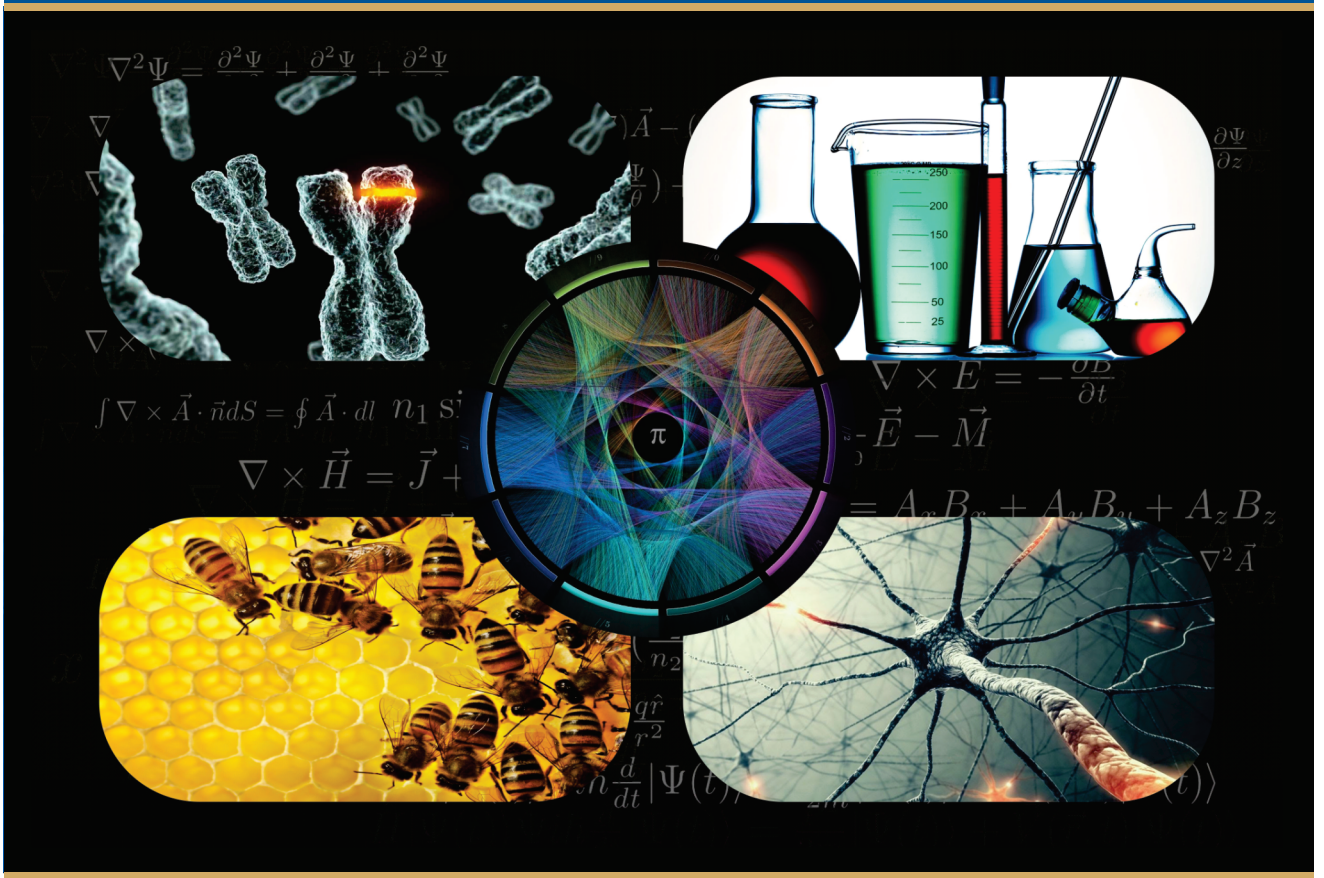
VOLUME 5
NUMBER 2
2021

Khazar Journal of Science and Technology



KHAZAR
UNIVERSITY PRESS

Vol. 5, №2, 2021 ISSN 2520-6133

KHAZAR JOURNAL OF SCIENCE AND TECHNOLOGY (KJSAT)

Hamlet Isayev/Isaxanli[1] *
Editor-in-Chief
Khazar University, Azerbaijan

[1] Also H. Issakhanly, H. Isakhanly, H. Isakhanli, H.A.Isayev, G.A.Isayev, or G.A.Isaev due to differences in transliteration. I use Hamlet Isayev in mathematics and science fields and Hamlet Isakhanli/Isaxanli in humanities.

Khazar University
41 Mehseti str., Baku AZ1096
Republic of Azerbaijan
Tel: (99412) 4217927
Website: **www.kjsat.com**

KHAZAR UNIVERSITY PRESS

Contents

Khazar Journal of Science and Technology Volume 5 №2 2021, 5-19
© Khazar University Press 2021 DOI: 10.5782/2520-6133.2021.5.2.5

Quantitative Trait Locus Mapping: Some Biological and Statistical Considerations

Seyyed Abolghasem Mohammadi[1,2,3,*], Javid Ojaghi[3]

[1]Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Tabriz, Tabriz, Iran

[2]Center of Excellence in Cereal Molecular Breeding, University of Tabriz, Tabriz, Iran

[3]Center for Cell Pathology, Department of Life Sciences, Khazar University, Baku AZ1096, Azerbaijan

**Corresponding author: mohammadi@tabrizu.ac.ir*

Abstract

Many important traits in plant and animal populations such as yield, quality, and resistance/tolerance to biotic and abiotic stresses are controlled by many genes with small effects and are known as quantitative traits (also 'polygenic,' 'multifactorial' or 'complex' traits). The regions within genomes that contain genes associated with variation of quantitative traits are known as quantitative trait loci (QTL). A key development in the field of complex trait analysis was the establishment of large collections of molecular/genetic markers, which could be used to construct detailed genetic maps of both experimental and domesticated species. These maps provided the foundation for the modern-day QUANTITATIVE TRAIT LOCUS (QTL) mapping methodologies. The identification of QTLs can help to understand how many genomic regions significantly contribute to the trait variation in a population and how much variation is due to additive, dominant or epistatic effects of QTLs. Although the basic principle of QTL mapping has been established in Sax's work on beans, the identification of QTLs based only on conventional phenotypic evaluation is possible. A major breakthrough in the characterization of quantitative traits was initiated by the development of RFLP markers which created opportunities to select QTLs. After that generally, biparental populations are used to map QTLs, in which the marker genotype and trait phenotype data are analyzed to detect the association between the two. The advent of molecular marker technology and the development of detailed linkage maps in various organisms made it possible to dissect QTs into discrete genetic factors. This review focuses the discussion on the biological considerations and statistical methods used for mapping QTLs.

Keywords: QTL mapping, RILs, SIM, molecular markers

1. Introduction

Much of the natural variation observed in the crops, domestic animals, and other populations is due to much more minor genetic changes in many genes. With laid down of the basic theoretical foundations of quantitative genetics by R.A. Fisher and the establishment of Quantitative Genetics, the focus was to partition the overall variation of quantitative traits into genetic and environmental ones. With the development and advancement of polymorphic markers in many species, one of the interesting and applied areas of researches in genetics, plant, and animal breeding is to partition genetic variation to individual quantitative trait loci (QTL) in the genome as well as interaction among them (Zeng et al., 1999; Doerge, 2002). A QTL is a region of the genome that is associated with an effect on a quantitative trait. Conceptually, a QTL can be a single gene, or it may be a cluster of linked genes that affect the trait. The aims behind QTL analysis are to:

- ✓ detect the genomic regions affecting the trait: where are the QTLs
- ✓ determine how much of the variation for the trait is caused by a specific region
- ✓ identify the gene action associated with the QTL additive/dominant effects
- ✓ identify the allele associated with the favorable effect
- ✓ determine whether there is a 'hot-spot on particular chromosomes for particular traits or is there a relatively random distribution
- ✓ assign breeding values to lines or families based on their genotypes at one or more QTLs.

The general steps in QTL mapping using bi-parental populations include (1) selection of parental lines that differ for traits of interest and generating segregating population, (2) selection of molecular markers such as SSR, and SNP for screening the two parental lines, (4) genotyping and phenotyping of the mapping population, and (5) detection of QTL using a suitable statistical method (Xu et al., 2017).

Mapping population

Detection of QTLs needs segregating populations which can be natural populations or populations developed from a cross between two lines. In practice, the most commonly used experimental designs for locating QTL start with two parental inbred lines, $P_1$and P_2, differing both in trait values and in the marker (M, N, ...) variants or alleles (M_1, M_2, N_1, N_2, ...) they carry. The F_1 individuals obtained from the cross of two homozygote lines are heterozygous at all markers and QTL regions. From the

F_1 population, crosses are made to generate F_2, backcrosses (BC), doubled haploids (DH), recombinant inbred lines (RIL), and near-isogenic lines (NIL) populations (figure 1) (Singh & Singh, 2015a; Xu et al., 2017).

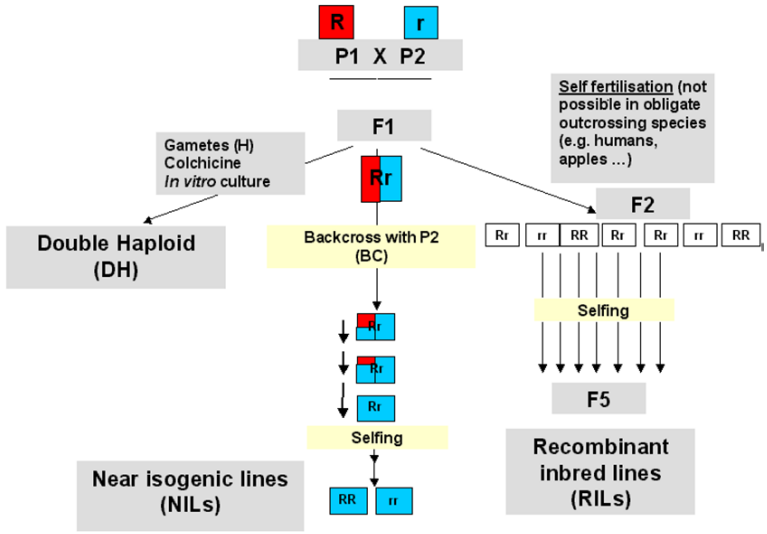


Figure 1. A schematic representation of the various bi-parental mapping populations.

