We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter 1

Introduction to Microsatellites: Basics, Trends and Highlights

Ibrokhim Y. Abdurakhmonov

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/66446

1. Introduction

A genome is written in four chemical letters (nucleotides designated as A, T, G, and C). Various combinations of these letters as a stretch of DNA molecule provide specificity and uniqueness of each gene sequence, cell types, and each individual genotype. Magic is that the order of each triplet (known as genetic code) of these four-letter DNA sequence stretches corresponds to one of 20 amino acids. A "spelling" order of nucleotides encodes the specific protein sequences determining a life function. Therefore, variations in these four-letter DNA sequences in a genome make the meaning and differentiation of the living organisms on Earth that generated biological diversity. Because of degeneracy of genetic code, some spelling changes in coding parts of a sequence (exons) may have no meaning and still code the same amino acid although changes may have evolutionary role and contribute to the biodiversity levels of populations. However, other changes may lead to generate novel proteins with new function and characteristics, stop the gene function and protein synthesis, or generate a partial protein sequences that is not sufficient for its functional activity—all these alter the function of the cell and generate a difference.

The pattern of DNA variations among individuals, generating beneficial and diseased phenotypes, is very useful to differentiate organisms in molecular level and to understand the evolutionary path of important genes as well as their functional and adaptive roles in the different eco-geographic environments. One of such abundant spelling variations in a genome is the 5- to 50-fold repetitions of the two to six nucleotide base pair (bp) motifs of DNA, such as $(GA)_n$ or $(GTACGT)_n$, which are called as microsatellites [1–4]. These tandem repeats are referred to as microsatellites, simple sequence length polymorphisms (SSLP), simple sequence repeats (SSR), or short tandem repeats (STR), which are used interchangeably among researchers. Microsatellites are abundantly found in all prokaryote and eukaryote genomes.

INTECH
open science | open minds

© 2016 The Author(s). Licensee InTech. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Litt and Luty [4] first coined the term "microsatellite" in 1989, where the word "satellite" was used due to fact that density gradient centrifugation separates DNA fragments with repetitive sequences into the upper "satellite" fraction with less density. As a genetic marker, microsatellites have been widely used in DNA-based genetic analyses for the past 25 years. Since the first paper by Litt and Luty [4] in 1989, as of October 2016, *Pubmed* [5] database search with the quoted keyword "microsatellite" found almost 44,000 research publications that have used or discussed microsatellites (**Figure 1**). Hodel et al. [1] reported that as of April 2016, they have found almost 225,000 published articles by searching Web of Science (WOS) database. Searching the WOS core collection for plant science-related articles, Vieira et al. [2] reported that for the past 5-year period from 2010 to 2015, there were 993 unique crop-related publications using microsatellites that demonstrate a wide utilization of SSRs in plant sciences. In this introductory chapter, I aimed to give an overview of definition, distribution, utility, and future of microsatellites, briefly highlighting chapter contents of this book.

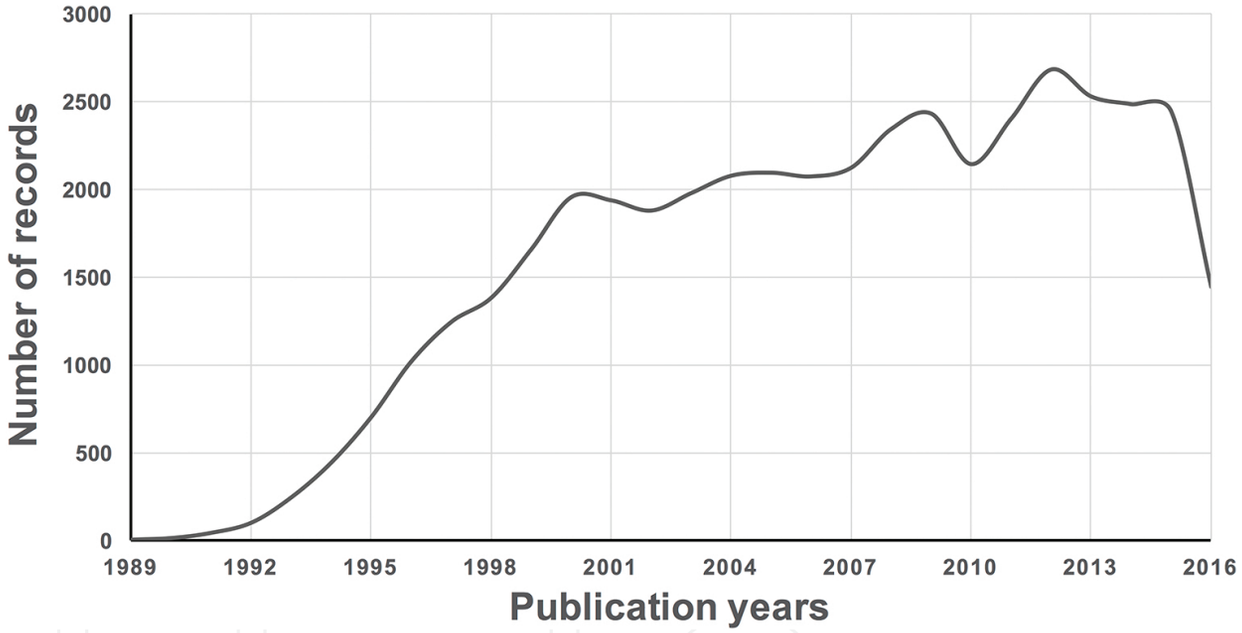

Figure 1. Number of publications retrieved with "microsatellite" keyword from PubMed [5].

