Functional Significance of Microsatellite Markers

Tatjana Sjakste¹, Natalia Paramonova¹, Nikolajs Sjakste²

¹Genomics and Bioinformatics, Institute of Biology, University of Latvia, Latvia
²Faculty of Medicine, University of Latvia, Latvia

Key Words: microsatellite; human diseases; promoter; intron; exon.

Summary. The review summarizes literature data on the positive results of association studies between the length of microsatellite repeats and predisposition to pathologies. Actually, the data can be classified according to the localization of the microsatellite: in the gene promoter, in the part of exon 1 coding the signal sequence, in gene introns, in the coding areas of genes, and in 3'-untranslated regions. The functional significance of microsatellite length changes can be evaluated in many cases. The authors came up to the conclusion that further studies on microsatellite associations with diseases remain prospective as they reflect changes in the gene functional activity.

Introduction

Microsatellites (MSs) are repeating sequences of 2–6 base pairs of DNA (1). They are used as molecular markers in genetics for kinship, population, and other studies. They can also be used for the studies of gene duplication or deletion, marker-assisted selection, and fingerprinting. Microsatellites are distributed throughout the genome (1). Being variable genetic elements, microsatellites provide a potent tool for the individual characterization of genomes. Variability is generated due to replication slippage caused by mismatches between DNA strands while being replicated during meiosis (2), and the event can occur once per 1000 generations (3). This slippage is much more common compared with point mutations (4). Microsatellite repeats mutagenize human genomes and alter the human genomic landscape across generations (5). The utility of microsatellites has been demonstrated by the study comprising 2058 germline changes discovered by analyzing 85 289 Icelanders at 2477 microsatellites. The paternal-to-maternal mutation rate ratio is 3.3, and the rate in fathers doubles from the age 20 to 58, whereas there is no association with age in mothers. Longer microsatellite alleles are more mutagenic and tend to decrease in length, whereas the opposite is seen for shorter alleles (6).

Microsatellites remain highly informative and useful measures of genomic variation for linkage and association studies despite the fact that general preference is given to single-nucleotide polymorphisms (SNPs). Microsatellites are much more genetically diverse compared to SNPs; they generate a greater haplotype diversity (7).

Microsatellites are known to form nuclear matrix anchorage sites (8), tissue-specific matrix attachment sites (9), and binding sites with vimentin and glial fibrillary acidic protein (10) and give rise to complex DNA spatial structures of extreme functional significance (10–12). MSs appear to be important components of insulators (13), silencers (12), and enhancers (14). MSs are also involved in the regulation of alternative splicing (15), mRNA stability (16), and recombination and repair (17, 18).

Expansions of microsatellite DNA repeats cause nearly 30 developmental and neurological inherited disorders (19). Further, we shall try to summarize and classify the data about the association between MSs and human diseases in connection to the localization of repeats in the genes and their possible functional role. Triplet expansion diseases will be excluded from our analysis, as the problem has been extensively reviewed (20, 21).

