1 gene. *J Hepatol* 33: 348-351.
  89. Iyer L, Das S, Janisch L, Wen M, Ramirez J, et al. (2002) UGT1A1\*28 polymorphism as a determinant of irinotecan disposition and toxicity. *Pharmacogenomics J* 2: 43-47.
  90. Gao PS, Heller NM, Walker W, Chen CH, Moller M, et al. (2004) Variation in dinucleotide (GT) repeat sequence in the first exon of the STAT6 gene is associated with atopic asthma and differentially regulates the promoter activity in vitro. *J Med Genet* 41: 535-539.
  91. Hickson N, Hampshire D, Castaman G, Eikenboom J, Rodeghiero F, et al. (2011) Effect of the VWF promoter (GT)<sub>n</sub> repeat and single-nucleotide polymorphism c.-2527G>A on circulating von Willebrand factor levels under normal conditions. *J Thromb Haemost* 9: 603-605.
  92. Shimajiri S, Arima N, Tanimoto A, Murata Y, Hamada T, et al. (1999) Shortened microsatellite d(CA)<sub>21</sub> sequence down-regulates promoter activity of matrix metalloproteinase 9 gene. *FEBS Lett* 455: 70-74.
  93. Peters DG, Kassam A, St Jean PL, Yonas H, Ferrell RE (1999) Functional polymorphism in the matrix metalloproteinase-9 promoter as a potential risk factor for intracranial aneurysm. *Stroke* 30: 2612-2616.
  94. Maeda S, Haneda M, Guo B, Koya D, Hayashi K, et al. (2001) Dinucleotide repeat polymorphism of matrix metalloproteinase-9 gene is associated with diabetic nephropathy. *Kidney Int* 60: 1428-1434.
  95. Pravica V, Asderakis A, Perrey C, Hajeer A, Sinnott PJ, et al. (1999) In vitro production of IFN-gamma correlates with CA repeat polymorphism in the human IFN-gamma gene. *Eur J Immunogenet* 26: 1-3.
  96. Saha A, Dhir A, Ranjan A, Gupta V, Bairwa N, et al. (2005) Functional IFNG polymorphism in intron 1 in association with an increased risk to promote sporadic breast cancer. *Immunogenetics* 57: 165-171.
  97. Dwivedi M, Laddha NC, Shah K, Shah BJ, Begum R (2013) Involvement of interferon-gamma genetic variants and intercellular adhesion molecule-1 in onset and progression of generalized vitiligo. *J Interferon Cytokine Res* 33: 646-659.
  98. Yim JJ, Lee HW, Lee HS, Kim YW, Han SK, et al. (2006) The association between microsatellite polymorphisms in intron II of the human Toll-like receptor 2 gene and tuberculosis among Koreans. *Genes Immun* 7: 150-155.
  99. Xiao Z, Nunome K, Yahara T, Inoue E, Nabeshima M, et al. (2014) Comparative studies of human UDP-glucuronosyltransferase 1A8 and 1A9 proximal promoters using single base substitutions. *Drug Metab Pharmacokinet* 29: 90-93.
  100. Nichenametla SN, Lazarus P, Richie JP (2011) A GAG trinucleotide-repeat polymorphism in the gene for glutathione biosynthetic enzyme, GCLC, affects gene expression through translation. *FASEB J* 25: 2180-2187.
  101. Buttica C, Gysin R, Cuenod M, Do KQ (2011) Interaction of GAG trinucleotide repeat and C-129T polymorphisms impairs expression of the glutamate-cysteine ligase catalytic subunit gene. *Free Radic Biol Med* 50: 617-623.
  102. Sano M, Inoue S, Hosoi T, Ouchi Y, Emi M, et al. (1995) Association of estrogen receptor dinucleotide repeat polymorphism with osteoporosis. *Biochem Biophys Res Commun* 217: 378-383.
  103. Cai Q, Gao YT, Wen W, Shu XO, Jin F, et al. (2003) Association of breast cancer risk with a GT dinucleotide repeat polymorphism upstream of the estrogen receptor-alpha gene. *Cancer Res* 63: 5727-5730.