2. Definition, occurrence, types, distribution, and density

A variable number tandem repeats (VNTRs) include both mini- and microsatellite DNAs. Minisatellites are the heterogeneous array of 10–60/100 core bp repeat motif sequences such as $(GGGCAGGTNG)_n$ that have repeat size of 1–15 kilobases (kb). In contrast, microsatellites, commonly, consist of a homogeneous array of core mono, di, tri, tetra, penta, and hexanucleotide motifs with repeat size of less than or around 1 kb. Controversially, some reports include all repeat arrays less than 9 bp into microsatellite category and those above nine core repeats into minisatellite group [1–3, 6]. However, the suggested core array of repeat motifs for

microsatellites are 2–6 bp, which non-randomly distributed throughout the genome [7] and vary largely in different regions of a genome or on different taxa [2]. Microsatellites can be found abundantly in non-coding parts of the genome such as introns, untranslated regions (UTR), and intergenic spaces, but they also occur in coding exonic sequences. Microsatellites also located within transposons and other dispersed repetitive elements [1–3, 6, 7].

Density of microsatellites considered to be highest in UTRs and in decreasing order in promoters, introns, intergenic regions, and coding sequences. Microsatellite repeat lengths in coding, non-coding and intergenic regions are reported to be species specific. For instance, generally, vertebrates (e.g., turtles) tend to have more and longer array microsatellites compared to plants and then invertebrates [6]. Although vary organism to organism and may not be true for all genomes under consideration, commonly trinucleotide motifs are more frequent than other types, being highest in plants to 61–73%, except Arabidopsis and potato (>30%) [6]. The distribution of di-nucleotide repeat microsatellites is higher in mammals (31%) and rodents (33%) than plants (in average ~24%) although Arabidopsis and potato have >30 and 50% dinucleotide repeats [6]. Plants have less frequency of GA dinucleotide repeat SSRs, while animal genomes abundantly contain these types of repeats. GC dinucleotide arrays are less frequent in coding sequences, but GC-rich trinucleotide arrays frequently occurred in exons, while AT-rich trinucleotides evenly distributed throughout all genomic regions. For example, ATT repeats are abundant in introns of genes of most organisms although some genomes like rodents tend to have AAG abundance in introns. Generally, ACG, ACC, ACT repeat SSRs are rare in all organisms [6].

Interestingly, there is a predominance of tri- and hexanucleotides in coding regions that is explained as a result of selection forces to keep reading frame not altered. However, microsatellites of such triplet expansions can cause harmful phenotypes such as Fragile X syndrome (FXS) and Huntington's disease [7]. Tetranucleotide motif SSRs located predominantly in noncoding regions with abundance of AT-rich motifs. Mammals have more tetranucleotide microsatellites compared to other organisms [6]. Similarly, pentanucleotide motif microsatellites abundantly represented in intronic regions of animal genomes compared to bacteria and plants.

According to a repeat motif pattern, microsatellites can be classified as (1) perfect with continuous repeat of single motif, (2) imperfect with a base pair disruption between repeats, (3) interrupted with insertion of a stretch of sequence of few nucleotides within repeats, or (4) composite with multiple SRR motif repeat types that vary among different taxa [2]. The density of SSRs also vary among different taxa and occurred one SSR in about 6.04 kb for Arabidopsis, 6–7 kb in mammals and could be less than 1 kb in puffer fish [2] or up to 212–292 kb in hexaploid wheat [8]. Microsatellites can be genomic, i.e., developed from genomic DNAs (gSSRs) or can be expressed, referred to as EST-SSRs, derived from expressed sequence tags (ESTs) [2, 9]. EST-SSRs have high power because of their associations with expressed genes, directly contributing to a phenotype [10]. In plants, SSRs can also be classified as nuclear SSRs if they occurred in nuclear DNA (nuSSRs) and chloroplast SSRs (cpSSRs), if they occurred in chloroplast DNA. cpSSR loci were first introduced by Powel et al. [11, 12] in 1995 as useful genetic markers with

broad applications in plants, in particular for measuring the cytoplasmic diversity and introgression in plant species, although there are some limitations highlighted below [12].