F2: A F_2 population is generated by selfing or sib-mating of the F_1 individuals from a cross between the two homozygote parents (figure 1). F_2 population consisted of all parental allele combinations and each individual is expected to have a unique combination of linkage blocks from the two parental lines. In the process of F_2 population generation, only a single meiotic cycle, in other words, only one round of recombination can occur between any two loci. Therefore, the estimates of linkage between pairs of loci based on F_2 populations are not more accurate and the estimated genetic distances between two loci are likely to be greater than those detected by the population undergoing several rounds of recombination. F_2 populations are the best suited for preliminary linkage analysis and QTL mapping. The expected ratios for dominant and codominant markers in an F_2 population are 3:1 and 1:2:1, respectively. The development of F_2 populations requires the minimum times compare to the other segregating populations and it takes only two generations. The presence of all possible genotypes at a locus in F_2 populations provides the estimates of additive, dominance, and epistatic components of the

genetic variance. Each individual in the population captures the recombination events from both male and female parental gametes, which makes it ideal for identifying heterosis QTLs.

Backcrosses (BC): BC populations are generated by crossing F_1 plants with either of the two parental lines (figure 1). Due to the absence of one of the two possible homozygote genotypes at each locus in BC, genetic analysis using dominant markers can be performed only when the marker allele is absent in recurrent parent, but this is not the case in codominant markers. Backcross exhibits 1:1 ratio for dominant and codominant markers. The individuals in BC populations undergone only one cycle of recombination. Therefore, in BC populations like F_2 populations, recombination is not fixed and cannot be evaluated in replicated trials, which makes them unsuitable for QTL mapping. Compared to an F_2 population, a backcross population is less informative for linkage mapping because recombination among markers occurs in only one set of gametes (*i.e.*, the F_1) (Lander et al., 1987).

Doubled haploid (DH)*:* DH plants are usually produced by anthers/pollen culture of F_1 plants followed by chromosome doubling of haploid plants using colchicine (figure 1). The DH lines are completely homozygous at all the loci and recombination in DH lines is fixed, therefore they can be multiplied and maintained indefinitely and can be evaluated in replicated trials. The expected Mendelian ratio for each locus in a DH population irrespective of the dominant or codominant nature of the genes is 1:1. DHs. DH populations, like RILs, are perpetual as they can be multiplied and maintained indefinitely and can be evaluated in replicated trials.

Recombinant inbred lines (RILs)*:* RILs are a set of homozygous lines produced by continuous inbreeding/selfing of individual F_2 plants *via* the single seed descent (SSD) method (figure 1). The RIL population generated by adequate generations of selfing consists of homozygotes lines with different recombination from parental genomes. The expected ratio of the two homozygotes lines at each locus in the population is 1:1. Due to several cycles of recombination, RILs enable the detection of markers located much closer to the target gene than is possible with F_2, DH, and BC populations. Since RILs are homozygous, like DHs they can be propagated indefinitely without any further change in their genotype and recombination structure; this makes RILs essentially a perpetual or permanent mapping population (Burr et al., 1988).

Near-isogenic lines (NILs): NILs are developed by continuous backcrossing to recurrent parent followed by a single generation of selfing. They are homozygous lines that are identical in genotype, except for a single gene/locus or a variable length of the genomic regions flanking the target locus. They may also differ for some random genomic segments located elsewhere in the genome. Like DHs and RILs, NILs are homozygous and perpetual mapping resources, but they are usually used for fine mapping of a specific gene/genomic region and are not common populations for the construction of linkage map and QTL mapping (Muehlbauer et al., 1988).

Over the years, bi-parental mapping populations have been used to map QTLs for various traits in crop plants such as barley (Hussain et al., 2016; Barati et al., 2017; Du et al., 2019; Capo-Chichi et al., 2021) wheat (Badakhshan et al., 2008; Azadi et al., 2015; Ehdaie et al., 2016; Guo et al., 2020; Wang et al., 2021), rice (Li et al., 2016; Amoah et al., 2020; Li et al., 2020).

Principle of QTL mapping

Quantitative trait loci (QTL) analysis is a methodology that combines DNA marker and phenotypic trait data to locate and characterize genes that influence quantitative traits. The individuals in the mapping population are partitioned into different groups based on their marker genotypes to determine whether significant differences exist between groups with respect to the trait being measured. A significant difference between phenotypic means of the groups (either 2 or 3) indicates that the marker locus is used to partition the mapping population is linked to a QTL controlling the trait (Tanksley, 1993; Young, 1996; Collard et al., 2005).

For genotyping of the population, DNA markers are first used to 'screen' (or evaluate) the parents of a mapping population for polymorphisms, detectable differences in marker patterns. After polymorphic markers are identified, they are used to evaluate each line or individual of the mapping population. Each line is scored for having the marker pattern corresponding to one or the other parent. The number of polymorphic markers needed for a QTL study will depend on the genome size of the species, the average spacing between markers, and the objectives of the study. In the next step, polymorphic markers will be used for the construction of a linkage/genetic map consisting of linkage groups in which relative positions and distances (cM) of loci are determined. In the saturated linkage map, the number of linkage groups is equal to the number of haplotype chromosomes in the species. Before linkage analysis, each marker locus is generally analyzed for evidence of segregation distortion, the deviation of observed segregation ratios from the ratios expected with Mendelian inheritance.

Statistical methods to map QTLs

Three commonly used methods for detecting QTLs in the bi-parental populations are single-marker analysis, simple interval mapping, and composite interval mapping (Liu, 1998; Tanksley, 1993).

Single-marker analysis: It also called 'single-point analysis' is the simplest method that examines the association of single marker variants with the trait variation at a

time. To calculate the strength of the association between genotype and phenotype, the mapping population is split into two or three groups, according to their marker genotypes, then the mean trait value of these two or three classes is compared. If the difference is significant, then this provides initial evidence for the location of a QTL in the neighborhood of the marker (Young, 1996). In the simplest case, linear equations can be developed to describe the relationship between a trait and each molecular marker using the following form: Y = μ + f(marker) + e, where, Y is the trait value, μ is the population mean, f(marker) is a function of the molecular marker, and e is an error. The linear model can be assessed by using t-tests, analysis of variance (ANOVA), and linear regression.

The t-test is the simplest method to test for trait men difference between two marker groups. For example, let $\hat{\mu}_{MM}$ and $\hat{\mu}_{Mm}$ be the observed trait means of individuals with marker genotypes *MM* and *Mm* for a marker in a backcross population, the *t* statistics for testing significance between $\hat{\mu}_{MM}$ and $\hat{\mu}_{Mm}$ is:

$$t = \frac{\hat{\mu}_{MM} - \hat{\mu}_{Mm}}{\sqrt{s^2\left(\frac{1}{n_{MM}} + \frac{1}{n_{Mm}}\right)}}$$

Where s^2 is the pooled sampling variance, and n_{MM} and n_{Mm} are corresponding sample size in each marker class. Significant *t* statistics show the presence of putative QTL in the vicinity of the tested marker locus. After locating the QTL, it is possible to estimate the effect of detected QTL on trait variance using the following formula: $\epsilon(\hat{\mu}_{MM} - \hat{\mu}_{Mm)=(1-2r)a}$, where ϵ denote the expectation, *a* is the effect of identified QTL and *r* is the recombination frequency between marker locus and QTL. It should be considered that using the backcross population it is not possible to separate QTL additive and dominance effects and *a* indicates the QTL genetic effect. In the case of DH and RIL populations, $\epsilon(\hat{\mu}_{MM} - \hat{\mu}_{mm)/2=(1-2r)a}$, where *a* is an estimate of the QTL additive effect.

In the F_2 population, individuals are classified into three genotypes groups based on each codominance marker locus, therefore ANOVA will be powerful than *t*-test for the significant test of the phenotypic means of genotypic groups. The analysis gives an F statistic and provides a quick and simple method to detect which markers are associated with a QTL. Since F_2 population consisted of all three possible genotypes in marker locus, therefore additive and dominance effects of QTL can be estimated as: $\hat{\mu}_{MM} - \hat{\mu}_{mm})/2 = a(1-2r)$ and $\hat{\mu}_{Mm} - (\hat{\mu}_{MM} + \hat{\mu}_{mm})/2 = d(1-2r)^2$, where let $\hat{\mu}_{MM}$, $\hat{\mu}_{mm}$ and $\hat{\mu}_{Mm}$ be the observed trait means of individuals with marker genotypes *MM*, *mm* and *Mm* for a marker in an F_2 population

Linear regression is the most commonly used statistical method for detection association between a marker locus and traits variation because the marker coefficient of determination (R^2) explains how much of the phenotypic variation is associated with the QTL linked to the marker. The marker and trait association is tested using the linear model of $y_i = \beta_0 + \beta_{x_i} + e_i$, where y_i is the trait value of the i^{th} individual in a population, β_0 is the mean (intercept), β is coefficient of regression showing the association, x_i is a dummy variable related to the i^{th} individual marker genotype taking a value of 1 and 0 for *MM*, and *Mm* marker genotypes in a BC population, respectively and e_i is a random residual variable for the i^{th} individual. The marker with significant regression coefficient is the one that is linked to the QTL.

In genetic terms, this method relies on the linear relationship between the size differences in the marker classes phenotypic means and the recombination frequency between the QTL and the individual marker, which is expressed as $1/2D_i = a_i(1 - 2r_i)$, where D_i and r_i are the difference between the phenotypic means of i^{th} markers classes and the recombination frequency between the QTL and the i^{th} marker, respectively (Kearsey & Hyne, 1994). At the true position of the QTL, this is linear regression of $½D_i$ on $(1- 2r_i)$ with sloped which passes through the origin.