Promoter Microsatellites. MSs can determine the activity of the upstream gene regulation elements like the locus control region of the beta-globin gene domain. The (AT)(8)N(12)GT(AT)(7) configuration of a microsatellite found in the hypersensitive site of the structure is associated with a special form of sickle cells, Tunisian βs chromosomes (22). The most common allele of the MS marker in STAT4, the STAT4-MS1-254 allele, located in the 5' flanking region of the gene, is significantly associated with sarcoidosis (23). Changes in the length of microsatellites within promoters and other cis-regulatory regions can also change the level of gene expression, and they are linked to abundant variations in cis-regulatory control regions in the human genome (24). For example, a CA-repeat mi-
Microsatellite in the insulin-like growth factor 1 promoter is associated with the level of this growth factor. It has turned out that the intensity of the gene transcription is regulated by the interaction of several SNPs and microsatellite-generating haplotypes with lower or higher levels of the gene transcription (25). Promoter microsatellites tend to be guanine-cytosine rich; they are often found at the start of genes and are probably associated with the regulatory elements such as CpG islands, G-quadruplexes (G4), and untranslated regulatory regions. Numerous promoter microsatellites possess the potential to influence human phenotypes by generating mutations in regulatory elements, which may ultimately lead to a disease (26). A CpG-CA repeat within the human endothelin-converting enzyme 1 promoter is highly polymorphic, harbors transcriptional start sites, is able to recruit the transcription factors and poly(ADP-ribose) polymerase-1 and splicing factors, and is functional regarding haplotype-specific promoter activity. The overall CpG-CA repeat composition of patients with Alzheimer's disease and nondemented control individuals has been found to be distinct (27). A length polymorphism of GT repeats in the promoter region of the human heme oxygenase-1 (HO-1) gene modulates the transcription of this gene (28). Numerous studies have linked human HO-1 gene promoter polymorphisms to a risk of vascular diseases (29). Persons carrying longer (GT)(n) repeats in the HMOX1 gene (L allele) promoter may be at a higher risk of type 2 diabetes mellitus (30). Functional analyses have shown that the persons with impaired glucose regulation type 2 diabetes mellitus, carrying the L/L (GT) (n) genotype, have significantly lower HO-1 protein expression levels than those with the S/S genotype (31). The same microsatellite is associated with susceptibility to cardiovascular complications of the disease. Patients with longer lengths of GT repeats in the heme oxygenase-1 gene promoter exhibit higher inflammation and oxidative stress. These patients have a higher risk of long-term cardiovascular events and mortality (32). A short allele of the same microsatellite might be associated with an abdominal aortic aneurysm (33). Long (GT)n repeats in the microsatellite polymorphism region of the HMOX1 gene have been reported to be associated with symptomatic malaria (34).

The aldose reductase (AKR1B1) gene promoter harbors a (CA)n microsatellite significantly associated with diabetic retinopathy. The z-2 microsatellite has been found to confer risk in type 1 and type 2 diabetes mellitus and the z-2 microsatellite to confer protection against diabetic retinopathy in type 2 diabetes mellitus regardless of ethnicity (35, 36). The S allele of a ((CCTTT)(n) repeat in the promoter of the NOS2 gene is associated with both hypertension and responsiveness to antihypertensive drug therapy (37). The same microsatellite is also associated with diabetic retinopathy, as well as the (GT)n promoter repeat in the tumor necrosis factor β gene (36).

Promoter MSs might also be associated with mental problems; the promoter TA microsatellite repeat in the estrogen receptor alpha gene is significantly associated with postpartum depression (38). The arginine vasopressin receptor 1A gene (AVPR1A) is widely expressed in the brain and is considered to be a key receptor in the regulation of social behavior. 5'-Flanking region polymorphisms in the human AVPR1A, RS3, and RS1 show differences in relative promoter activity by length. Shorter repeat alleles of RS1 and RS3 have decreased relative promoter activity in the human neuroblastoma cell line SH-SY5Y. The short alleles of RS1 are associated with autism (39).

A recent meta-analysis by Shen et al. has reported that CYP11A1 promoter microsatellite [TTTA]n repeat polymorphisms may contribute to increasing susceptibility to the risk of polycystic ovary syndrome (40).

Signal Sequence Microsatellites. Some microsatellites are localized in translated areas of the genes. For example, the carnosinase gene contains a D18S880 microsatellite formed of a leucine triplet repeat in its signal sequence. Homozygotes for 5 trinucleotide repeats in this microsatellite are susceptible to diabetic nephropathy (41). The human signal transducer and activator of transcription 6 (STAT6) gene represents one of the most promising candidate genes for asthma and other inflammatory diseases on the chromosomal region 12q13-q24. The gene exon 1 contains a GT repeat upstream the first methionine codon. Allele A4 of the GT repeat polymorphism is associated with an increase in the eosinophil cell count (42). The genotype of 13/15-GT repeat allele heterozygosity is significantly associated with allergic subjects (43). Microsatellites of Coding Regions. Besides trinucleotide expansion diseases, characterized mostly by polyglutamine tracts (poly-Q), which cannot be analyzed here due to space limitations, an interesting trinucleotide repeat has been identified in the MIC-A gene. The exon 5 microsatellite polymorphism of the MIC-A gene consists of 5 alleles based on the number of GCT triplet repeat units (alleles A4, A5, A6, and A9) and the presence of an additional nucleotide insertion (allele A5.1). CGT repeats regulate the number of Ala residues in the protein, and the A5.1 leads to a frameshift mutation. The exon encodes the membrane-binding domain of the protein (44). The microsatellite alleles are associated with Addison’s disease (44) and type 1 diabetes mellitus (45–47). Some alleles are protective against...
juvenile idiopathic arthritis (48). Variations of CAG (Gln) repeats in the androgen receptor gene in physiological limits, not causing insensitivity to androgens, can influence certain physiological parameters. Shorter androgen receptor (AR) CAG is associated with low HDL-C and testosterone levels (49).

