Journal of Medical Systems (2019) 43: 261
https://doi.org/10.1007/s10916-019-1408-0

PATIENT FACING SYSTEMS

Spectral Analysis of Codons in the DNA Sequence of Fragile X Syndrome

Jeya Bharathi S[1] · Balamanikandan P[1]

Received: 20 March 2019 / Accepted: 26 June 2019 / Published online: 2 July 2019

Abstract

There are frequent studies undergoing related to the Fragile X syndrome caused due to the triplet CGG replicates on the X chromosome of Fragile X Mental Retardation 1 (FMR1) gene. Mutations of this chromosome can lead to Fragile X syndrome, rational disability, and other cognitive discrepancies. A novel approach based on Rajan Transform is proposed to analyze the spectral density of codons. The traditional transform like Fourier transform provides imaginary values whereas the Rajan Transform exhibits only the real values. The mutations there in the DNA are successfully distinguished by using the Rajan Transform which is suitable tool for the spectral analysis of DNA sequences. The utilization of the Rajan Transform urges larger profits in terms of minimal false alarm rate and thereby leading to an increase in the accuracy of the spectral analysis.

Keywords Fragile X syndrome · Codon · Autocorrelation · Fourier transform · Rajan transform

Introduction

Mendelian disorders

Nowadays genetic disorders are major reasons of human fatality. Genetic disorders are caused by a mutation or alteration in one gene is known as monogenic disorder. According to WHO, over 10,000 of human diseases are identified as monogenic. Monogenic disorders are also commonly quoted as Mendelian disorders. Mendelian patterns signify the biological inheritance satisfying the laws of Gregor Mendal. A mutation in the single gene, which results in the gene deviating from the Mendelian pattern, sets the road to disorders in human [2]. The mutation may occur either due to single chromosome arising from one of the parent or from both. The monogenic disorder may be either a dominant one or recessive one. The classification is typically based on whether the mutation is inherited from one chromosome or both the chromosomes of the parent. The various monogenic disorders are namely such as sickle cell disease, down syndrome, polycystic kidney disease, Haemophilia, Duchenne Muscular Dystrophy, Fragile X syndrome etc.

Fragile X syndrome

Fragile X syndrome (FXS) is an innate disorder. It is typically owed to a raise of the CGG triplet replicates inside the Fragile X mental retardation 1 (FMR1) gene which is present on the X chromosome. FMR1 gene is accountable for the protein predictable as Fragile X Mental Retardation protein (FMRP) (www.wikipedia.org/wiki/FMR1). FMRP addresses in the brain of human beings have the key role in normal development of cognitive skills. FMR1 mutations lead to various distresses in human body like FXS, Parkinson's disease, premature ovarian failure etc. An expansion occurs in the trinucleotide repeat (CGG) of 55–230 repeats is the prime causative agent of Fragile X Syndrome (www.ncbi.nlm.nih.gov/gene/2332). FXS can be roughly seen in 2% of population [3]. FXS leads to diverse range of deficiencies sketched in the Table 1 (https://en.wikipedia.org/wiki/Genetic_disorder#Single_gene_disorder). FXS is the main reason of intellectual disability in human race thereby affecting 1 out of 4000–6000 women and 1out of 5000–7000 men [4]. Owing to the facts spelled out above, it can be seen that the fragile X syndrome is one of the most crucial

This article is part of the Topical Collection on *Patient Facing Systems*

✉ Balamanikandan P
mailtobalamaths@gmail.com

Jeya Bharathi S
sjbmat@tce.edu

[1] Department of Mathematics, Thiagarajar College of Engineering, Madurai, Tamil Nadu, India

Table 1 FXS Deficits

Deficiency	Description
Intellectual disability	Difficulties in learning due to insufficient levels of adaptive and cognitive ability etc.
Behaviour anomalies	Hyperactivity, inferior eye contact, withdrawal from social activities, anxiety, broken speech etc.
Connective tissue abnormalities	Low muscle tone, hyper extensible joints, prolapsed in mitral valve etc.
Phenotypic features	Elongated face, velvety skin, prominent ears, Hyper extensible finger joints, face and jaw etc

areas to be researched in order to aid people affected by FMR1 gene mutation. There is a necessity to have an in depth analysis of the gene structure and minimal error. Error minimization is important as this area is a most sensitive and crucial field related to the health of human race. There should be no room for errors as they can lead to improper diagnosis which may have adverse effects on the people.