3. Origin, evolution, and mutation mechanisms and rates

Due to sequence mutation, the genesis of microsatellite locus can be started from a mutated site with minimum of eight nucleotide repeats or from de novo points without repeat motifs leading to formation of "proto-microsatellites/SSRs" sites. Some reports discuss the potential of minisatellites as a progenitor of SSRs, while others suggest the contribution of transposons to the birth of SSRs although it is not evident in birds and plants [6]. Following the next generation of DNA replication process, "proto-SSR" sites can get expanded due to errors caused by DNA polymerase strand slippage [2, 3, 6, 7]. Moreover, based on "transposon-mediated" microsatellite birth, SSRs can be born due to transposon movement, exemplified by the origin of *Alu* element-derived AT-rich SSRs [3, 6, 7].

Replication slippage is considered the main mechanism [2, 3, 6, 7] of microsatellite genesis, repeat expansion or reduction, and generating a variability that all are contributing to the molecular evolution of microsatellites. Besides replication-associated slippage, the birth of microsatellites can occur during transcription-coupled DNA repair and/or repair of double-stranded breaks where repetitive sequences are preferably used for filling the gaps [6]. Further, insertion and deletions (indels) or single nucleotide substitutions, which occur in increased rates [13, 14], can generate new repeat arrays in microsatellites. For instance, a comparative human and primate genome analysis revealed that the repeat number change in short microsatellites mostly occurs because of single nucleotide polymorphism (SNP) mutations rather than slippage [15].

Microsatellite mutations can simultaneously change one, two, or more repeat unit(s), providing a higher mutation rates of 10^{-2} to 10^{-6} per microsatellite locus per generation [6] than other mutation types, such as point mutation rates, which is approximately 10^{-9} nucleotides per generation for entire genome in eukaryotes [1]. The base position relative to the microsatellite, genomic location, repeat type and number, base identity, flanking sequence, speed of recombination and transcription, and heterozygosity of microsatellite alleles greatly affect the microsatellite mutation rates [6, 16]. In particular, microsatellites in noncoding regions tend to mutate frequently than those in coding regions, and/or changes in perfect repeats can generate new SSRs [6]. In addition, recombination with unequal-length SSR alleles can increasingly cause SSR instability during meiosis [15, 16]; dinucleotide repeats mutate frequently than tri- and tetra-nucleotide arrays, and/or longer and purer repeats can mutate in high rate than shorter repeats with low purity [2]. SSR mutation rate is species and gender specific where human males have higher SSR mutation rate ($0–7 \times 10^{-3}$ per locus per gamete per generation) compared to females [17].

The rate of both repeat motif expansion and contraction is also species specific. For example, repeat expansion mutations are faster in humans than chimpanzees, or there is a loss of two repeat units per mutation in yeast compared to a loss of 1.4 repeats in Drosophila [6]. Repeat

expansion mutations predominantly found in primates, while bacteria have repeat contractions. Longer repeat arrays in SSRs are considered to be recent origin and long repeat SSRs are biased toward repeat expansion mutation. Statistically, the patterns of variation in SSR loci can be studied and predicted using "Stepwise Mutation Model" (SMM) or its two-step modification and "Infinite Allele Model" (IAM) [6].

4. Biological function

Due to common understanding that repetitive DNAs are "junk" and nonfunctional, tandem repeat microsatellites have been considered as neutral elements in a genome without distinct biological function although there were numerous observations that microsatellite mutations can lead to many diseased phenotypes and change the function of proteins. The occurrence of microsatellites in coding and regulatory gene regions (as well as introns or in intergene regions) supported the biological function of microsatellites in such processes as (1) gene expression including transcription and translation, (2) gene silencing, (3) alternative splicing and mRNA transport, (4) chromatin organization, and (5) regulation of cell cycle [2, 6]. Involvement of microsatellite repeat motives in these key biological processes of cell life not only leads to the cell phenotype change and cause disease and unwanted traits but also determines the evolutionary fate, survival, plasticity, and adaptation of organisms in changing and potentially harmful environments [2, 3, 6]. Discovery of the co-localization of SSR with pre-microRNAs and influence of CUU repeat numbers to the loop size of pri-microRNAs in orange plants [6, 18] or involvement of certain r(CGG)-derived microRNAs such as miR-fmr1s in FXS-pathogenesis demonstrated a possible role of microsatellites in many developmental processes regulated by microRNAs [19].