Single marker analysis is simple and does not need a linkage map, and QTL mapping was initially carried out by looking for an association between genotypes at individual markers and phenotypic variation of target traits. There are three problems with this approach (Lander & Botstein 1989):

i) the analysis cannot determine whether a significant marker effect is due to one or QTLs.

ii) even in the cases of a single QTL, it cannot determine the significant marker effect is due to closely linked QTL with small effect or distantly linked with large effect.

iii) the method cannot estimate the likely positions of the QTLs, and the QTL effect is confounded with the QTL distance from marker, i.e., recombination frequencies

Simple interval mapping (SIM): To overcome some problems of single-marker analysis, Lander and Botstein (1989) proposed simple interval mapping based on the maximum likelihood method which makes use of linkage maps and analyses intervals between a pair of adjacent linked markers along chromosomes simultaneously, instead of analyzing single markers. In SIM, the interval between two adjacent markers is tested for the presence of a putative QTL by performing a likelihood ratio test (LRT).

To perform the test, the Likelihood is calculated for a given set of parameters (particularly QTL effect and QTL position) given the observed data on phenotypes and marker genotypes. The estimates for the parameters are those where the likelihood is highest. Finally, measure of the strength of evidence for the presence of a QTL at give interval, *e.g.,* λ can be tested with a likelihood ratio test using likelihood ratio (LR); $LR = ln\frac{\Pr(y|QTL\ at\ \lambda, \mu_{MM\lambda}, \mu_{Mm\lambda}, \sigma_\lambda}{(y|no\ QTL,\ \mu,\ \delta)}$ or logarithm of odd (LOD), $LOD = log_{10}\frac{\Pr(y|QTL\ at\ \lambda, \mu_{MM\lambda}, \mu_{Mm\lambda}, \sigma_\lambda}{(y|no\ QTL,\ \mu,\ \delta)}$, where $\mu_{MM\lambda}$, $\mu_{Mm\lambda}$ and σ_λ are the maximum likelihood estimates, assuming a single QTL at position λ. There is a one-to-one correspondence between LR and LOD, and LR can be translated into LOD as $LOD = 0.217LR$. LOD is the most commonly used test statistics in SIM and LOD =3 means that the top model (presence of QTL) is 1000 times more likely than the bottom model (absence of QTL). This test is performed at any two adjacent marker intervals in each linkage group. If the likelihood ratio test statistic at a given markers interval exceeds a predefined critical threshold or equal to the critical threshold, a QTL is estimated at the position of the maximum test statistic at that interval. Depend on the size of genome, density of markers in linkage map and type and size of population, the threshold at %5 significant level over a whole genome is generally between 2 and 3.5 on LOD score (Zeng, 1994). Alternatively, using QTL mapping software the relevant threshold for a given data set can be determined from the data by using permutation or bootstrap test. The QTL likelihood curve of the LOD is derived by plotting the LOD score against marker position in the genome. The LOD curve achieves the critical threshold or above it indicates the presence of a QTL at this position (figure 2).

Some computer software uses multiple regression of phenotypic data on marker genotypic data to perform SIM as described by Haley and Knott (1992). The results of SIM based on regression analysis are very similar to those obtained with the maximum likelihood approach, except in the presence of a large proportion of missing marker data. However, computationally the multiple regression approach is faster than the maximum likelihood method and also more robust if the assumption of a Gaussian distribution of residuals is violated. The F-values obtained in the regression analyses are converted into LOD scores by using the transformations (Haley & Knott, 1992).

Simple interval mapping is statistically more powerful than single-marker analysis and provides a LOD score curve that allows localization of the QTL onto the linkage map. In addition, the estimate of QTL effect is more reliable and is not confounded with the QTL distance from the marker. Finally, the missing marker genotype data are taking into account, which enhances reliability of the findings. However, SIM has some limitations. The method assumes a single QTL in the interval of two

adjacent markers. The QTL effect can be biased when more than one QTL present at the market interval and if two QTLs locate close to each other, it will detect a single "ghost" QTL. Implementation of SIM requires more computation time than single-marker analysis (Singh & Singh, 2015b).

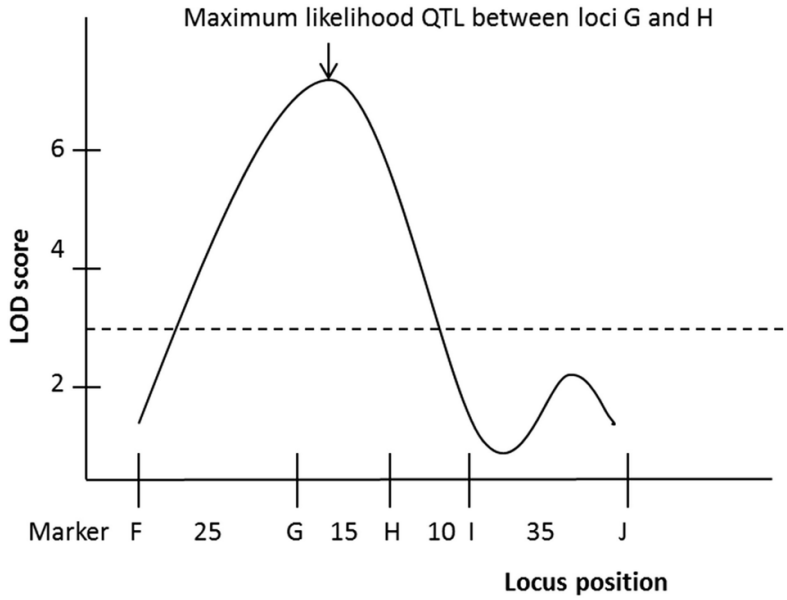


Figure 2. The LOD curve indicating that the most likely QTL position (peak of the curve). Adapted from Boopathi (2020).

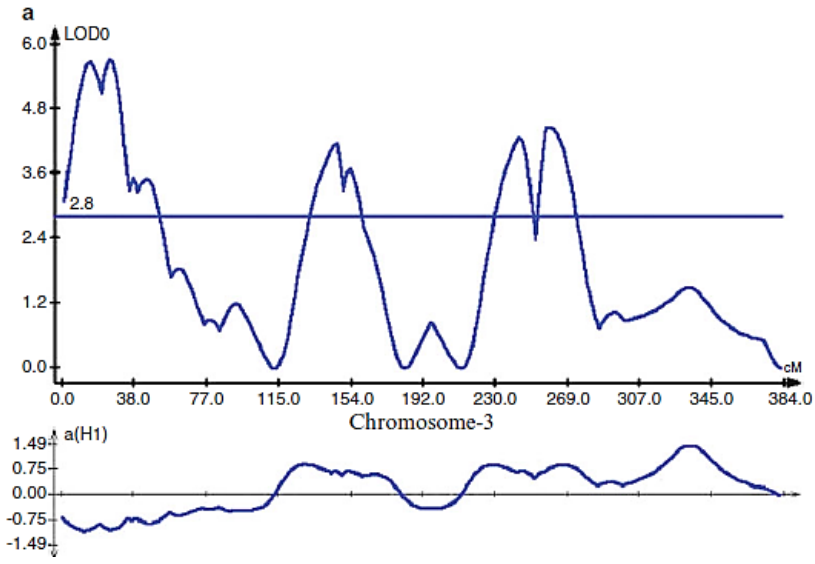
Composite interval mapping (CIM)

Simple interval method uses genome-wide scanning for detecting the position of a QTL throughout the genome. But this method can lead to biased estimates of QTL positions and effects when multiple QTLs occur on the interval of adjacent markers because it makes use of a single-marker interval between adjacent markers at a time and has no test has been performed to eliminate the effect of other QTLs outside the interval. Therefore, if a real QTL is located near a marker interval with no QTL, interval mapping may still detect a "ghost" QTL due to the linkage between the real QTL and the interval being tested (We et al., 2007). To overcome this problem, two almost identical methods namely "composite interval mapping" and "multiple-QTL model" or "marker-QTL- marker" (MQM) were proposed by Zeng (1994) and Jansen and Stam (1994), respectively. CIM combines interval mapping for a single

QTL in a given interval with multiple regression analysis on markers associated with other QTLs to control the effects of QTLs present in other marker intervals of the same or other chromosomes.

As CIM is an extension of SIM and uses some selected markers as cofactors (covariates) in the model to control for the genetic variation in other possibly linked and unlinked QTL based on the following model, $y_i = \mu + Z_i B + \sum_{r=1}^{m} X_{ir}\beta_r + e_i$, where yi is the ith individual phenotypic trait value; m μ is the overall mean; B is a column vector for the effects of a putative QTL, which depends on the type of mapping population; Zi is a row vector of predictor variables corresponding to the effects of the putative QTL; Xir is a row vector of predictor variables corresponding to the rth cofactor marker; βr is a column vector with the coefficient of the rth cofactor marker; and ei is the random error (Silva et al., 2012).

The main advantages of CIM are (1) By the search in one-dimension, multiple QTLs can be mapped, (2) By using linked markers as cofactors, the test is not affected by QTL outside the region, thereby the precision of QTL mapping is increased, (3) By eliminating much of genetic variance by other QTL, the residual variance is reduced and, consequently, the power of QTL detection is improved (Zeng, 1994). Due to these reasons, the CIM is more accurate and powerful than SIM in detecting QTLs (figure 3). However, CIM algorithm has some limitations such as (1) The different methods of cofactor selection, e.g., unlinked marker control, all marker control, and stepwise regression may produce different and sometimes contradictory results, and (2) In the presence of epistasis, CIM is inefficient because the method is unable to detect interacting QTLs.



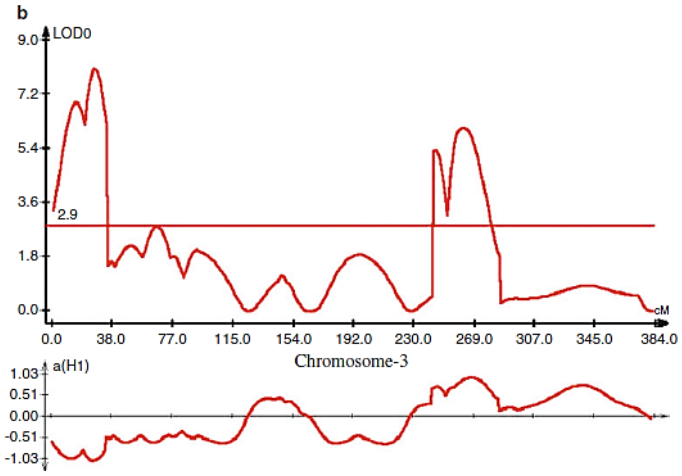


Figure 3. LOD score graphs obtained by (a) simple interval mapping (SIM) and (b) composite interval mapping (CIM) for 1,000-grain weight in rice (Adopted from Singh and Singh, 2015)

2. Bootstrap Method in QTL Mapping

One important problem in QTL mapping is large confidence intervals (CI) associated with QTL locations in segregation populations (Kearsey 1998a). The reliability depends on the heritability of individual QTL. Simulations have shown that the 95% CI of QTLs in an F_2 population of 300 individuals is more than 30 cM while it is difficult to reduce the CI to much less than 10 cM even for a very highly heritable QTL; more markers beyond a density of one every 15 Cm do not help much.