Intrinsic Microsatellites. Microsatellites within introns also influence a phenotype, through ways that are not currently understood; this is the cause of numerous associations of microsatellite repeat polymorphisms with human diseases. For example, a GAA triplet expansion in the first intron of the X25 gene appears to interfere with transcription and causes Friedreich ataxia (50). Subjects having more CA repeats in the first intron of the type 2 11β-hydroxysteroid dehydrogenase gene (HSD11B2) are susceptible to developing abnormal glucose tolerance (51). A repeat polymorphism in the fourth intron of the NOS3 gene is linked to hypertension (52). We have detected an association between type 2 diabetes mellitus and microsatellite markers of the region 14q13 localized in the introns of the PSMA6 and KIAA0391 genes, rs63749745, markers of the region 14q13 localized in the introns between type 2 diabetes mellitus and microsatellite repeat polymorphisms with human diseases. One allele of the D6S1276 microsatellite in intron 1 of the BMP5 gene is associated with the risk of osteoarthritis in women; 2 alleles are protective (63).

3'-UTR Microsatellites. Microsatellites localized in the 3'-untranslated regions (3'-UTRs) may affect the final mRNA stability, the localization, the export from the nucleus and the translation efficiency. The androgen receptor CAG repeat polymorphism (AR CAG) affects receptor transcriptional activity (the shorter repeats, the more sensitive AR) and is associated with androgenic parameters and obesity (49). For the conserved regulatory sequences within the 3'-UTRs and the specific elements binding to them enable gene expression control at the posttranscriptional level, and all these processes reflect the actual state of the cell (64). Shorter alleles of the microsatellites in the 3' flanking region of the leptin gene, coding for a protein hormone, mainly synthesized in adipocytes, which regulates the food intake and energy expenditure of the body, are significantly associated with hypertension (65). Reduced repeat lengths in the EGFR gene 3'-UTR polyA repeat are linked with osteosarcomas (66). Several alleles of the microsatellite (AT)n in the 3'-UTR of anticytotoxic T lymphocyte antigen-4 (CTLA-4) gene, namely 104-, 106-, 110-, and 116-bp alleles, were observed to be predisposing to recurrent miscarriage (67).

Remote and Locus-Specific Microsatellites. In some cases, the association with diseases is found for microsatellites localized far from the candidate genes. Marker D12S96 is localized 5,653 cM downstream the vitamin D receptor (VDR) gene. Despite this long distance and obscure functional relations, statistically significant linkage disequilibrium has been detected between allele 22 of locus D12S96 and osteoporosis (68). In some cases, the association between microsatellites and candidate genes is not traced at all as these are sooner the locus than gene markers. The 8p21-23 region microsatellites D8S136 and D8S520 are consistently and strongly related with prostate cancer (69). The locus has been traced due to a frequent loss of heterozygocity in tumors, but not as a result of association studies. The D1S2726 microsatellite, located 30 kb from the KCNA3 gene, which encodes the voltage-gated potassium channel Kv1.3, is associated with susceptibility to autoimmune pancreatitis (70).
Conclusions
The above data clearly indicate that most microsatellites manifesting the association with human pathologies are harbored in genes encoding enzymes involved in pathogenesis of the pathologies; in many cases, the impact of the changes in the microsatellite length on the gene function can be evaluated. Thus, further studies on microsatellite associations with diseases remain prospective, despite numerous whole-genome association studies. Contribution of studies with individual polymorphisms to the understanding of genetic background of diseases should not be underestimated.