There are many examples for distinctive phenotypic changes that directly associated with the increases or decreases of microsatellite repeat arrays. For instance, more than 40 neurological diseases in humans, such as FXS and spinocerebellar ataxia (SCA1) with a polyglutamine tracts, are caused by microsatellite motif length changes in trinucleotide arrays [20]. Microsatellite repeat changes determine morphological features, for example, repeat expansion of microsatellite stretches in the *aristaless-like 4 (ALX-4)* and *runt-related transcription factor 2 (RUNX2)* genes of domesticated dogs (*Canis familiaris*) is associated with limb and skull morphology [21] with interesting correlation between longer sequence lengths of *RUNX2* microsatellites and longer faces of dogs, which is observed in 30 naturally evolving Carnivora species [22]. Microsatellite repeat polymorphism in control regions of the *Vasopressin 1a receptor* gene affects social behavior, level of monogamy [23], and autism and socialization skills [24] in humans, and the courtship behaviors in other mammals [25]. Repeat number changes in microsatellites control the duration of its circadian clock cycles in a fungus *Neurospora crassa* [26]. SSR expansions in noncoding regions also generate diseased phenotypes. For instance, Friedreich Ataxia is caused by a GAA triplet expansion in the first intron of the *X25* gene that is explained by its influence with transcription [27]. Repeat number expansions/reductions in introns of several genes such as *Asparagine synthetase, NOS3*, and *EGRF* genes cause acute lymphoblastic leukemia [28], hypertension [29], and osteosarcomas [30], respectively.

5. Utility of microsatellites as genetic markers

As a genetic marker, microsatellites can be widely applied for solving a numerous type of different tasks. These include the construction of genetic linkage groups and integrated maps; correlation of phenotypic and genotypic variations using quantitative trait locus (QTL) and/or linkage disequilibrium (LD)-based association mapping approaches; analyses of parentage and/or ancestry; DNA barcoding for plant varieties and germplasm; evaluation of gene flow and variety/seed purity; breeding using marker-assisted selection tools; estimation of genetic diversity, phylogeography, conservation and restoration of biodiversity, molecular evolution, taxonomy, and phylogenetic features of biological species; detection of genetic structure of native plant populations and crop germplasm, origin and domestication of crop species, migration, demographic process, population differentiation and kinship; assessment of impacts of mutagenic contaminants; and application in forensics and disease diagnostics [1, 2, 31].

5.1. Marker development

Microsatellites are polymerase chain reaction (PCR)-based markers and require a prior knowledge on sequence structure before using them as a genetic marker. There are two ways to develop SSR markers: (1) necessary genome or its part should be sequenced following screening for microsatellite repeat arrays; or (2) preliminary sequenced genomes databases can be mined using variety of *in silico* bioinformatics software packages. As further steps, both approaches, however, require designing and synthesis of marker primers, genotyping, scoring, and assessing the polymorphism levels in diverse genotypes under study using PCR to apply for a specific genetic study. There are various methods and approaches [1, 2] available for construction/enrichment (e.g., selective hybridization or biotin-captured) and sequencing [e.g., Sanger or next generation sequencing (NGS)-based] of genomic libraries for SSRs as well as genotyping (e.g., agarose, polyacrylamide (PAG), and capillary electrophoresis with fluorescent detection), which we skip the details here. These approaches have been historically well optimized and used depending on the purpose of study, expertise, availability of necessary equipment and reagents, and targeted types of SSR arrays.

When sequences are generated de novo or available as genome databases in the National Center of Biotechnology Information NCBI [32], most important step is to efficiently screen microsatellite containing sequences and design markers. For this purpose, there are many SSR array searching algorithms available such as tandem repeat finder (TRF), MIcroSAtellite identification tool (MISA), SSRFinder, and PALFinder [1, 2]. Besides there are several web servers based online tools such as CID [33] and WebSat [34]. Each of these bioinformatics tools has its advantages and disadvantages, can address various aspects of microsatellite mining and marker development and be used according to study/task objectives, expertise and availability. There is some recommended software for efficient screening microsatellite repeats from DNA sequences such as MISA or Phobos [1]. Further, there is a list of many other useful bioinformatics resources for the genetic analyses of microsatellite data [35].

5.2. Advantages

Among all other type of molecular markers, for past three decades, microsatellite markers were the marker of choice because they are PCR based; abundant and dispersed throughout a genome; highly mutagenic, polymorphic, and informative; co-dominant, suitable for detecting heterozygotes, and multi-allelic; experimentally reproducible; transferable among related taxa; cost-effective and easy to detect; amplified from low quality and low quantity of DNAs; and presumably neutral markers. In addition, microsatellites are of particularly useful to construct a genetic map of large genomes when a reference genome is absent [1]. They are favored markers for small-scale genetic studies with limited budget, potentially detecting large genetic information and physiological parameters of a genome [3], do not require high marker density, especially if LD block sizes of a genome are long [31] and benefit from inclusion of additional samples for the project without significant costs [1]. Microsatellites can be also used for testing non-neutrality and subjected to automated florescent dye-based band scoring through multiplexed genotyping for large-scale studies, which help to cut the time and cost of the study [2]. Unipartental cytoplasmic inheritance with presumably no recombination history of cpSSR [12] further provide a great advantage to develop universal primers to genotype and genetically analyze distantly related plant taxa although there are some limitations, too (see below). Importantly, EST-SSR markers developed from coding genes can be a great tool to directly tag and map-based cloning of functionally meaningful "candidate genes" through genotype to phenotype correlations in genetic mapping studies [10].