The related works are conferred in the next section which is related to Fragile X Syndrome, section III explains the codons and autocorrelation for codons of FMR1 gene and also discusses the proposed Fourier Transform in detail, section IV illustrates the experimental set up utilized to evaluate the proposed Rajan Transform method, it also illustrates the results that got from the evaluation and also discusses the results, and the final section is the concluding section of the paper.

Related works

A great deal of research is being done across the world in the realm of Fragile X Syndrome (FXS). Dean et al. have presented a detailed review report on the FXS. It included the summarization of the research findings for the challenges in the field of therapeutics, pathophysiology and technical problems found in the Fragile X Syndrome. The author has stated that the disorder has been discovered more than two decades ago but still has lot of surprises in its sleeves that are yet to be solved. The author also states the necessity of early diagnosis of the syndrome [1]. Pugin et al. has in detail stated about the FMR1 disorders in the ambit of clinical, pharmacological and molecular aspects. The author has stated that 98% of the FXS cases have been found to be due to CGG triplet expansion of the FMR1 gene. The author also has in detail expounded the neurobiological aspects of the FMR1 including the FMR protein (FMRP) roles in human cognitive development and FMRP related disorders. The author has also illustrated the action of the Fragile X Syndrome which is about genetic counseling [5]. Mayo et al. have done various experiments on the mice with normal gene and the mice that have mutation in their FMR1 gene. The mice were subject to various stress conditions like no stress, restraint stress and unpredictable stress conditions for a total of 12 days. The authors have observed that chronic stress lead to the induction of the social interaction and working memory [6]. Jiraanont et al. have performed a series of experimental analysis and concluded that the deletion of 300 KB of DNA also leads to the Fragile X syndrome [7].

Wang et al. have devised powerful methodologies for large scale analysis of the individual synapses. Their analysis revealed distinct mutations in the synaptic proteins across 6000 metrics [9]. Clark et al. have analyzed nearly 321 cases of CGG repeats of FMR1 gene. The authors have evaluated the relation between the Essential Tremor and FMR1 alleles. The findings of the author have suggested that FMR1 CGG repeats do not form a genetic risk factor for Essential Tremor [10]. K. Vijayakumar et al. deals with the need for computerized threat estimation with the help of NLP to auto recognize the threat on the scrutiny of fault and susceptibilities [11]. Dorota Bielinska-Waz et al. described a moment of inertia method preserves as functional to an subjective system of distinct objects, in meticulous to the spectrum representing the DNA Sequence [10, 17]. K. Vijayakumar et al. proposed feature selection algorithm process for text classification by means of the algorithms ant colony optimization (ACO) and artificial neural network (ANN). This crossbreed approach is pretended using Reuter's data set and established its competence [13]. Nilay Chheda, Naman Turakhia et al. developed the data using Digital Signal Processing (DSP) techniques for the conversion of symbolic data into numerical data using encodings and transformations such as Fast Fourier Transform (FFT) and export it with dissimilar size windows and search for significant spectral prototype by selecting time delay in the plot using Biospectrogram [14]. Durbin et al. described that large number of influential sequence analysis methods are based on doctrine of probabilistic representation of derived score matrices to find out the implication of sequence alignments [15]. Shinahara Kumi et al. proposed revoke transcription-PCR revealed that a large quantity of the mutant transcripts joined unexpectedly, causing premature execution of the protein synthesis. Although uncommon, point mutations in the FMR1gene may be a cause of autism and mental retardation in Japanese patients [16]. Gustin, S. L., et al. examined that the association among Fragile X Mental Retardation 1 gene (FMR1) Cytosine-Guanine-Guanine (CGG) repeat number and ovarian

Table 2 Total number of Codons for FMR1 Gene

aaa	aac	aag	aat	aca	acc	acg	act
75	29	66	42	36	22	12	21
aga	agc	agg	agt	ata	atc	atg	att
70	32	52	35	29	23	47	38
caa	cac	cag	cat	cca	ccc	ccg	cct
41	25	49	24	35	10	41	22
cga	cgc	cgg	cgt	cta	ctc	ctg	ctt
9	6	61	11	23	14	31	23
gaa	gac	gag	gat	gca	gcc	gcg	gct
61	20	51	49	32	11	13	31
gga	ggc	ggg	ggt	gta	gtc	gtg	gtt
52	22	22	25	24	20	30	25
taa	tac	tag	tat	tca	tcc	tcg	tct
36	17	23	21	36	25	11	17
tga	tgc	tgg	tgt	tta	ttc	ttg	ttt
50	27	36	28	21	32	33	32