Several approaches have been explored to overcome this problem. A statistical method called ***'bootstrapping'*** is used to overcome this problem in QTL mapping. Bootstrap is a resampling method, which provides a very robust procedure for constructing CI for QTL position (Walling et al., 1998). It involves resampling points from one's data, with replacement, to create a series of bootstrap samples of the same size as the original data. Suppose the original data set consists of n individuals. A bootstrap sample is generated by drawing n values, with replacement, from the original data set. Such a sample will have some of the original values present multiple times and others not present at all. A series of N such samples are generated and an estimate (map position in this case) is computed for each, generating a distribution of estimates (the empirical distribution). The variation among the resulting estimates is taken to indicate the size of the error involved in estimating from original data. The resulting 95% bootstrap confidence interval has its lower

value the estimate corresponding to the 2.5% cumulative frequency point of the empirical bootstrap distribution, while the upper value is corresponding to the upper 97.5% 0f the bootstrap distribution. Simulation studies showed that this approach usually yields CI very close to the correct length when at least 200 bootstrap samples are used (Lynch & Walsh, 1998).

Software for QTL mapping

A number of freely available and commercial software packages are used for mapping QTLs in experimental populations. Hereby, some most commonly used and freely available software will be introduced.

WinQTL Cartographer is the most commonly used software for mapping QTL mapping which is available free at http://statgen.ncsu.edu/qtlcart/WQTLCart.htm (Wang et al., 2012). Data imports and exports can be performed in a variety of formats and empirical threshold LOD scores are calculated by permutation, and confidence intervals for QTL positions are estimated by the bootstrap method. Various types of QTL mapping methods including single-marker analysis, SIM, CIM, MIM with epistasis, Bayesian interval mapping, multiple trait analysis, and multiple trait MIM analysis are implemented in the software.

QTL IciMapping is freely available public software capable of building high-density linkage maps and mapping quantitative trait loci (QTL) in biparental populations (Meng et al., 2015). Multiple functionalities for linkage analysis are available in the software including segregation distortion analysis (SDL), construction of linkage maps in biparental populations (MAP), construction of consensus map from multiple linkage maps with common markers (CMP). The QTL mapping options such as mapping of additive, dominant, and digenic epistasis genes (BIP), mapping of additive and digenic epistasis genes with chromosome segment substitution lines (CSL), and QTL-by-environment interaction analysis (MET) are also implemented in *QTL IciMapping*. It can use marker and phenotypic data files in form of plain text or MS Excel formats. The outputs of software contain the summary of the completed linkage maps, Mendelian ratio test of individual markers, estimates of recombination frequencies, LOD scores, genetic distances, results at all scanning positions, identified QTL, permutation tests, and detection powers for up to six mapping methods.

R/qtl is an extensible, interactive environment for mapping QTLs in experimental populations derived from inbred lines crosses. It has been designed as an add-on program to the statistical language/software R, which is freely available from http://www.r-project.org. The single-QTL genome scans and two-dimensional, two-QTL genome scans, under a normal model, with the possible inclusion of covariates,

by the EM algorithm, Haley–Knott regression, and multiple imputation are various functions for QTL mapping in *R/qtl*. 2001). Further, non-parametric interval mapping for performing single-QTL genome scans as well as binary trait mapping are also implemented in R/qtl package (Broman et al., 2003).

Conclusions

QTL mapping studies have a long and rich history and have played important roles in plant and animal breeding as well as in gene cloning and characterization; however, there is still a great deal of work to be done. Through development of appropriate segregation populations and selection of effective statistical and molecular biological tools are essential for practical implementation of QTL mapping in applied sciences. Therefore, understanding the biological and statistical basis of methods applied for QTL mapping will enable breeders to determine the ideal genes and genotypes from QTL studies and utilizes them in breeding program.

References

Amoah, N.K.A., Akromah, R., Kena, A.W., Manneh, B., et al. (2020). Mapping QTLs for tolerance to salt stress at the early seedling stage in rice (*Oryza sativa* L.) using a newly identified donor 'Madina Koyo'. Euphytica, 216, 156. https://doi.org/10.1007/s10681-020-02689-5.

Azadi, A., Mardi, M., Majidi Hervan, E., Mohammadi, S.A., et al. (2015). QTL mapping of yield and yield components under normal and salt-stress conditions in bread wheat (*Triticum aestivum* L.). Plant Mol. Biol. Rep., 33:102-120. https://doi.org/10.1007/s11105-014-0726-0.

Badakhshan, H., Mohammadi, S.A., Aharizad, A., Moghaddam, M. JalalKamali, M.R. & Khodarahmi, M. (2008). Quantitative trait loci in bread wheat (*Triticum aestivum* L.). associated with resistance to stripe rust. Biotechnol. Biotechnol. Eq., 22: 901-906.

Barati, A., Moghaddam, M., Mohammadi, S.A., Ghazvini, H.A. & Sadeghzadeh, B. (2017). Identification of QTLs associated with agronomic and physiological traits under salinity stress in barley. J. Agr. Sci. Tech., 19: 185-200.

Boopathi, N.M. (2020). QTL analysis. In: Boopathi, N.M. (ed.). Genetic Mapping and Marker Assisted Selection: Basic, Practice and Benefits. Springer, p.p. 253-326.

Broman, K.W., Wu, H., Sen, Ś. & Churchill, G.A. (2003). R/qtl: QTL mapping in experimental crosses. Bioinformatics, 19: 889-890. https://doi.org/10.1093/bioinformatics/btg112.

Burr. B. & Burr, F.A. (1991). Recombinant inbreds for molecular mapping in maize: theoretical and practical considerations. Trends Genet., 7: 55-60.

Capo-Chichi, L.J.A., Eldridge, S., Elakhdar, A., Kubo, T., Brueggeman, R. & Anyia, A.O. (2021). QTL mapping and phenotypic variation for seedling vigour traits in barley (*Hordeum vulgare* L.). Plants (Basel). 10(6):1149. https://doi.org/10.3390/plants10061149.

Carbonell, E.A. & Asins, M.J. (1996). Statistical methods for the detection of genes

controlling quantitative trait loci expression, In Mohan Jain, S., Sopory, S.K. and Veilleux, R.E. (eds.) Current Plant Science and Biotechnology in Agriculture: In vitro Haploid Production in Higher Plants, Vol. 2, Kluwer Academic Publishers, pp: 255-285.

Doerge, R.W. (2002). Mapping and analysis of quantitative trait loci in experimental populations. Nat. Rev. Genet., 3: 43-52.

Du, B., Wang, Q., Sun, G., Ren, X., et al. (2019). Mapping dynamic QTL dissects the genetic architecture of grain size and grain filling rate at different grain-filling stages in barley. Sci. Rep., 9, 18823. https://doi.org/10.1038/s41598-019-53620-5.

Ehdaie, B., Mohammadi, S.A., Nouraein, M., Bektas, H. & Waines, J. G. (2016). QTLs for root traits at mid-tillering and for root and shoot traits at maturity in a RIL population of spring bread wheat grown under well-watered conditions. Euphytica, 211:39. https://doi.org/10.1007/s10681-016-1736-9B.

Guo, Y,. Zhang, G., Guo, B., Qu, C., et al. (2020). QTL mapping for quality traits using a high-density genetic map of wheat. PLoS ONE 15(3): e0230601. https://doi.org/10.1371/journal.pone.0230601.

Haley, C.S. & Knott., S.A. (1992). A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity, 69: 315-324.

Hussain, H., Rengel, Z., Mohammadi, S.A., Ebadi-Segherloo, A. & Maqsood, M.A. (2016). Mapping QTL associated with remobilization of zinc from vegetative tissues into grains of barley (*Hordeum vulgare*). Plant Soil, 399:193-208. https://doi.org/10.1007/s11104-015-2684-1.

Kearsey, M.J. & Farquhar, A.G.L. (1998a). QTL analysis in plants; where are we now? Heredity, 80: 137 – 142.

Kearsey, M.J. & Hyne, V. (1994). QTL analysis: a simple marker-regression approach. Theor. Appl. Genet., 89; 698-702.

Lander, E.S. & Botstein, D. (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. Genetics, 121: 185-199.

Li, B., Zhang, Y., Mohammadi, S.A., Huai, D., Zhou, Y. & Kliebenstein, D.J. (2016). An integrative genetic study of rice metabolism, growth and stochastic variation reveals potential C/N partitioning loci. Sci. Rep., 6:30143. https://doi.org/10.1038/srep30143.

Li, X., Zheng, H., Wu, W., Liu, H., et al. (2020). QTL mapping and candidate gene analysis for alkali tolerance in japonica rice at the bud stage based on linkage mapping and genome-wide association study. Rice, 13, 48. https://doi.org/10.1186/s12284-020-00412-5.

Liu, B. (1998). Statistical Genomics: Linkage, Mapping and QTL Analysis. CRC Press, Boca Raton.

Lynch, M. & Walsh, B. (1998). Genetics and Analysis of Quantitative Traits. Sinauer Associates, Inc. Massachusetts.

Meng, L., Li, H., Zhang, L. & Wang, J. (2015). QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. Crop J., 3: 269-283. https://doi.org/10.1016/j.cj.2015.01.001.

Muehlbauer, G.J., Specht, J.E., Thomas-Compton, M.A., Staswick, P.E. & Bernard, R.L. (1988). Near isogenic lines – a potential resource in the integration of conventional

and marker linkage maps. Crop Sci., 28: 729-735.