5.3. Disadvantages

There are various concerns and caveats to use microsatellites, too. Some of these include but not limited to (1) need for a priori genomic sequence information that is not available for most prokaryotes where specific effort can be costly and time consuming; (2) PCR failure due to point mutations in primer sequences resulting in 'null' alleles and falsely hiding the reality when applying PCR primers across different species with mutated primer binding sites, or because of environmental degradation of long repeat arrays; (3) PCR stutters of short SSR arrays giving multiple bands from single locus; (4) abundance for rare, private or minor alleles; (5) issues with assigning of multiple band SSRs alleles in the absence of correct parentage and pedigree information; and (6) size homoplasy, heteroplasmy and cytoplasmic introgression (in particular with cpSSR) due to back mutations during replication slippage [1, 2, 3, 9, 12]. All these complicate and bias downstream genetic analyses, inflate F-statistics or *p*-values, falsify the diversity levels, relatedness, divergence, and true evolution or phylogenetic grouping. Due to homoplasy and high rate of polymorphism in SSRs, phylogenetic studies should be carried with cautiousness for distantly related species [1].

However, these all do not void the usefulness of SSR markers, rather call attention of researchers using this marker system. There are several approaches to take into consideration of these caveats when SSRs are used that include verification of size homoplasy, heteroplasmy and primer site point mutations using additional cloning and re-sequencing including NGS [12]; exclusion of problematic, rare, and private alleles from the analyses based on specific objective

[29] (e.g., haplotype networks due to high potential of repeated evolution [12]); use of more samples and markers for genotyping rather concluding based on few samples and small number of markers; and reanalysis of the results with or in combination of different type DNA markers such as RFLPs, SNPs, etc. [12].

5.4. Future utilization

Because of past 5 years' successful and wide application of SNP markers for genome-wide applications and the emergence of cost-effective and large-scale sequencing, as well as SNP detection and genotyping methodology such as NGS, NGS-based genotyping by sequencing (GBS), and restriction site–associated DNA sequencing (RAD-Seq) techniques circumvented a rapid shift of SSR-based molecular marker studies to SNP-based research. This was evidenced by sharp decrease of a number of publications using microsatellites in crop species during 2010–2015 [2]. However, as discussed and highlighted by Hodel et al. [1], microsatellite markers will continue to be useful and favorable markers. This is due to the fact that (1) not all studies require in-depth genotyping as provided by NGS-based approaches where SSRs remain a suitable choice, and (2) sample size can be largely expanded without significant cost when SSRs are used, which is costly with NGS-based approaches. Further, (3) additional large sample inclusion can increase the power of microsatellite-based studies, which perform similarly with SNPs; (4) existed SSR-marker data can be readily incorporated and used with new studies; (5) multi-allelic nature of SSRs makes them highly suitable for studying small subpopulations; and (6) microsatellites are the best markers of choice for small-scale laboratories with limited budget.

Additionally, SSRs are still efficient markers for (1) marker-assisted selection (MAS) programs to mobilize QTL blocks using small number of SSR markers based on LD information, (2) germplasm characterization using evenly spaced core set of few SSRs, (3) seed or variety purity testing, and (4) SSR indexing of cultivars (barcoding) and plant germplasm resource. All these invalidate any emerging opinion on prospective total "death" of microsatellites as useful genetic markers and demonstrate the future benefit of microsatellites in many genetic studies. Highlights and some updates on advantages, disadvantages, and usefulness of SSRs for various applications in agricultural and biomedical fields have been presented in following book chapters of this book, which I provide a brief information below to introduce them to readers.

6. Highlights from chapters

In this context, with the objective to provide current updates on microsatellite applications in genetic studies as well as re-highlight the usefulness of microsatellites in current and future genetic analyses, in this edited volume, we compiled 10 chapters describing the wide utilization of microsatellite markers in different biological taxa. Generally, chapters presented research studies and review discussions on following three directions: (1) micro-

satellite markers in plants and genetic diversity research, (2) microsatellite markers in animal genetics and breeding, and (3) microsatellites in cancer research.

In the first section, the chapter by Jamila Bernardia and her team, Universita Cattolica del Sacro Cuore, Piacenza, Italy, presents the use of microsatellites in livestock and illustrated exploitation and versatility of microsatellites for the characterization of agricultural diversity and food traceability. Authors studied the assessment of genetic diversity in apple, pear, and sweet and sour cherry trees and explored the molecular authentication of wheat food chain of plant cultivars and farm animals. The chapter discusses that a small number of SSR markers can be efficiently used to differentiate and link each tree cultivar to its corresponding genotypic profile and be useful for molecular traceability of the whole production chain from durum wheat raw material to processed pasta despite food processing degrades DNAs.