reserve, with a particular focus solely on the assortment of CGG repeat number below the premutation (PM) range (< 55 CGG repeats). It is reported that statistically significant correlation of ovarian reserve and CGG repeat number in women with < 55 CGG repeats [17]. Kremer, E. J., et al. determined that the repeat show signs of instability both when replica in a non-homologous host and after intensification by the polymerase chain reaction. These consequences imply difference in the trinucleotide repeat copy number as the molecular basis for the instability and perhaps the fragile site [18]. Hauberg, M. E., et al. observed the role of miRNAs in schizophrenia in the perspective of disease-associated hereditary deviation and utilized an diagnostic framework that generally measured the role of miRNAs in common-variant schizophrenia vulnerability and initiate further substantiation for their association [19]. Amancio, A. P., et al. validated the molecular inheritance verdict of patients alleged of Fragile X Syndrome (FXS) and two PCR analyses were carry out using dissimilar primers, one for screening (PCR-T) and other for the detection of the pre-mutation (PCR-P) [13].

From the above concise discussion, it can be seen that different research groups are actively involved in finding various aspects of the syndrome and the FMR1 gene. To aid in the diagnosis of the Fragile X Syndrome, we have proposed an mathematical method using Rajan Transform in analyzing the spectral density of codons in FMR1gene.

Codons

The inherited code is a set of regulations which converts the information encoded inside DNA sequences into proteins. A three letter nucleotide forms a component of inherited code in a DNA or RNA known as codons. Amongst all organisms, the genetic code is highly analogous and it is articulated in Table 2 with 64 ways. The codons illustrate the process in addition of amino acids sequentially during protein synthesis. The amino acids are carried to the ribosomes by Transfer RNA (tRNA) and interpret the Messenger RNA (mRNA) one codons at an instance. The translated initial codon called as start codons such as AUG and termination called as stop codons such as UAG, UGA and UAA. There are totally 64 codons of Fragile X Mental Retardation 1(FMR1) which is tabulated in Table 2.

Autocorrelation

Autocorrelation is the connection between observation and time lag. The scrutiny of autocorrelation is a statistical technique for resulting repeating patterns of the given

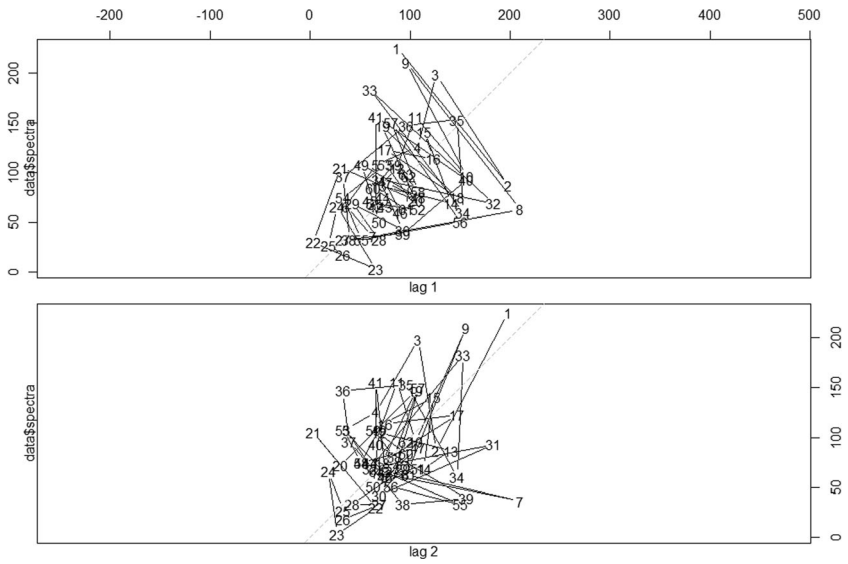


Fig. 1 Fourier Transform Auto correlation plot for Fragile Mental Retardation1 (FMR1) gene