Silva, L.D.C.E., Wang, S. & Zeng, Z.B. (2012). Composite interval mapping and multiple interval mapping: Procedures and guidelines for using windows QTL cartographer. In: Rifkin, S.A. (ed.), Quantitative Trait Loci (QTL): Methods and Protocols, Methods in Molecular Biology. Springer, New York, p.p. 75-119.

Singh, B.D. & Singh, A.K. (2015a). Mapping population. In: Singh, B.D., and Singh, A.K. (eds.) Marker-Assisted Plant Breeding: Principles and Practices. Springer, pp. 125-150.

Singh, B.D. & Singh, A.K. (2015b). Mapping of quantitative trait loci. In: Singh, B.D., and Singh, A.K. (eds.) Marker-Assisted Plant Breeding: Principles and Practices, Springer, pp. 185-216.

Tanksley, S.D. (1993). Mapping polygenes. Annu. Rev. Genet., 27: 205 – 233.

Walling, G.A., Visscher, P.M. & Haley, C.S. (1998). A comparison of bootstrap methods to construct confidence intervals in QTL mapping. Genet. Res., 71: 171-180.

Wang, S., Basten, C.J. & Zeng Z-B. (2012). Windows QTL Cartographer 2.5_011. Department of Statistics, North Carolina State University, Raleigh.

Wang, Y., Xu, X., Hao, Y., Zhang, Y., et al. (2021). QTL mapping for grain zinc and iron concentrations in bread wheat. Front. Nutr., 2 9;8:680391. https://doi.org/10.3389/fnut.2021.680391.

We, R., Casella, G. & Ma, C.X. (2007). Composite QTL mapping. In: We, R., Casella, G., and Ma, C.X. (eds.). Statistical Genetics of Quantitative Traits. Statistics for Biology and Health. Springer, New York, NY. https://doi.org/10.1007/978-0-387-68154-2_13.

Wu, W.R. & Li, W.M. (1996). Model fitting and model testing in the method of joint mapping of quantitative trait loci. Theor. Appl. Genet., 92:477-482.

Xiao, J., Grandillo, S., Ahn, S.N., McCouch, A.R., Tanksley, S.D., Li, J. & Yuan, L. (1996). Genes from wild rice improve yield. Nature, 384: 223-224.

Xu, Y., Li, P., Yang, Z. & Xu, C. (2107). Genetic mapping of quantitative trait loci in crops. Crop J., 5: 175-184. https://doi.org/10.1016/j.cj.2016.06.003.

Young, N.D. (1996). QTL mapping and quantitative disease resistance in plants. Annu. Rev. Phytopathol., 34: 479–501.

Zeng, Z.B. (1994). Precision mapping of quantitative trait loci. Genetics, 136: 1457-1468.

Zeng, Z.B., Kao C.H. & Basten, C.J. (1999). Estimating the genetic architecture of quantitative traits. Genet. Res., 74: 279-289.

Khazar Journal of Science and Technology Volume 5 №2 2021, 20-33

DOI: 10.5782/2520-6133.2021.5.2.20

Epigenetic Modifications: Basic Mechanisms in Normal and Cancerous Cells

Morvarid Soleiman
Department of Life Sciences, Khazar University, Baku, Azerbaijan

Abstract

Epigenetics is one of the most rapidly expanding fields in biology that refers to the somatically heritable differences in gene expression that are not coded in the DNA sequence itself and work with genetic mechanisms to determine transcriptional activity. Aberrant epigenetic modifications including, DNA methylation, histone modifications, and small noncoding microRNAs (miRNA), probably occur at a very early stage in neoplastic development, and they are widely described as important players in cancer progression. However, the reversible nature of epigenetic alternations has encouraged the development of pharmacologic inhibitors as anti-cancer therapeutics. In this case, histone deacetylase inhibitors and DNA methylation inhibitors have been FDA-approved for several years and are clinically successful. In this article, we review the mechanisms of each epigenetic modification in both normal and tumor cells. Also, the potential of epigenetic alternations as a new emerging target in cancer therapy has been discussed.

Keywords: Epigenetics, Epigenetic modifications, Cancer, Epigenetic therapy

Introduction

The term ‘epigenetics’ defines all heritable changes in gene expression and chromatin organization that is not coded in the DNA sequence itself (Egger et al., 2004). Recent progresses have highlighted the key role of epigenetic mechanisms in ensuring the appropriate control of biological processes, such as imprinting, X chromosome inactivation, or the establishment and maintenance of cell identity (Altucci & Minucci, 2009). Epigenetic inheritance is an essential mechanism that allows the stable propagation of gene activity states from one generation of cells to the next (Herceg, 2007). Methylation of DNA, chemical modification of the histone proteins, and RNA-dependent regulation have been investigated as important mechanisms of epigenetic regulation.

Traditionally, cancer has been viewed as a genetic disease, and it is now becoming apparent that the onset of cancer is preceded by epigenetic abnormalities. All critical changes in every aspect of tumor biologies such as cell growth and differentiation, cell cycle control, DNA repair, angiogenesis, and migration, are caused not only by genetic but also by epigenetic mechanisms (Sharma et al., 2010). The genetic origin of cancer is widely accepted; however, recent studies suggest that epigenetic alterations may be the key initiating events in some forms of cancer. These findings have led to a global initiative to understand the role of epigenetics in the initiation and propagation of cancer.

In addition, epigenetic alternations, unlike genetic mutations, with changeable nature can be restored to their normal state (Jones & Martienssen, 2005; Yoo & Jones, 2006). They are reversible by pharmacological manipulation of the enzymes responsible for chromatin modification: indeed, epigenetic drugs (histone deacetylase inhibitors and DNA demethylating agents) are currently on the market, inducing proliferative arrest and death of tumor cells (Altucci & Minucci, 2009). In this review, we take a comprehensive look at the current epigenetic mechanisms in normal cells and their comparative aberrations that occur during carcinogenesis. We also discuss the great potential lies in the development of epigenetic therapies for some malignancies.

1- Epigenetic features

Chromatin is organized by repeating units of nucleosomes, which consist of 145_147 base pairs of DNA wrapped around an octamer of two copies of each histone protein (H3, H4, H2A, and H2B) (Luger et al., 1997). Despite the stability of the nucleosome and the high degree of compactness of the nucleus, chromatin is astonishingly dynamic. Epigenetic modifications include DNA methylation, histone modifications, and non-coding RNAs such as microRNAs (miRNAs) can be modified chromatin structure. The N-terminal tails of the histone proteins are protruding out from the nucleosomal core particles, and these tails serve as regulatory registers onto which epigenetic signals can be written. (Lund & van Lohuizen, 2004). Due to their impact on the genome, epigenetic modifications are involved in the regulation of many pathways including apoptosis, cell proliferation, and differentiation (Miceli et al., 2014; Sharma et al., 2010). Additionally, epigenetic aberrations, including global hypomethylation and miRNA deregulation, as well as promoter hypermethylation and deacetylation have been displayed in many human cancer types (Taby & Issa, 2010).

1.1. DNA methylation

DNA methylation was the first epigenetic modification found in humans in the early 1980s (Cooper, 1983). DNA methylation, the covalent addition of a methyl group to the cytosine base in DNA, has been recognized as a critical regulatory mechanism during development, cellular differentiation, and tissue homeostasis. It has been connected to different physiological and pathological processes, including genomic imprinting, X chromosome inactivation in females, tissue-specific gene expression, chromosome stability, and a number of abnormalities, including cancer (Bernstein et al., 2007; Hathaway et al., 2012; Kanwal & Gupta, 2010). In humans, DNA methylation occurs in the C^5 position of cytosines that precede guanines (CpGs) which are catalyzed by DNA methyltransferases (DNMTs) using S-adenyl methionine (SAM) as the methyl donor (Herman & Baylin, 2003; Ohm et al., 2007). Three enzymes are involved in the generation and maintenance of DNA methylation patterns. DNMT1 has a strong preference for hemimethylated CpG dinucleotides, thus can methylate CpG in a newly synthesized DNA strand based on the presence of methylation in the complementary template (Roberti et al., 2019) and Dnmt3a and Dnmt3b, defined as a de novo methyltransferase, show no preference for hemimethylated target sites and are present mainly in the early development stages and germ cells, whereas they are largely suppressed in the differentiated somatic cells (Wang & Shen, 2004) The majority of CpG dinucleotide are concentrated within CpG-rich DNA regions known as "CpG islands" that have been evolutionarily conserved to promote gene expression by regulating the chromatin structure and transcription factor binding (Bennett & Licht, 2018; Ramirez-Carrozzi et al., 2009). The methylation of CpG islands results in stable silencing of gene expression. During gametogenesis and early embryonic development, CpG islands undergo differential methylation and regulate gene expression during development and differentiation. As CpG islands are associated with the control of gene expression, it would be expected that CpG islands might display tissue-specific patterns of DNA methylation (Meissner et al., 2008; Moore et al., 2013). Fully methylated CpG islands in the silenced allele for specific imprinted autosomal genes and multiple silenced genes on the female inactivated X chromosome and deacetylation of histone proteins is the first step in the recruitment of methyltransferase to the CpG islands, resulting in hypermethylation of the promoter, which are two critical roles of DNA methylation that have been reported (Verma & Srivastava, 2002).