Further, the chapter by Maria Eugenia Barrandeguy and Maria Victoria Garcia, Universidad Nacional de Misiones, Instituto de Biologia Subtropical Nodo Posadas, Argentina, has covered the development of microsatellite markers, genotyping, data analysis, and interpretation of obtained results in the examples of nuSSRs and cpSSRs. The chapter discusses the usefulness of microsatellite markers for the analysis of past and present microevolutionary forces in native forest pant populations and making inferences about future of these natural populations.

In their chapter, Rodolphe Laurent Gigant and his team from France have assessed the mating system of the natural populations of *Vanilla mexicana* (Orchidaceae) in the island of Guadeloupe. Using only six transferable SSRs out of 33 developed in other *Vanilla* species, authors successfully genotyped a set of 51 *V. mexicana* samples, which helped to differentiate *V. mexicana* samples, assess the genetic diversity and other genetic characteristics, determine a heterozygote deficiency, and estimate self-pollination rates. Results showed that "*V. mexicana* is mainly reproducing by autogamy via spontaneous self-pollination in Guadeloupe," which is reported as a useful trait for interspecific breeding of *Vanilla* species. Justyna Anna Nowakowska, from Forest Research Institute, Poland, has studied genetic structure of fourteen Scots pine populations from North-eastern Poland using SSR markers that revealed high genetic indices for the mean polymorphic information content, genetic diversity and heterozygosity. There was low population differentiation identified among stands, which were clustered into one genetically similar group. The chapter concludes that the present distribution of genetically related populations of Scots pine in North-eastern Poland seems to reflect the historical events such as post-glacial colonization of Poland from different European refugia and/or human management carried out in the past.

The last chapter in this section by Beyene Amelework, University of KwaZulu-Natal, South Africa, and Ethiopian Institute of Agricultural Research, Ethiopia, reviewed the use of microsatellite markers in genetic diversity analysis and heterotic grouping of sorghum and maize through the estimation of molecular-based genetic distance. The chapter also discusses the existing challenges with the use of SSR markers in heterotic grouping in studied crops.

The second section of the book covers microsatellite marker application in animal sciences. The chapter by Yuta Seki and his colleagues, Tokyo Metropolitan Institute of Medical Science and Tokyo University of Agriculture, Japan, has provided a review on the currently available studies on domestic goat (*Capra hircus*) breeding using microsatellite markers to demonstrate exploitation of these markers for the assessment of intra- and inter-population genetic diversity, QTL mapping, and marker-assisted selection of favorable phenotypes. Authors also stated that despite SNPs may be favorable marker for animal studies because of their large-scale genomic converge, microsatellites remain as a marker of choice for small scale genetic studies, owing to economic concerns such as cost, time and labor as well as because of their genotyping simplicity.

Further, Emil J. Hernandez-Ruz and his colleagues, Federal University of Para (UFPA), Brazil, presented a research study on microsatellite marker development and evaluation of the genetic structure of the Amazonian fish *Hypophthalmus marginatus* from the Tocantins and Araguaia River in the Eastern Amazonia. Although genetic analyses were performed using only two polymorphic microsatellite loci out of 17 developed for this fish species, results not only provided evidence on the existence of (1) low levels of genetic diversity in H. marginatus of the Tocantis basin possibility related to the Dam construction and 2) a gene flow mainly in the upstream or downstream directions but also were consistent with data from mitochondrial markers. Authors recommend the use of more markers to validate the influence of dam for reduction of genetic diversity of the Amazonian fish species.

In the chapter by Hongyu Ma and his colleagues, Shantou University and Chinese Academy of Fishery Sciences, China, authors presented a research study on the development and characterization of microsatellite markers for genetic study of the mud crab (*Scylla paramamosain*). Efforts have helped to isolate and characterize 302 polymorphic microsatellite markers. Authors have evaluated polymorphism and genetic differentiation of the mud crab wild populations, established microsatellite-based parentage assignment of the mud crab offspring, identified a marker associated with growth performance, and constructed a first preliminary genetic linkage map for *S. paramamosain* using microsatellite and amplified fragment length polymorphism (AFLP) markers. The chapter concludes that these findings should provide novel insights into genome biology, wild resource background, and molecular marker-assisted selection in S. paramamosain

The third section of the book includes two similar topic chapters that describe the impact of microsatellite instability (MSI) in causing the cancer diseases. In particular, Jeffery W. Bacher and his team from Promega Corporation and University of Wisconsin, Madison, USA, provided a detailed review on the role and significance of MSI in hereditary and sporadic type of cancers. They have discussed the discovery of MSI and its association with colorectal cancer or Lynch syndrome, and the use of SSR marker in disease screening. In addition, emerging and alternative NGS-based methods in detecting both tumor MSI status and germline mutations in a single test for LS are reviewed. The chapter concludes that MSI detection is poised to take on an even greater role in prediction of responses to the new immunotherapies targeted at MSI-positive tumors Similarly, the following chapter by Narasimha Reddy Parine and Mohammad

Saud Alanazi, King Saud University, Saudi Arabia, described the role of genetic instability, including MSI in colorectal cancer. Differing from previous chapter, this chapter reviews the major molecular mechanisms causing genomic and microsatellite instability, including a mismatch repair (MMR) system and cancer formation.