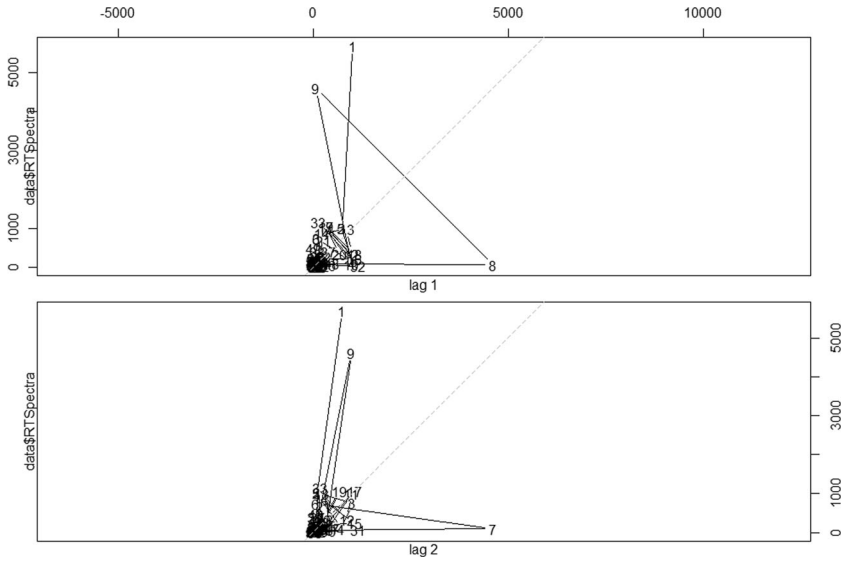


Fig. 2 Rajan Transform Auto correlation plot for Fragile Mental Retardation1 (FMR1) gene

observations. The mean and variance regularizes the autocorrelation function. For m observations$\{A_1, A_2, A_3, ..., A_m\}$, the autocorrelation can be acquired as

$$S(t) = \frac{1}{(m-t)\sigma^2} \sum_{p=1}^{m-t} (A_p - \mu)(A_{p+t} - \mu) \quad (1)$$

For data articulated as a discrete sequence, it is often necessary to calculate the autocorrelation with elevated computational competence (See Figs. 1, 2, 3 and 4).

Fourier transform

The Fourier transform is the frequency realm depiction of a novel key sequence and is given by

$$a(t) = \sum_{s=0}^{T-1} x_s e^{-j\frac{2\pi st}{T}} \quad (2)$$

The sum of spectrum of all 64 codons forms the sequences which are as follows:

{225,87,198,126,108,66,36,63,210,96,156,105,87,69,141,114,123,75,147,72,105,30,3,66,27,18,33,33,69,42,93,69,183,60,153,147,96,33,39,93,156,66,66,75,72,60,90,75,108,51,69,63,108,75,33,51,150,81,108,84,63,96,99,96}.

This sequence is called the characteristic sequence of the strand. The graphical representation of the Fourier Transform for Fragile Mental Retardation1 (FMR1) gene is illustrated (in Fig. 1).

Rajan transform

Rajan Transform is a novel transform and a variant of Decimation-In-Frequency Fast Fourier Transform (DIF-FFT)

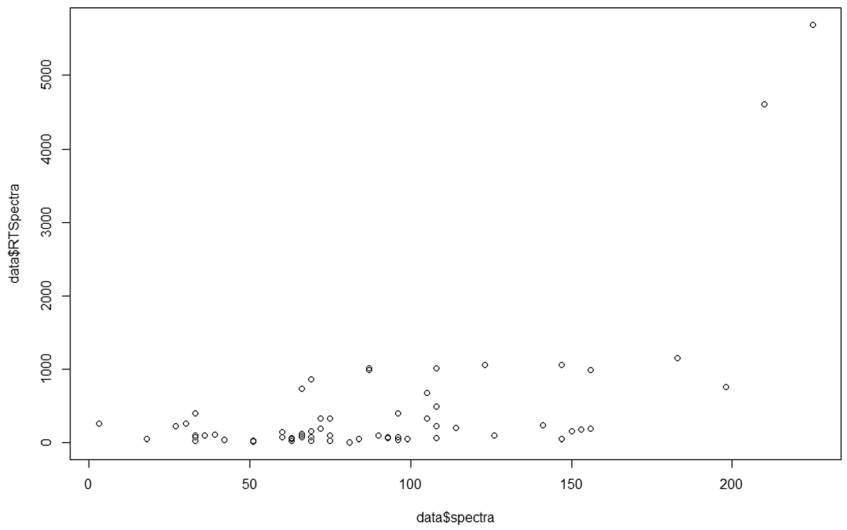
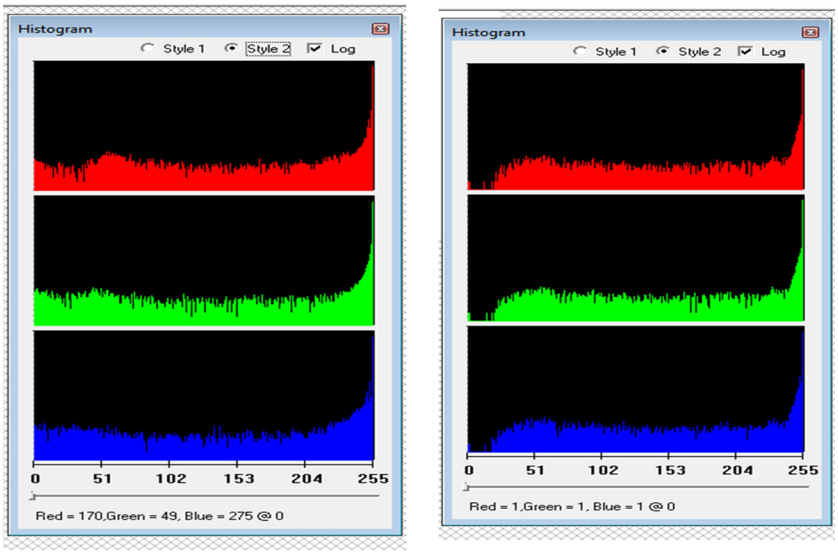