  104. Gade-Andavolu R, Macmurray J, Comings DE, Calati R, Chiesa A, et al. (2009) Association between the estrogen receptor TA polymorphism and Harm avoidance. *Neurosci Lett* 467: 155-158.
  105. Pinsonneault JK, Sullivan D, Sadee W, Soares CN, Hampson E, et al. (2013) Association study of the estrogen receptor gene ESR1 with postpartum depression—a pilot study. *Arch Womens Ment Health* 16: 499-509.
  106. Golubic K, Smalcelj A, Sertic J, Juricic L (2014) Estrogen receptor 1 gene (TA)<sub>n</sub> polymorphism is associated with lone atrial fibrillation in men. *Croat Med J* 55: 38-44.
  107. Matsuyama N, Mishihiro S, Sugimoto M, Furuichi Y, Hashimoto M, et al. (2003) The dinucleotide microsatellite polymorphism of the IFNAR1 gene promoter correlates with responsiveness of hepatitis C patients to interferon. *Hepatology Res* 25: 221-225.
  108. Yoshida K, Alagbe O, Wang X, Woolwine B, Thornbury M, et al. (2005) Promoter polymorphisms of the interferon-alpha receptor gene and development of Interferon-induced depressive symptoms in patients with chronic hepatitis C: preliminary findings. *Neuropsychobiology* 52: 55-61.
  109. Kanchan K, Jha P, Pati SS, Mohanty S, Mishra SK, et al. (2015) Interferon-gamma (IFNG) microsatellite repeat and single nucleotide polymorphism haplotypes of IFN-alpha receptor (IFNAR1) associated with enhanced malaria susceptibility in Indian populations. *Infect Genet Evol* 29: 6-14.
  110. Zecevic M, Amos CI, Gu X, Campos IM, Jones JS, et al. (2006) IGF1 gene polymorphism and risk for hereditary nonpolyposis colorectal cancer. *J Natl Cancer Inst* 98: 139-143.
  111. Reeves SG, Rich D, Meldrum CJ, Colyvas K, Kurzawski G, et al. (2008) IGF1 is a modifier of disease risk in hereditary non-polyposis colorectal cancer. *Int J Cancer* 123: 1339-1343.
  112. Chen HY, Huang W, Leung VH, Fung SL, Ma SL, et al. (2013) Functional interaction between SNPs and microsatellite in the transcriptional regulation of insulin-like growth factor 1. *Hum Mutat* 34: 1289-1297.
  113. Kaczmarek M, Pacholska-Bogalska J, Kwasniewski W, Kotarski J, Halerz-Nowakowska B, et al. (2015) A microsatellite polymorphism in IGF1 gene promoter and timing of natural menopause in Caucasian women. *Int J Med Sci* 12: 32-41.
  114. Kwasniewski W, Gozdzicka-Jozefiak A, Wolun-Cholewa M, Polak G, Sierocinska-Sawa J, et al. (2016) Microsatellite polymorphism in the P1

- promoter region of the IGF1 gene is associated with endometrial cancer. *Mol Med Rep* 13: 4950-4958.
115. Borrás T, Bryant PA, Chisolm SS (2006) First look at the effect of overexpression of TIGR/MYOC on the transcriptome of the human trabecular meshwork. *Exp Eye Res* 82: 1002-1010.
  116. Pova CA, Malta RF, Rezende Mde M, de Melo KF, Giannella-Neto D (2006) Correlation between genotype and phenotype in primary open angle glaucoma of Brazilian families with mutations in exon 3 of the TIGR/MYOC gene. *Arq Bras Oftalmol* 69: 289-297.
  117. Rahmanian M, Kargar M (2014) Tumor necrosis factor-alpha polymorphism and susceptibility to multiple sclerosis in the Iranian population. *Iran Red Crescent Med J* 17: e18247.
  118. Wang P, Wang J, Yu M, Li Z (2016) Tumor Necrosis Factor-alpha T-857C (rs1799724) Polymorphism and Risk of Cancers: A Meta-Analysis. *Dis Markers* 2016: 4580323.