A vast amount of knowledge has been gained in the last few years about altered methylation patterns, both hyper- and hypo-methylation, in many different types of cancers including prostate, breast, gastric, liver, lung, glioblastoma, and leukemia (Park & Han, 2019). There are two types of general changes in DNA methylation that occur in tumor cells in comparison to normal cells of the same tissue type:

demethylation within many regions of the genome in coordination with de novo methylation of select CpG islands. Much of the hypomethylation is concentrated within broad late-replicating Lamin-associated domains that make up about 40% of the genome and contain many repetitive sequences. More striking is hypermethylation of a wide range of CpG islands that are usually unmethylated in every tissue that this occurs widely at promoters of tumor suppressor genes that cause uncontrolled growth cancerous cells (Klutstein et al., 2016; Siegfried et al., 1999). Tumor-specific methylation changes in different genes have been identified and documented. Despite no evidence of clearly identified actors in DNA demethylation, alteration of global DNA methylation patterns in cancer is often associated with an over-expression of DNMTs as described in various tumors (Akhavan-Niaki & Samadani, 2013). The exact degree of overexpression of DNMTs in tumors remains unclear but a low-level over-expression seems to be common (Delpu et al., 2013). Current studies have shown that the duplication of the DNMT3b gene in different cancer cell lines where copy number correlates to increased mRNA and protein levels, also, DNMT3b over-expression occurs through the stabilization of its mRNA in human colorectal carcinoma (López de Silanes et al., 2009; Simó-Riudalbas et al., 2011)

1.2. Histon Modifications

Histones are highly conserved small basic proteins that are found exclusively in eukaryotic cells, predominantly in the nucleus. Histones –and especially their N-terminal "tails" are recognized as being dynamic regulators of gene activity that undergo many post-translational chemical modifications, including acetylation, methylation, phosphorylation, ubiquitylation, and sumoylation. Histone modifications have been linked to a number of chromatin-dependent processes, including X chromosome inactivation, genome stability, and meiotic chromosome dynamics (Esteller, 2007; Goll & Bestor, 2002; Xhemalce et al., 2006). Given the number of sites and the variety of possible modifications, the combinatorial possibilities are extremely large and it is tempting to believe that histone modification has a regulatory role (Lagger et al., 2002). It has been found that individual modifications can be associated with transcriptional activation or repression. Acetylation and phosphorylation generally accompany transcription, methylation, and ubiquitination are implicated in both activation and repression of transcription (Karlić et al., 2010).

1.2.1. Histon acetylation

Among all histone modifications, the status of acetylation and methylation of specific lysine residues have a crucial role in regulating chromatin structure and gene

expression (Esteller, 2007). Histone acetylation is catalyzed by a number of enzymes that act on the four core histones and transfer the acetyl moiety from Acetyl-CoA to the ε-amino group of lysine residues (HATs). Histone acetylation is generally associated with transcription activation due to destabilizing nucleosomes and promoting both nucleosomal rearrangements of chromatin remodeling complexes and binding of a diverse set of DNA-binding factors involved in transcription, DNA repair, and other processes (Fuchs et al., 2009). On the other hand, histone deacetylase enzymes (HDACs) are associated with nucleosome stabilization and repression of remodeling activities so, are generally thought to act in a repressive manner (Kurdistani & Grunstein, 2003).

More recently, an increasing number of disease processes have been observed to involve abnormalities of acetylating and deacetylating cellular events. Recent studies have been illustrated that a decrease in the amount of functionally available HDACs or an increase of functionally active HAT enzymatic activities affects the conformation and activity of associated transcription factors. As a consequence, genes that were previously silenced may now be activated or even overexpressed, whereas other genes, which were previously expressed, may now secondarily be repressed (Mahlknecht & Hoelzer, 2000)

1.2.2. Histone Methylation

Histone methylation, perhaps more than any other form of modification, has demonstrated an essential role in diverse biological processes ranging from transcriptional regulation to heterochromatin formation (Bannister & Kouzarides, 2005). Histones may be methylated on lysine (K), arginine (R), and/or histidine (H) residues using S-adenosylmethionine (SAM or AdoMet) as a methyl group donor. At present, there are 24 known sites of methylation on histones which 17 are lysine residues and 7 are arginine residues (Kim et al., 2014; Margueron et al., 2005). While the methylation of histidine is one of the rare histone modifications and has not been well characterized (Greer & Shi, 2012), Lysine methylation of histones is a remarkable and complex epigenetic mark that makes up both transcriptionally silenced and active chromatin domains, depending on which lysine residues are methylated and the degree of methylation (Greer & Shi, 2012). Lysine methylation has three methyl additions including mono-(H3K9me1), di-(H3K9me2), and tri-(H3K9me3) that each level of methylation produces different outcomes (Martin & Zhang, 2005). Among different types of lysin methylation including histone H3 lysine 4 (H3K4), H3K9, H3K27, H3K36, H3K79, and H4K20, histone H3K9 and H3K27 methylation is associated with silenced regions, whereas H3K4 and H3K36 methylation is correlated well with active genes (Berger, 2007; Tan et al., 2011). Arginine is another residue that can be mono(me1) symmetrically dimethylated

(me2s), or asymmetrically dimethylated(me2a) on their guanidinyl group (Greer & Shi, 2012).

The steady-state level of a covalent histone modification is controlled by a balance between enzymes that catalyze the addition and removal of a given modification. Although this notion is generally true for many histone modifications, an enzyme capable of removing methyl groups from a methyl-lysine residue has remained elusive, until recently (Tsukada et al., 2006). Histone demethylases, first described by Shi et al, have the opposite effect on transcription. The histone demethylase LSD1 is responsible for H3K4 demethylation, which leads to transcriptional inactivation. However, when LSD1 forms a complex with androgen receptors, it demethylates H3K9 and activates transcription. Other histone demethylases, such as JHDM1, can convert active chromatin marks H3K36me2 to an unmodified state (Bártová et al., 2008).

Aberrant histone modifications are known to play a key role in the pathogenesis of several human diseases such as cancer (Shanmugam et al., 2018). Promoters carry on important regulatory sequences for transcription control of their genes. Generally, the vast majority of deregulation of histone modifications at an individual promoter is intimately linked to misexpression of the downstream gene, which may have critical consequences for the cancer phenotype (Kurdistani, 2011). In this case, loss of trimethylation of H4K20 was identified as one of the common hallmarks of human cancers (Huang et al., 2017). Also, high levels of H3K27me2/3, and H3K79me3, as well as modest levels of H3K9me2/ 3 are linked to gene repression or silencing (Kanwal & Gupta, 2010). In addition, altered global levels of histone acetylation, particularly acetylation of H4K16, have been linked to a cancer phenotype in a variety of cancers. While gene expression especially in proto-oncogene can be activated by hyperacetylation, hypoacetylation of tumor suppressors often localizes to promoters, co-occurring with DNA methylation, causing the genes to be silenced (Audia & Campbell, 2016).

1.3. MicroRNAs

MicroRNAs (miRNA) comprise a class of short non-coding RNAs with 18–25 nucleotides in length that can have a profound effect in controlling gene expression post-transcriptionally (Chuang and Jones, 2007). miRNAs are involved in RNA interference (RNAi) machinery and they contribute to diverse physiological and pathophysiological functions, including the regulation of developmental timing and pattern formation, restriction of differentiation potential, cell signaling, and carcinogenesis (Sato et al., 2011). The transcription of most miRNA genes is mediated by RNA polymerase II (Pol II) and the primary transcripts (pri-miRNAs)

are usually several kilobases long and contain local stem-loop structures and are then processed in the nucleus by the RNase III Drosha and DGCR8 into the precursor miRNAs (pre-miRNAs) (Chuang and Jones, 2007, Kim et al., 2009). Pre-miRNAs are then processed by the RNAse III Dicer to generate a double-stranded (ds)RNA approximately 22 nucleotides long. After generation of the miRNA duplex, one strand (the miRNA-guide strand) is loaded onto the RNA-induced silencing complex (RISC) that is able to modulate the expression of target protein-coding mRNAs by base-pairing to partially complementary regions frequently located at the 3'-untranslated regions (3'-UTR) of the target transcript (Malumbres, 2013; Romero-Cordoba et al., 2014). It has been investigated that an individual miRNA can dramatically control the expression of more than one target mRNAs and that each mRNA may be regulated by multiple miRNAs (Cai et al., 2009). Several studies have reported links between aberrant miRNA expression and several aspects of cellular function from proliferation and differentiation to apoptosis (Ha, 2011).

In 2002, the role of miRNAs in cancer was suggested by Croce and colleagues with the discovery that regions of miR-15 and miR-16 frequently deleted in chronic lymphocytic
leukemia (CLL). Recently, deregulation of miRNAs has been definitively linked to the initiation and progression of tumorigenesis and the role-play of miRNAs have been investigated in many types of human cancer, including breast, colon, gastric, lung, prostate, and thyroid (Ha, 2011; Hatziapostolou & Iliopoulos, 2011; Reddy, 2015). Recent researches have been illustrated that in different types of cancers, particular miRNAs may show oncogenic or tumor-suppressive function. For instance, miR-29 was reported as an oncogene in breast cancer while acting as a tumor-suppressor gene in lung tumors. Additionally, loss of miRNA-23b caused migration and invasion in bladder cancer cells while decreasing the expression of miRNA-23b induced apoptosis and reduced invasive capabilities in renal cell carcinoma cell lines (Campos-Viguri et al., 2015; Reddy, 2015; Zaman et al., 2012).

In addition, miRNAs can both regulate and be regulated by other epigenetic mechanisms. In this case, dysregulation of miR-101 leads to reduced H3K27me3 and inhibits cancer cell proliferation. Expression of miR-143 in colorectal cancer cells and the miR-29 family in lung cancer cells reduces DNMT3A and DNMT3B levels, respectively, and consequently, cell growth and colony formation were decreased (Fabbri et al., 2007; Kelly et al., 2010; Ng et al., 2009).

Recently, miRNAs have been recommended as epigenetic biomarkers in the diagnosis of cancer. For example, miR-199a, miR-200a, miR-146, miR-214, miR-221, and miR-222 have been investigated to be overexpressed, whereas miR-100 is down-regulated in human cancers. Furthermore, 27 miRNAs are noticeably related

to chemotherapy response and consider as possible prognostic and diagnostic biomarkers (Kanwal & Gupta, 2012; Paranjape et al., 2009; Peter, 2009).