Fig. 3 Correlation of codon spectrum and Rajan Transform spectrum



(a)Normal (b)Mutant

Fig. 4 (**a**) & (**b**) Histogram Variations for Normal and Mutant DNA Sequence of Fragile X Syndrome of Secondary Structures

algorithm. Rajan Transform is homomorphic and it is more appropriate for pattern recognition. x(n) represents a sequence of number at T length, which is a power of 2, it is first alienated into two parts a(s) and b(s). Each of the first and the second parts include (T/2) points is obtained as follows:

$$a(s) = x(r) + x\left(r + \left(\frac{T}{2}\right)\right); 0 \le s \le \frac{T}{2}, 0 \le r \le \frac{T}{2} \tag{3}$$

$$b(s) = |x(r) - x\left(r - \frac{T}{2}\right)|; \quad 0 \le s \le \frac{T}{2}, \left(\frac{T}{2}\right) \le r \le T \tag{4}$$

Each of the T/2 point section is further separated into two parts, where each includes of T/4 points as follows:

$$a_1(t) = a(s) + a\left(s + \left(\frac{T}{4}\right)\right), 0 \le t \le \frac{T}{4}, 0 \le s \le \frac{T}{4} \tag{5}$$

$$a_2(s) = a(r) - g\left(a - \left(\frac{T}{4}\right)\right), 0 \le t \le \frac{T}{4}, \frac{T}{4} \le s \le \frac{T}{2} \tag{6}$$

$$b_1(t) = b(s) + b\left(s + \left(\frac{T}{4}\right)\right), 0 \le t \le \frac{T}{4}, 0 \le s \le \frac{T}{4} \tag{7}$$

$$b_1(s) = b(r) - b\left(s - \left(\frac{T}{4}\right)\right), 0 \le t \le \frac{T}{4}, \frac{T}{4} \le s \le \frac{T}{2} \tag{8}$$

This iteration is processed until there is no other division left to execute. Hence, the total number of iterations are log T. Here + and − symbols are represented as the sum and difference operators, respectively. x(n) is the sequence for the number at length $T = 2^t$; $t > 0$, then its Rajan transform is denoted by x(t). Rajan Transform can be applied to any of the number sequence since it promotes isomorphism of the sequence class, i.e. it helps in designing a domain set that has cyclic and dyadic variations in its sequence. When t = 0, the value of x(t) = 5691, likewise we repeat the steps further till the value of k = 63. The Rajan Transform spectrum sequence are given by.

{5691,1017,753,99,1011,729,93,63,4605,69,987,333,987,867,231,201,1059,333,1059,333,675,261,261,123,219,51,69,21,159,39,69,27,1155,141,171,45,399,393,111,57,189,75,99,93,189,75,99,93,489,15,75,27,219,21,93,21,153,3,57,45,51,33,45,39}.

Rajan Transform spectrum (See Fig. 2) discloses additional information as it is an enhanced method for extracting the concealed information within the sequences. Due to the simple preservative method adopted for the extraction of the Rajan Transform of any given sequence, we will enclose calculated cumulative points of significance on the frequency axis. When the length of the input sequence is N, then the calculated points on the frequency alignment will be N/2, N/4, 3 N/4, N/8, 7 N/8. At these positions, we ought to expect the highest point and appropriate information. Rajan transform being homomorphic in nature, easy for computation (no aggravate of unreal numbers), permutation invariant, possessing dual property and a unique inheritance property; all these possessions that construct Rajan Transform as the most ideal tool for spectral analysis rather than other transforms like Discrete Fourier Transform (DFT) or Fast Fourier Transform (FFT).