  119. Hua XP, Qian J, Cao CB, Xie J, Zeng XT, et al. (2016) Association between TNF-alpha rs1800629 polymorphism and the risk of myocardial infarction: A meta-analysis. *Genet Mol Res* 15: 10.4238/gmr.15037292.
  120. Santos AC, de Moura EL, Ferreira JM, Santos BR, Alves VM, et al. (2017) Meta-Analysis of the Relationship between TNF-alpha (-308G/A) and IL-10 (-819C/T) Gene Polymorphisms and Susceptibility to Dengue. *Immunol Invest* 46: 201-220.
  121. Yang ZC, Xu F, Tang M, Xiong X (2017) Association Between TNF-alpha Promoter-308 A/G Polymorphism and Systemic Lupus Erythematosus Susceptibility: A Case-Control Study and Meta-Analysis. *Scand J Immunol* 85: 197-210.
  122. Dietzsch E, Parker MI (1999) CA repeat polymorphism in the promoter region of the COL1A2 gene. *J Hum Genet* 44: 419-420.
  123. Hata R, Akai J, Kimura A, Ishikawa O, Kuwana M, et al. (2000) Association of functional microsatellites in the human type I collagen alpha2 chain (COL1A2) gene with systemic sclerosis. *Biochem Biophys Res Commun* 272: 36-40.
  124. Ramirez F, Tanaka S, Bou-Gharios G (2006) Transcriptional regulation of the human alpha2(I) collagen gene (COL1A2), an informative model system to study fibrotic diseases. *Matrix Biol* 25: 365-372.
  125. Costa BM, Viana-Pereira M, Fernandes R, Costa S, Linhares P, et al. (2011) Impact of EGFR genetic variants on glioma risk and patient outcome. *Cancer Epidemiol Biomarkers Prev* 20: 2610-2617.
  126. Lai CY, Sung FC, Hsieh LL, Tang R, Chiou HY, et al. (2013) Associations between genetic polymorphisms of epidermal growth factor receptor (EGFR) and survival of colorectal cancer (CRC) patients treated with 5-fluorouracil-based chemotherapy. *Ann Surg Oncol* 20 Suppl 3: S599-606.
  127. Chen B, Luo J, Gu W, Shen L, Wang H, et al. (2015) Shorter EGFR Dinucleotide Repeat Length Predicts Better Response of Patients with Advanced Non-small Cell Lung Cancer to EGFR Tyrosine Kinase Inhibitor. *Cell Biochem Biophys* 73: 799-804.
  128. Conti S, Gallo E, Sioletic S, Facciolo F, Palmieri G, et al. (2016) Molecular genetic alterations in egfr CA-SSR-1 microsatellite and egfr copy number changes are associated with aggressiveness in thymoma. *J Thorac Dis* 8: 386-395.
  129. Jin P, Warren ST (2000) Understanding the molecular basis of fragile X syndrome. *Hum Mol Genet* 9: 901-908.
  130. Yang W, Fan C, Chen L, Cui Z, Bai Y, et al. (2016) Pathological Effects of the FMR1 CGG-Repeat Polymorphism (5-55 Repeat Numbers): Systematic Review and Meta-Analysis. *Tohoku J Exp Med* 239: 57-66.
  131. Tabassum R, Chavali S, Dwivedi OP, Tandon N, Bharadwaj D (2008) Genetic variants of FOXA2: risk of type 2 diabetes and effect on metabolic traits in North Indians. *J Hum Genet* 53: 957-965.
  132. Bekris LM, Shephard C, Janer M, Graham J, McNeney B, et al. (2007) Glutamate cysteine ligase catalytic subunit promoter polymorphisms and associations with type 1 diabetes age-at-onset and GAD65 autoantibody levels. *Exp Clin Endocrinol Diabetes* 115: 221-228.
  133. Gysin R, Kraftsik R, Sandell J, Bovet P, Chappuis C, et al. (2007) Impaired glutathione synthesis in schizophrenia: convergent genetic and functional evidence. *Proc Natl Acad Sci U S A* 104: 16621-16626.