7. Conclusions

Microsatellite markers have been one of the most reliable molecular markers derived from the DNA molecule, which were widely and successfully used for life science research directions including agriculture and biomedical fields. As a molecular marker, microsatellites have many advantages suitable for the wide types of genetic analyses, but do present concerns and caveats that require attention and corrections for the results and their interpretation in specific analyses, which were highlighted by chapters of this book. Although the trends of molecular marker application and use for past 5 years show a decreased utilization of microsatellite markers and present a shifted growth toward the use of SNP markers, that is due to the emergence of novel generation NGS-based genotyping technologies, microsatellite markers remain to be useful and choice of marker system for the specific genetic studies. This is because of multi-allelic nature, simplicity of genotyping procedures, cost-effectivity, and suitability of microsatellite markers for small-scale laboratories with limited budget. In this book, all chapters re-highlighted the usefulness of microsatellites in genetic analyses of various life science fields, providing updated discussions and reviews on current use and future prospects of these markers, which invalidate emerging opinion on "full-death" of microsatellites as useful genetic markers.

Acknowledgements

I am thankful to the Academy of Sciences of Uzbekistan, Committee for Coordination of Science and Technology Development, the Office of International Research Programs (OIRP) of the United States Department of Agriculture (USDA)—Agricultural Research Service (ARS), Texas A&M University, and U.S. Civilian Research & Development Foundation (CRDF) for financial support of SSR marker–based research of cotton in Uzbekistan.

Author details

Ibrokhim Y. Abdurakhmonov

Address all correspondence to: genomics@uzsci.net

Center of Genomics and Bioinformatics, Academy of Science of the Republic of Uzbekistan, Tashkent, Uzbekistan

References

[1] Hodel RG, Segovia-Salcedo MC, Landis JB, Crowl AA, Sun M, Liu X, Gitzendanner MA, Douglas NA, Germain-Aubrey CC, Chen S, Soltis DE, Soltis PS. The report of my death was an exaggeration: A review for researchers using microsatellites in the 21st century. Appl Plant Sci. 2016;4:1600025. DOI:10.3732/apps.1600025

[2] Vieira ML, Santini L, Diniz AL, Munhoz CF. Microsatellite markers: what they mean and why they are so useful. Genet Mol Biol. 2016;39:12–328. DOI: 10.1590/1678-4685-GMB-2016-0027

[3] Saeed AF, Wang R, Wang S. Microsatellites in pursuit of microbial genome evolution. Front Microbiol. 2016;6:1462. DOI: 10.3389/fmicb.2015.01462

[4] Litt M, Luty JA. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J Hum Genet. 1989;44:397–401

[5] PubMed database [Internet]. 2015. Available from: http://www.ncbi.nlm.nih.gov/pubmed [Accessed from 2016-10-14]

[6] Trivedi S. Microsatellites (SSRs): Puzzles within puzzle. Indian J Biotechnol. 2004;3:331–347

[7] Richard GF, Kerrest A, Dujon B. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. Microbiol Mol Biol Rev. 2008;72:686–727. DOI:10.1128/MMBR.00011-08

[8] Ma ZQ, Röder M, Sorrells ME. Frequencies and sequence characteristics of di-, tri-, and tetra-nucleotide microsatellites in wheat. Genome. 1996;39:123–30

[9] Ellis JR, Burke JM. EST-SSRs as a resource for population genetic analyses. Heredity (Edinb). 2007;99:125–132

[10] Varshney RK, Graner A, Sorrells ME. Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 2005;23:48–55.

[11] Powell W, Morgante M, Andre C, McNicol JW, Machray GC, Doyle JJ, Tingey SV,Rafalski JA. Hypervariable microsatellites provide a general source of polymorphic DNA markers for the chloroplast genome. Curr Biol. 1995;5:1023–1029

[12] Wheeler GL, Dorman HE, Buchanan A, Challagundla L, Wallace LE. A review of the prevalence, utility, and caveats of using chloroplast simple sequence repeats for studies of plant biology. Appl Plant Sci. 2014;2:1400059. DOI: 10.3732/apps.1400059.