References
1. King DG. Evolution of simple sequence repeats as mutable sites. Adv Exp Med Biol 2012;769:10-25.
2. Tautz D, Schlötterer C. Simple sequences. Curr Opin Genet Dev 1994;4:832-7.
3. Weber JL, Wong C. Mutation of human short tandem repeats. Hum Mol Genet 1993;2:1123-8.
4. Jarne P, Lagoda PJ. Microsatellites, from molecules to populations and back. Trends Ecol Evol 1996;11:424-9.
5. Grandi FC, An W. Non-LTR retrotransposons and microsatellites: partners in genomic variation. Mob Genet Elements 2013;3:e25674.
6. Sun JX, Helgason A, Masson G, Ehrenesdottir SS, Li H, Mallick S, et al. A direct characterization of human mutation based on microsatellites. Nat Genet 2012;44:1161-5.
7. Gulcher J. Microsatellite markers for linkage and association studies. Cold Spring Harb Protoc 2012;2012:425-32.
8. Boulíkas T. Nature of DNA sequences at the attachment regions for the nuclear matrix. J Cell Biochem 1993;52:14-22.
9. Lenartowski R, Goc A. Tissue-specific association of the human tyrosine hydroxylase gene with the nuclear matrix. Neurosci Lett 2002;330:151-4.
10. Tolstogov GV, Li G, Shoeman RL, Traub P. Interaction in vitro of type III intermediate filament proteins with higher order structures of single-stranded DNA, particularly with G-quadruplex DNA. DNA Cell Biol 2005;24:85-110.
11. Li G, Tolstogov GV, Traub P. Interaction in vitro of type III intermediate filament proteins with Z-DNA and B-Z-DNA junctions. DNA Cell Biol 2003;22:141-69.
12. Rothenburg S, Koch-Nolte F, Haag F. DNA methylation and Z-DNA formation as mediators of quantitative differences in the expression of alleles. Immunol Rev 2001;184:286-96.
13. Filiippova GN, Thiesen CP, Penn BH, Cho DH, Hu YJ, Moore JM, et al. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. Nat Genet 2001;28:353-45.
14. Bassuny WM, Ibara K, Sasaki Y, Kuromaru R, Kohn H, Matsuura, N, et al. A functional polymorphism in the promoter/enhancer region of the FOXP3/Scurf1 gene associated with type 1 diabetes. Immunogenetics 2003;55:149-56.
15. Hui J, Hung LH, Heiner M, Schreiner S, Neumüller N, Reither G, et al. Intronic CA-repeat and CA-rich elements: a new class of regulators of mammalian alternative splicing. EMBO J 2005;24:1988-98.
16. Lee JH, Jeon MH, Seo YJ, Lee YJ, Ko JH, Tsujimoto Y, et al. CA repeats in the 3′-untranslated region of bcl-2 mRNA mediate constitutive decay of bcl-2 mRNA. J Biol Chem 2004;279:42758-64.
17. Wang G, Vasquez KM. Z-DNA, an active element in the genome. Front Biosci 2007;12:4424-38.
18. Chin YJ, Schleifman EB, Glazer PM. Repair and recombination induced by triple helix DNA. Front Biosci 2007;12:4288-97.
19. Kim HC, Min JH. The balancing act of DNA repeat expansions. Curr Opin Genet Dev 2013;23:280-8.
20. Wang YH. Chromatin structure of repeating CTG/CAG and CGG/CCG sequences in human disease. Front Biosci 2007;12:4731-41.
21. Fondon JW, Hammad EA, Hanan AJ, King DG. Simple sequence repeats genetic modulators of brain function and behavior. Trends Neurosci 2008;31:328-34.
22. Ben Mustapha M, Mounini I, Zorai A, Douzi K, Ghanem A, Abbès S. Microsatellite and single nucleotide polymorphisms in the β-globin locus control region-hypersensitive Site 2: SPECIFICITY of Tunisian βs chromosomes. Hemoglobin 2012;36:533-44.
23. Tanaka G, Matsushita I, Ohashi J, Tsuchiya N, Ikushima S, Oritsu M, et al. Evaluation of microsatellite markers in association studies: a search for an immune-related susceptibility gene in sarcoidosis. Immunogenetics 2005;56:861-70.