  134. Nichenametla SN, Muscat JE, Liao JG, Lazarus P, Richie JP, Jr (2013) A functional trinucleotide repeat polymorphism in the 5'-untranslated region of the glutathione biosynthetic gene GCLC is associated with increased risk for lung and aerodigestive tract cancers. *Mol Carcinog* 52: 791-799.
  135. D'Silva S, Borse V, Colah RB, Ghosh K, Mukherjee MB (2011) Association of (GT)<sub>n</sub> repeats promoter polymorphism of heme oxygenase-1 gene with serum bilirubin levels in healthy Indian adults. *Genet Test Mol Biomarkers* 15: 215-218.
  136. Bestach Y, Sieza Y, Attie M, Riccheri C, Verri V, et al. (2015) Polymorphisms in TNF and IFNG are associated with clinical characteristics of aplastic anemia in Argentinean population. *Leuk Lymphoma* 56: 1793-1798.
  137. Kader AK, Liu J, Shao L, Dinney CP, Lin J, et al. (2007) Matrix metalloproteinase polymorphisms are associated with bladder cancer invasiveness. *Clin Cancer Res* 13: 2614-2620.
  138. Tamura K, Suzuki M, Arakawa H, Tokuyama K, Morikawa A (2003) Linkage and association studies of STAT6 gene polymorphisms and allergic diseases. *Int Arch Allergy Immunol* 131: 33-38.
  139. Nagarkatti R, B-Rao C, Vijayan V, Sharma SK, Ghosh B (2004) Signal transducer and activator of transcription 6 haplotypes and asthma in the Indian population. *Am J Respir Cell Mol Biol* 31: 317-321.
  140. Qian X, Gao Y, Ye X, Lu M (2014) Association of STAT6 variants with asthma risk: a systematic review and meta-analysis. *Hum Immunol* 75: 847-853.
  141. Lee EY, Yim JJ, Lee HS, Lee YJ, Lee EB, et al. (2006) Dinucleotide repeat polymorphism in intron II of human Toll-like receptor 2 gene and susceptibility to rheumatoid arthritis. *Int J Immunogenet* 33: 211-215.
  142. Yim JJ, Ding L, Schaffer AA, Park GY, Shim YS, et al. (2004) A microsatellite polymorphism in intron 2 of human Toll-like receptor 2 gene: functional implications and racial differences. *FEMS Immunol Med Microbiol* 40: 163-169.
  143. Bochud PY, Hawn TR, Siddiqui MR, Saunderson P, Britton S, et al. (2008) Toll-like receptor 2 (TLR2) polymorphisms are associated with reversal reaction in leprosy. *J Infect Dis* 197: 253-261.
  144. Takagi Y, Masamune A, Kume K, Satoh A, Kikuta K, et al. (2009) Microsatellite polymorphism in intron 2 of human Toll-like receptor 2 gene is associated with susceptibility to acute pancreatitis in Japan. *Hum Immunol* 70: 200-204.
  145. Nischalke HD, Berger C, Aldenhoff K, Thyssen L, Gentemann M, et al. (2011) Toll-like receptor (TLR) 2 promoter and intron 2 polymorphisms are associated with increased risk for spontaneous bacterial peritonitis in liver cirrhosis. *J Hepatol* 55: 1010-1016.
  146. Kutikhin AG (2011) Association of polymorphisms in TLR genes and in genes of the Toll-like receptor signaling pathway with cancer risk. *Hum Immunol* 72: 1095-1116.
  147. Bosma PJ, Chowdhury JR, Bakker C, Gantla S, de Boer A, et al. (1995) The genetic basis of the reduced expression of bilirubin UDP-glucuronosyltransferase 1 in Gilbert's syndrome. *N Engl J Med* 333: 1171-1175.
  148. Liu W, Ramirez J, Gamazon ER, Mirkov S, Chen P, et al. (2014) Genetic factors affecting gene transcription and catalytic activity of UDP-glucuronosyltransferases in human liver. *Hum Mol Genet* 23: 5558-5569.