The correlation value of codon spectrum and Rajan Transform spectrum (in Fig. 3) is 0.5989654. The results of Fragile X Syndrome Normal and Mutant are plotted by using Logical Image Processing System. The histogram (See Fig. 4(a) & (b)) is illustrated for FMR 1 Normal and Mutant which is a representation of a distribution consists of group of data points.

Conclusion

Fragile X Syndrome was first discovered twenty decades back. Since the discovery of the syndrome various researches has been going on in the field of the Fragile X Syndrome. Currently there has been various researchers are trying to find remedy that will induce fast recovery. At the same time early diagnosis is very much necessary to ensure quick start of the therapy for the affected children and also frequent counselling for the affected child's parents. The results illustrate that the spectrum attained by using Rajan Transform gives improved perceptive of the hidden information in the DNA sequences. Thus we have proposed a novel approach of Rajan Transform Model in the spectral analysis of the codons in FMR1 gene.

Compliance with ethical standards

Conflict of interest The authors have no conflict of interest.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Dean, D. D. et al., Fragile X syndrome: Current insight. *The Egyptian Journal of Medical Human Genetics* 17:303–309, 2016.
2. Guo, M. H. et al., Determinants of power in gene-based burden testing for monogenic disorders. *The American Journal of Human Genetics* 99:527–539, 2016.
3. Brosco, J. P., Mattigly, M., and Sanders, L. M., Impact of apecific medical interventions on reducing the prevalence of medical retardation. *Arch Pediatr Adolesc Med* 160:302–309, 2006.
4. Ciaccio, C. et al., Fragile X syndrome: A review of clinical and molecular diagnoses. *Italian Journal of Pediatrics* 43:39, 2017.
5. Pugin et al., Clinical, molecular, and pharmacological aspects of FMR1-related disorders. Neurología. 32(4):241–252, 2017.
6. Mayo et al., Behavioral effects of chronic stress in the Fmr1 mouse model for fragile X syndrome. *Behavioural Brain Research* 320: 128–135, 2017.
7. Jiraanont et al, Germinal mosaicism for a deletion of the FMR1 gene leading to fragile X syndrome, *European Journal of Medical Genetics*, 2016.
8. Wang et al., FMR1 KO and Fenobam treatment differentially impact distinct synapse populations of mouse neocortex. *Neuron* 84: 1273–1286, 2014.
9. Clark et al., Genetic analysis of FMR1 repeat expansion in essential tremor. *Neuroscience Letters* 593:114–117, 2015.
10. Vijayakumar, K. et al. Automated risk identification using NLP in cloud based development environments. *J Ambient Intell Human Comput*, Springer 2017.
11. Bielinska-Waz, D., Waz, P., Spectral-Dynamic representation of DNA sequences. *Journal of Biomedical Informatics*, 2017.
12. Vijayakumar, K. et al. An ACO–ANN based feature selection algorithm for big data. *Cluster Comput*, Springer 2018.
13. Chheda, N., Turakhia, N., et al. Biospectrogram: a tool for spectral analysis of biological sequences, *Bioinformatics* 2012.
14. Durbin, R., Eddy, S., Krogh, A., and Mitchison, G., Biological sequence analysis. Cambridge, UK: Cambridge University Press, 1998.
15. Kumi, S., Takahiko, S., Kenji, M., and Yasuhiro, K., Single-strand conformation polymorphism analysis of the *FMR1*gene in autistic and mentally retarded children in Japan. *The Journal of Medical Investigation* 51:52, 2004.
16. Gustin, S. L., et al. Evidence of an age-related correlation of ovarian reserve and FMR1 repeat number among women with "normal" CGG repeat status. *J Assist Reprod Genet*, 2015.
17. Kremer, E. J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S. T., Schlessinger, D., Sutherland, G. R., and Richards, R. I., Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. *Science* 252:1711–1714, 1991.
18. Hauberg, M. E., et al. Analyzing the Role of MicroRNAs in Schizophrenia in the Context of Common Genetic Risk Variants. *JAMA Psychiatry*, 2016
19. Amancio, A. P., et al. Molecular analysis of patients suspected of Fragile X Syndrome. *Genet Mol Res* 2015