24. Rockman MV, Wray GA. Abundant raw material for cis-regulatory evolution in humans. Mol Biol Evol 2002;19:1991-2004.
25. Chen HY, Huang W, Leung VH, Fung SL, Ma SL, Jiang H, et al. Functional interaction between SNPs and microsatellite in the transcriptional regulatory region of insulin-like growth factor 1. Hum Mutat 2013;34:1289-97.
26. Sawaya S, Bagshaw A, Buschiazzo E, Kumar P, Chowdhury S, Black MA, et al. Microsatellite tandem repeats are abundant in human promoters and are associated with regulatory elements. PLoS One 2013;8:e54710.
27. Li Y, Seidel K, Marschall P, Klein M, Hope A, Schachtj J, et al. A polymorphic microsatellite repeat within the ECE-1c promoter is involved in transcriptional start site determination, human evolution, and Alzheimer’s disease. J Neurosci 2012;32:16807-20.
28. Hu YF, Lee KT, Wang HH, Ueng KC, Yeh HI, Chao TF, et al. The association between heme oxygenase-1 gene promoter polymorphism and the outcomes of catheter ablation of atrial fibrillation. PLoS One 2013;8:e56440.
29. Wu ML, Ho YC, Yet SE. A central role of heme oxygenase-1 in cardiovascular protection. Antioxid Redox Signal 2011;15:1835-46.
30. Bao W, Song F, Li X, Rong S, Yang W, Wang D, et al. Association between heme oxygenase-1 gene promoter polymorphisms and type 2 diabetes mellitus: a HuGE review and meta-analysis. Am J Epidemiol 2010;172:631-6.
31. Song F, Li X, Zhang M, Yao P, Yang N, Sun X, et al. Association between heme oxygenase-1 gene promoter polymorphisms and type 2 diabetes in a Chinese population. Am J Epidemiol 2009;170:747-56.
32. Chen YH, Hung SC, Tarag DC. Length polymorphism in heme oxygenase-1 and cardiovascular events and mortality in hemodialysis patients. Clin J Am Soc Nephrol 2013;8:1756-63.
33. Gregorek AC, Goralk KC, Polance DS, Dalbelic S. GT microsatellite repeats in the heme oxygenase-1 gene promoter associated with abdominal aortic aneurysm in Croatian patients. Biochem Genet 2013;51:482-92.
34. Mendoza VR, Luz NF, Santos NJ, Borges VM, Gonçalves MS, Andrade BB, et al. Association between the haptoglobin and heme oxygenase 1 genetic profiles and soluble CD163 in susceptibility to and severity of human malaria. Infect Immun 2012;80:1445-54.
35. Abhay S, Hewitt AW, Burdon KP, Craig JE. A systematic meta-analysis of genetic association studies for diabetic retinopathy. Diabetes 2009;58:2137-47.
36. Uthra S, Raman R, Mukesh BN, Rajkumar SA, Kumari F, Lakshmi P, et al. Diabetic retinopathy: validation study of ALR2, RAGE, INOS and TNFB gene variants in a south Indian cohort. Ophthalmic Genet 2010;31:244-51.
37. Oliveira-Paula GH, Lacchini R, Coeli-Lacchini FB, Junior HM, Tanus-Santos JE. Inducible nitric oxide synthase haplotype associated with hypertension and responsiveness to antihypertensive drug therapy. Gene 2013;515:391-5.
38. Pinsonneault JK, Sullivant D, Sades W, Soares CN, Hampson E, Steiner M. Association study of the estrogen receptor gene ESR1 with postpartum depression – a pilot study. Arch Womens Ment Health 2013;16:499-509.
39. Tansey KE, Hill MJ, Cochrane LE, Gill M, Anney RJ, Galagher L. Functionality of promoter microsatellites of arginine vasopressin receptor 1A (AVPR1A): implications for autism. Mol Autism 2011;2:3.
40. Shen W, Li T, Hu Y, Liu H, Song M. Common polymorphism of the 33 bp repeat polymorphism in the 3'UTR of endothelial constitutive nitric oxide synthase gene in type 2 diabetes mellitus in Latvian and Finnish populations. Prostate Cancer Prostatic Dis 2007;10:71-7.