1.4. Epigenetic Therapy

Unlike stable genetic mutations, epigenetic modifications like DNA methylation and histone modifications have reversible nature (Ali et al., 2015). This fundamental difference between genetic and epigenetic alterations makes the epigenome much more amenable to the development of therapeutic strategies (Altucci & Minucci, 2009). The so-called epi-drugs are an elevating exciting field in anticancer research and therapy (Miceli et al., 2014). Epigenetic modifier drugs capable of reversing aberrant DNA methylation and histone acetylation patterns by inhibiting DNMTs and HDACs have been extensively discovered (Kristensen et al., 2009). The first successful drugs developed as epigenetic agents were DNA methyltransferase inhibitors; these were followed by histone deacetylase inhibitors (HDIs) (Yanis Boumber & Issa, 2011). DNA demethylating drugs, 5-azacytidine (Vidaza), and 5-aza-2′- deoxycytidine (decitabine) have been approved in 2004 and 2006 for myelodysplastic syndrome and leukemia (Yang et al., 2006). These small molecules are analogues of cytidine and they inhibit DNA methylation by the irreversible inhibition of DNMTs, causing hypomethylation of DNA, and reactivation of silenced genes (Platzbecker et al., 2012). DNMT inhibitors can be subdivided into three groups, nucleoside analogue, non-nucleoside analogue, and antisense oligonucleotides, based on their structures and functions (Peedicayil, 2006). While the first group of DNMT inhibitors is, S-phase specific drugs, phosphorylated to the deoxynucleotide triphosphate and then incorporated instead of cytosine into replicating DNA and inhibit DNMTs (Egger et al., 2004), the non-nucleoside analogue like RG108 is DNMT inhibitor that designed to target human DNMT1 at its active site without high levels of cytotoxicity and affecting the methylation status of centromeric repeats (Graça et al., 2014). Antisense oligonucleotides as a third group of DNMT inhibitors can block translation by inactivation of mRNA through complementary hybridization (Yan et al., 2003).

Histone deacetylase (HDAC) inhibitors belong to a group of small-molecule drugs that induce a broad range of effects on cancer cells, including cell cycle arrest, apoptosis, cell differentiation, autophagy, and anti-angiogenic effects (Khan & La Thangue, 2012). Targeting HDACs is more complex than targeting DNA methyltransferases because this group of proteins has multiple subclasses with mechanisms of action still under contention (Azad et al., 2013). However, the HDAC inhibitor suberoylanilide hydroxamic acid (SAHA) was approved in 2006 for the treatment of persistent or cutaneous T cell lymphoma (Moufarrij et al., 2019).

Recently, regarding the parallel function of DNA methylation and histone modifications, is now being paid to testing combinations of drugs, which may increase the efficacy of each of the single agents. For example, the combination of HDACi and DNMTi can elevate the expression levels of tumor suppressor genes (Jones et al., 2019). Furthermore, combining conventional cancer therapies including chemotherapy and radiotherapy with the use of epigenetic therapy, reversing the changes of DNA methylation and histone acetylation patterns, holds a huge potential for successful treatment of hematological malignancies as well as solid tumors that may allow for lower dosing which can minimize side effects of treatment improving quality of life and treatment compliance (Kelly et al., 2010; Kristensen et al., 2009).

However, deeper understandings of the global patterns of these epigenetic modifications and their corresponding changes in cancer have enabled the design of better treatment strategies.

Conclusion

The recognition of epigenetics as a significant contributor to normal development and disease has opened new avenues for drug discovery and therapeutics. In this paper, we have summarized the complex epigenetic regulatory pathways in normal cells. Also, the key role of DNA methylation, histone modification, and miRNAs as epigenetic modifications in cancer progression and therapeutics were reviewed. Based on the above, epigenetic alterations in comparison with genetic changes are reversible and are typically acquired in a gradual manner. Given that the epigenetic changes induced by DNMTs and HDACi are transient and reversible, a number of studies are ongoing to help define the optimal doses and treatment schedules for these agents. Besides their methylation and acetylation, histones can be phosphorylated, ubiquitylated, and sumoylated. These modifications, which have been less well studied in the context of disease, may expand current possibilities for therapeutic intervention. So, the integration of the latest advances in epigenomic approaches like whole-genome microarray expression profiling and chromatin immunoprecipitation-based sequencing (ChIP-seq) methods will be allow mapping of all types of histones and DNA modification's state and miRNA levels in the genome with high accuracy, which will be helpful in the identification of biomarkers and the development of epigenetic drugs with greater specificity.

References

Akhavan-Niaki, H. & Samadani, A. A. (2013). DNA methylation and cancer development: molecular mechanism. Cell biochemistry and biophysics, 67, 501-513.

Ali, M., Hanif, M. & Farooqi, A. A. (2015). Epigenetic therapy for cancer. Pak. J. Pharm. Sci, 28**,** 1023-1032.

Altucci, L. & Minucci, S. (2009). Epigenetic therapies in haematological malignancies: searching for true targets. European Journal of Cancer, 45, 1137-1145.

Audia, J. E. & Campbell, R. M. (2016). Histone modifications and cancer. Cold Spring Harbor perspectives in biology, 8, a019521.

Azad, N., Zahnow, C. A., Rudin, C. M. & Baylin, S. B. (2013). The future of epigenetic therapy in solid tumours—lessons from the past. Nature reviews Clinical oncology, 10**,** 256-266.

Bannister, A. J. & Kouzarides, T. (2005). Reversing histone methylation. Nature, 436, 1103-1106.

Bártová, E., Krejčí, J., Harničarová, A., Galiová, G. & Kozubek, S. (2008). Histone modifications and nuclear architecture: a review. Journal of Histochemistry & Cytochemistry, 56, 711-721.

Bennett, R. L. & Licht, J. D. (2018). Targeting epigenetics in cancer. Annual review of pharmacology and toxicology, 58, 187-207.

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. Nature, 447, 407-412.

Bernstein, B. E., Meissner, A. & Lander, E. S. (2007). The mammalian epigenome. Cell, 128**,** 669-681.

Cai, Y., Yu, X., Hu, S. & Yu, J. (2009). A brief review on the mechanisms of miRNA regulation. Genomics, proteomics & bioinformatics*,* 7, 147-154.

Campos-Viguri, G. E., Jiménez-Wences, H., Peralta-Zaragoza, O., Torres-Altamirano, G., Soto-Flores, D. G., Hernández-Sotelo, D., Alarcón-Romero, L. D. C., Jiménez-López, M. A., Illades-Aguiar, B. & Fernández-Tilapa, G. (2015). miR-23b as a potential tumor suppressor and its regulation by DNA methylation in cervical cancer. Infectious agents and cancer, 10, 1-10.

Chuang, J. C. & Jones, P. A. (2007). Epigenetics and microRNAs. Pediatric research, 61**,** 24-29.

Cooper, D. N. (1983). Eukaryotic DNA methylation. Human genetics, 64, 315-333.

Delpu, Y., Cordelier, P., Cho, W. C. & Torrisani, J. (2013). DNA methylation and cancer diagnosis. International journal of molecular sciences, 14, 15029-15058.

Egger, G., Liang, G., Aparicio, A. & Jones, P. A. (2004). Epigenetics in human disease and prospects for epigenetic therapy. Nature, 429, 457-463.

Esteller, M. (2007). Cancer epigenomics: DNA methylomes and histone-modification maps. *Nature reviews genetics,* 8, 286-298.

Fabbri, M., Garzon, R., Cimmino, A., Liu, Z., Zanesi, N., Callegari, E., Liu, S., Alder, H., Costinean, S. & Fernandez-Cymering, C. (2007). MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. Proceedings of the National Academy of Sciences, 104, 15805-15810.

Fuchs, S. M., Laribee, R. N. & Strahl, B. D. (2009). Protein modifications in transcription elongation. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 1789, 26-36.

Goll, M. G. & Bestor, T. H. (2002). Histone modification and replacement in chromatin activation. Genes & development, 16, 1739-1742.

Graça, I., J Sousa, E., Baptista, T., Almeida, M., Ramalho-Carvalho, J., Palmeira, C., Henrique, R. & Jerónimo, C. (2014). Anti-tumoral effect of the non-nucleoside DNMT inhibitor RG108 in human prostate cancer cells. Current pharmaceutical design, 20, 1803-1811.

Greer, E. L. & Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. Nature Reviews Genetics, 13, 343-357.

Ha, T.-Y. (2011). MicroRNAs in human diseases: from cancer to cardiovascular disease. Immune network, 11, 135-154.

Hathaway, N. A., Bell, O., Hodges, C., Miller, E. L., Neel, D. S. & Crabtree, G. R. (2012). Dynamics and memory of heterochromatin in living cells. Cell, 149, 1447-1460.

Hatziapostolou, M. & Iliopoulos, D. (2011). Epigenetic aberrations during oncogenesis. Cellular and Molecular Life Sciences, 68, 1681-1702.

Herceg, Z. (2007). Epigenetics and cancer: towards an evaluation of the impact of environmental and dietary factors. Mutagenesis, 22, 91-103.

Herman, J. G. & Baylin, S. B. (2003). Gene silencing in cancer in association with promoter hypermethylation. New England Journal of Medicine, 349, 2042-2054.

Huang, T., Lin, C., Zhong, L. L., Zhao, L., Zhang, G., Lu, A., Wu, J. & Bian, Z. (2017). Targeting histone methylation for colorectal cancer. Therapeutic advances in gastroenterology, 10, 114-131.

Jones, P. A. & Martienssen, R. (2005). A blueprint for a human epigenome project: the AACR human epigenome workshop. Cancer research, 65, 11241-11246.

Jones, P. A., Ohtani, H., Chakravarthy, A. & De Carvalho, D. D. (2019). Epigenetic therapy in immune-oncology. Nature Reviews Cancer, 19, 151-161.

Kanwal, R. & Gupta, S. (2010). Epigenetics and cancer. Journal of applied physiology, 109, 598-605.

Kanwal, R. & Gupta, S. (2012). Epigenetic modifications in cancer. Clinical genetics, 81, 303-311.

Karlić, R., Chung, H.-R., Lasserre, J., Vlahoviček, K. & Vingron, M. (2010). Histone modification levels are predictive for gene expression. Proceedings of the National Academy of Sciences, 107, 2926-2931.

Kelly, T. K., De Carvalho, D. D. & Jones, P. A. (2010). Epigenetic modifications as therapeutic targets. Nature biotechnology, 28, 1069-1078.

Khan, O. & La Thangue, N. B. (2012). HDAC inhibitors in cancer biology: emerging mechanisms and clinical applications. Immunology and cell biology, 90, 85-94.

Kim, V. N., Han, J. & Siomi, M. C. (2009). Biogenesis of small RNAs in animals. Nature reviews Molecular cell biology, 10, 126-139.