[13] Pumpernik D, Oblak B, Borstnik B. Replication slippage versus point mutation rates in short tandem repeats of the human genome. Mol Genet Genomics. 2008;279:53–61. DOI: 10. 1007/s00438-007-0294-1

[14] Siddle KJ, Goodship JA, Keavney B, Santibanez-Koref MF. Bases adjacent to mononucleotide repeats show an increased single nucleotide polymorphism frequency in the human genome. Bioinformatics. 2011;27:895–898. DOI: 10.1093/bioinformatics/btr067

[15] Amos W. Mutation biases and mutation rate variation around very short human microsatellites revealed by human-chimpanzee-orangutan genomic sequence alignments. J Mol Evol. 2010;71:192–201. DOI: 10.1007/s00239-010-9377-4

[16] Amos W. Heterozygosity increases microsatellite mutation rate. Biol Lett. 2006;12: 20150929. DOI:10.1098/rsbl.2015.0929

[17] Brinkmann B, Klintschar M, Neuhuber F, Hühne J, Rolf B. Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. Am J Hum Genet. 1998;62:1408–1415

[18] Liu Y, Wang L, Chen D, Wu X, Huang D, Chen L, Li L, Deng X, Xu Q. Genome-wide comparison of microRNAs and their targeted transcripts among leaf, flower and fruit of sweet orange. BMC Genomics. 2014;15:695. DOI: 10.1186/1471-2164-15-695

[19] Lin SL. microRNAs and Fragile X Syndrome. Adv Exp Med Biol. 2015;888:107–121. DOI: 10.1007/978-3-319-22671-2_7

[20] Pearson CE, Nichol Edamura K, Cleary JD. Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet. 2005;6:729–742

[21] Fondon JW 3rd, Garner HR. Molecular origins of rapid and continuous morphological evolution. Proc Natl Acad Sci U S A. 2004;101:18058–18063. DOI: 10.1073/pnas.0408118101

[22] Sears KE, Goswami A, Flynn JJ, Niswander LA. The correlated evolution of Runx2 tandem repeats, transcriptional activity, and facial length in carnivora. Evol Dev. 2007;9:555–565.DOI:10.1111/j.1525-142X.2007.00196.x

[23] Hammock EA, Young LJ. Microsatellite instability generates diversity in brain and sociobehavioral traits. Science. 2005;308:1630–1634

[24] Yirmiya N, Rosenberg C, Levi S, Salomon S, Shulman C, Nemanov L, Dina C, Ebstein RP. Association between the arginine vasopressin 1a receptor (AVPR1a) gene and autism in a family-based study: mediation by socialization skills. Mol Psychiatry. 2006;11:488–94

[25] Graham BM, Solomon NG, Noe DA, Keane B. Male prairie voles with different avpr1a microsatellite lengths do not differ in courtship behaviour. Behav Processes. 2016;128:53–57. DOI: 10.1016/j.beproc.2016.04.006

[26] Michael TP, Park S, Kim TS, Booth J, Byer A, Sun Q, Chory J, Lee K. Simple sequence repeats provide a substrate for phenotypic variation in the *Neurospora crassa* circadian clock. PLoS One. 2007;2:e795

[27] Bidichandani SI, Ashizawa T, Patel PI. The GAA triplet-repeat expansion in Friedreich ataxia interferes with transcription and may be associated with an unusual DNA structure. Am J Hum Genet. 1998;62:111–21

[28] Akagi T, Yin D, Kawamata N, Bartram CR, Hofmann WK, Song JH, Miller CW, den Boer ML, Koeffler HP. Functional analysis of a novel DNA polymorphism of a tandem repeated sequence in the asparagine synthetase gene in acute lymphoblastic leukemia cells. Leuk Res. 2009;33:991–996. DOI: 10.1016/j.leukres.2008.10.022

[29] Jemaa R, Ben Ali S, Kallel A, Feki M, Elasmi M, Taieb SH, Sanhaji H, Omar S, Kaabachi N. Association of a 27-bp repeat polymorphism in intron 4 of endothelial constitutive nitric oxide synthase gene with hypertension in a Tunisian population. Clin Biochem. 2009;42:852–856. DOI:10.1016/j.clinbiochem.2008.12.002

[30] Kersting C, Agelopoulos K, Schmidt H, Korsching E, August C, Gosheger G, Dirksen U, Juergens H, Winkelmann W, Brandt B, Bielack S, Buerger H, Gebert C. Biological importance of a polymorphic CA sequence within intron 1 of the epidermal growth factor receptor gene (EGFR) in high grade central osteosarcomas. Genes Chromosomes Cancer. 2008;47:657–64. DOI: 10.1002/gcc.20571

[31] Abdurakhmonov IY, Abdukarimov A. Application of association mapping to understanding the genetic diversity of plant germplasm resources. Int J Plant Genomics. 2008;2008:574927. DOI: 10.1155/2008/574927

[32] The National Center of Biotechnology Information (NCBI) [Internet]. 2016. Available from: http://www.ncbi.nlm.nih.gov [Accessed: 2016-10-14]

[33] Freitas PD, Martins DS, Galetti PM Jr. cid: a rapid and efficient bioinformatic tool for the detection of SSRs from genomic libraries. Mol Ecol Resour. 2008;8:107–108. DOI: 10.1111/j.1471-8286.2007.01950.x

[34] Martins WS, Lucas DC, Neves KF, Bertioli DJ. WebSat—a web software for microsatellite marker development. Bioinformation. 2009;3:282–283

[35] SoftLinks [Internet]. 2016. Available from: http://softlinks.amnh.org/microsatellites.html [Accessed: 2016-10-14]