41. Cao F, Zhang H, Feng J, Gao C, Li S. Association study of three microsatellite polymorphisms located in introns 1, 8, and 9 of DISC1 with schizophrenia in the Chinese Han population. Genet Test Mol Biomarkers 2013;17:407-11.
42. Zakhei A, Simin H, Foroussan S, Manoochehr T. Polymorphic CT dinucleotide repeat in the GATA3 gene and risk of breast cancer in Iranian women. Med Oncol 2013;30:504.
43. Zheng Y, Huo D, Zhang J, Yoshimatsu TF, Niu Q, Olopadie OI. Microsatellites in the estrogen receptor (ESR1, ESR2) and androgen receptor (AR) genes and breast cancer risk in African American and Nigerian women. PLoS One 2012;7:e40494.
44. Xie GB, Xu P, Che YN, Xia YJ, Cao YX, Wang WJ, et al. Microsatellite polymorphism in the fibrillin 3 gene and susceptibility to PCOS: a case-control study and meta-analysis. Reprod Biomed Online 2013;26:168-74.
45. Belguith-Maalej S, Kallel R, Minif M, Abid M, Ayadi H, Kacem HH. Association of intronic repetition of SLC26A4 gene with Hashimoto thyroiditis disease. Genet Res (Camb) 2013;95:38-44.
46. Sjäkte T, Eglitte J, Sochnevs A, Marga M, Piragas V, Collan Y, et al. Microsatellite genotyping of chromosome 14q13.2-14q13 in the vicinity of proteasomal gene PSA16 and association with Graves’ disease in the Latvian population. Immuno genetics 2004;56:238-43.
47. Sjäkte T, Triapina I, Rumba-Rozenfelde I, Lunin R, Sugoka O, Sjäkte N. Identification of a novel candidate locus for juvenile idiopathic arthritis at 14q13.2 in the Latvian population by association analysis with microsatellite markers. DNA Cell Biol 2010;29:543-51.
48. Silva GA, Santos MP, Mota-Passos I, Boechat AL, Malheiro A, Naveca FG, et al. IFN-γ +875 microsatellite polymorphism as a potential protection marker for leprosy patients from Amazonas state, Brazil. Cytokine 2012;60:493-7.
49. Hijikata M, Shoijima J, Matsushita I, Tokunaga K, Ohashi J, Hang NT, et al. Association of IFNGR2 gene polymorphisms with pulmonary tuberculosis among the Vietnamese. Hum Genet 2012;131:675-82.
50. Rodriguez-Fontenla C, Carr A, Gomez-Reino JJ, Tsezou A, Loughlin J, Goelz I. Association of a BMP5 microsatellite with knee osteoarthritis: case-control study. Arthritis Res Ther 2012;14:R257.
51. Michalova E, Vojtesek B, Hrastka R. Impaired pre-mRNA processing and altered architecture of 3' untranslated regions contribute to the development of human disorders. Int J Mol Sci 2013;14:15681-94.
52. Akther Q, Masood A, Ashraf R, Majid S, Rasool S, Khan T, et al. Polymorphisms in the 3'UTR of the human leptin gene and their role in hypertension. Mol Med Rep 2012;5:1058-62.
53. Kersting C, Ageolopoulos K, Schmidt H, Korschning E, August G, Goshgeher G, et al. Biological importance of a polymorphic CA sequence within intron 1 of the epidermal growth factor receptor gene (EGFR) in high grade central osseousomas. Genes Chromosomes Cancer 2008;47:657-64.
54. Gupta R, Prakash S, Parvesh P, Agrawal S. Association of CTLA-4 and TNF-α polymorphism with recurrent miscarriage among North Indian women. Cytokine 2012;60:456-62.
55. Raje M, Botre C, Ashra R. Genetic epidemiology of osteoporosis across four microsatellite markers near the VDR gene. Int J Mol Epidemiol Genet 2013;4:101-8.
56. Zeegers MP, Nekeman D, Khan HS, van Dijk BA, Goldbohm RA, Schalken J, et al. Prostate cancer susceptibility genotypes on 8p21-23 in a Dutch population. Prostate Cancer Prostatic Dis 2012;15:248-53.
57. Ota M, Ito T, Umemura T, Katsuyama Y, Yoshizawa K, Hamano H, et al. Polymorphism in the KCNA3 gene is associated with susceptibility to autoimmune pancreatitis in the Japanese population. Dis Markers 2011;31:223-9.