Kim, W., Choi, M. & Kim, J.-E. (2014). The histone methyltransferase Dot1/DOT1L as a critical regulator of the cell cycle. Cell Cycle, 13, 726-738.

Klutstein, M., Nejman, D., Greenfield, R. & Cedar, H. (2016). DNA methylation in cancer and aging. Cancer research, 76, 3446-3450.

Kristensen, L. S., Nielsen, H. M. & Hansen, L. L. (2009). Epigenetics and cancer treatment. European journal of pharmacology, 625, 131-142.

Kurdistani, S. K. (2011). Histone modifications in cancer biology and prognosis. Epigenetics and Disease, 91-106.

Kurdistani, S. K. & Grunstein, M. (2003). Histone acetylation and deacetylation in yeast. Nature reviews Molecular cell biology, 4, 276-284.

Lagger, G., O'carroll, D., Rembold, M., Khier, H., Tischler, J., Weitzer, G., Schuettengruber, B., Hauser, C., Brunmeir, R. & Jenuwein, T. (2002). Essential function of histone deacetylase 1 in proliferation control and CDK inhibitor repression. The EMBO journal, 21, 2672-2681.

López De Silanes, I., Gorospe, M., Taniguchi, H., Abdelmohsen, K., Srikantan, S., Alaminos, M., Berdasco, M., Urdinguio, R. G., Fraga, M. F. & Jacinto, F. V. (2009). The RNA-binding protein HuR regulates DNA methylation through stabilization of DNMT3b mRNA. Nucleic acids research, 37, 2658-2671.

Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature, 389**,** 251-260.

Lund, A. H. & Van Lohuizen, M. (2004). Epigenetics and cancer. Genes & development, 18, 2315-2335.

Mahlknecht, U. & Hoelzer, D. (2000). Histone acetylation modifiers in the pathogenesis of malignant disease. Molecular Medicine, 6, 623-644.

Malumbres, M. (2013). miRNAs and cancer: an epigenetics view. Molecular aspects of medicine, 34, 863-874.

Margueron, R., Trojer, P. & Reinberg, D. (2005). The key to development: interpreting the histone code? Current opinion in genetics & development, 15**,** 163-176.

Martin, C. & Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nature reviews Molecular cell biology, 6, 838-849.

Meissner, A., Mikkelsen, T. S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B. E., Nusbaum, C. & Jaffe, D. B. (2008). Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature, 454, 766-770.

Miceli, M., Bontempo, P., Nebbioso, A. & Altucci, L. (2014). Natural compounds in epigenetics: A current view. Food and chemical toxicology, 73, 71-83.

Moore, L. D., Le, T. & Fan, G. (2013). DNA methylation and its basic function. Neuropsychopharmacology, 38, 23-38.

Moufarrij, S., Dandapani, M., Arthofer, E., Gomez, S., Srivastava, A., Lopez-Acevedo, M., Villagra, A. & Chiappinelli, K. B. (2019). Epigenetic therapy for ovarian cancer: promise and progress. Clinical epigenetics, 11, 1-11.

Ng, E., Tsang, W., Ng, S., Jin, H., Yu, J., Li, J., Röcken, C., Ebert, M., Kwok, T. & Sung, J. (2009). MicroRNA-143 targets DNA methyltransferases 3A in colorectal cancer. British journal of cancer, 101, 699-706.

Ohm, J. E., Mcgarvey, K. M., Yu, X., Cheng, L., Schuebel, K. E., Cope, L., Mohammad, H. P., Chen, W., Daniel, V. C. & Yu, W. (2007). A stem cell–like chromatin pattern

may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nature genetics, 39, 237-242.

Paranjape, T., Slack, F. & Weidhaas, J. (2009). MicroRNAs: tools for cancer diagnostics. Gut, 58, 1546-1554.

Park, J. W. & Han, J.-W. (2019). Targeting epigenetics for cancer therapy. Archives of pharmacal research, 42, 159-170.

Peedicayil, J. (2006). Epigenetic therapy-a new development in pharmacology. Indian Journal of Medical Research, 123, 17.

Peter, M. E. (2009). Let-7 and miR-200 microRNAs: guardians against pluripotency and cancer progression. Cell cycle, 8, 843-852.

Platzbecker, U., Wermke, M., Radke, J., Oelschlaegel, U., Seltmann, F., Kiani, A., Klut, I., Knoth, H., Röllig, C. & Schetelig, J. (2012). Azacitidine for treatment of imminent relapse in MDS or AML patients after allogeneic HSCT: results of the RELAZA trial. Leukemia, 26, 381-389.

Ramirez-Carrozzi, V. R., Braas, D., Bhatt, D. M., Cheng, C. S., Hong, C., Doty, K. R., Black, J. C., Hoffmann, A., Carey, M. & Smale, S. T. (2009). A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. Cell, 138, 114-128.

Reddy, K. B. (2015). MicroRNA (miRNA) in cancer. Cancer cell international, 15, 1-6.

Roberti, A., Valdes, A. F., Torrecillas, R., Fraga, M. F. & Fernandez, A. F. (2019). Epigenetics in cancer therapy and nanomedicine. Clinical epigenetics, 11, 1-18.

Romero-Cordoba, S. L., Salido-Guadarrama, I., Rodriguez-Dorantes, M. & Hidalgo-Miranda, A. (2014). miRNA biogenesis: biological impact in the development of cancer. Cancer biology & therapy, 15, 1444-1455.

Sato, F., Tsuchiya, S., Meltzer, S. J. & Shimizu, K. (2011). MicroRNAs and epigenetics. The FEBS journal, 278, 1598-1609.

Shanmugam, M. K., Arfuso, F., Arumugam, S., Chinnathambi, A., Jinsong, B., Warrier, S., Wang, L. Z., Kumar, A. P., Ahn, K. S. & Sethi, G. (2018). Role of novel histone modifications in cancer. Oncotarget, 9, 11414.

Sharma, S., Kelly, T. K. & Jones, P. A. (2010). Epigenetics in cancer. Carcinogenesis, 31, 27-36.

Siegfried, Z., Eden, S., Mendelsohn, M., Feng, X., Tsuberi, B.-Z. & Cedar, H. (1999). DNA methylation represses transcription in vivo. Nature genetics, 22, 203-206.

Simó-Riudalbas, L., Melo, S. A. & Esteller, M. (2011). DNMT3B gene amplification predicts resistance to DNA demethylating drugs. Genes, Chromosomes and Cancer, 50, 527-534.

Taby, R. & Issa, J. P. J. (2010). Cancer epigenetics. CA: a cancer journal for clinicians, 60, 376-392.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S. & Rajagopal, N. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell, 146, 1016-1028.

Tsukada, Y.-I., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., Tempst, P. & Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature, 439, 811-816.

Verma, M. & Srivastava, S. (2002). Epigenetics in cancer: implications for early detection and prevention. The lancet oncology, 3, 755-763.

Wang, K.-Y. & Shen, C. J. (2004). DNA methyltransferase Dnmt1 and mismatch repair. Oncogene, 23, 7898-7902.

Xhemalce, B., Dawson, M. A. & Bannister, A. J. (2006). Histone modifications. Reviews in Cell Biology and Molecular Medicine.

Yan, L., Nass, S. J., Smith, D., Nelson, W. G., Herman, J. G. & Davidson, N. E. (2003). Specific inhibition of DNMT1 by antisense oligonucleotides induces re-expression of estrogen receptor a (ER) in ER-negative human breast cancer cell lines. Cancer biology & therapy, 2, 552-556.

Yang, A. S., Doshi, K. D., Choi, S.-W., Mason, J. B., Mannari, R. K., Gharybian, V., Luna, R., Rashid, A., Shen, L. & Estecio, M. R. (2006). DNA methylation changes after 5-aza-2′-deoxycytidine therapy in patients with leukemia. Cancer research, 66, 5495-5503.

Yanis Boumber, M. & Issa, J.P. J. (2011). Epigenetics in cancer: what's the future? Oncology, 25, 220.

Yoo, C. B. & Jones, P. A. (2006). Epigenetic therapy of cancer: past, present and future. Nature reviews Drug discovery, 5, 37-50.

Zaman, M. S., Thamminana, S., Shahryari, V., Chiyomaru, T., Deng, G., Saini, S., Majid, S., Fukuhara, S., Chang, I. & Arora, S. (2012). Inhibition of PTEN gene expression by oncogenic miR-23b-3p in renal cancer. PloS one, 7, e50203.

Khazar Journal of Science and Technology Volume 5 №2 2021, 34-48

DOI: 10.5782/2520-6133.2021.5.2.34

Investigation of Variability of Apricot (*Prunus armeniaca* L.) Using Morphological, Pomological Traits and Microsatellite Markers

Amina Rakida

Genetic Resources Institute of ANAS, Baku, AZ1106, Azadliq ave., 155

Abstract

Based on 10 (morphological, pomological, phenological) traits and using 10 microsatellite molecular markers, 17 Azerbaijani apricot cultivars and accessions have been evaluated. All the genotypes manifested a high level of variability. Based on the size, fruits were divided into 2 groups: small fruits (< 40 g) and large fruits (> 40 g). Cultivars such as Ordubad eriyi, Ag erik, Mayovka (Terter) and Ag erik Gulnar with fruit weight above 70 g were estimated as very large**.** In general, fruits had yellow skin ground color and flesh color as well as high TSS. A high correlation was observed between bud break season and blossom season, bud break season and harvest season, bud break season and leaf fall season, blossom season and harvest season, blossom season and leaf fall season, harvest season and leaf fall season. However, a low or insignificant correlation was found between other pomological or phenological characteristics. According to the PCA results, 100% of the total variance among cultivars is attributed to the first seven components. NJ cluster analysis divided apricot cultivars into three main groups. The number of cultivars in the I, II and III clusters were, respectively, eight, seven and two. A total of 60 alleles, ranging from 3 to 9 alleles were revealed by molecular data obtained from microsatellite markers. The average of expected heterozygosity (He), observed heterozygosity (Ho) and polymorphism information content (PIC) were found to be 0.68, 0.77, and 0.63, respectively. The article presents the results of the first genetic diversity analysis of apricot cultivars from the regions. We believe the study will contribute to the effective management and sustainable utilization of apricot germplasm in future breeding programs in the regions.

Keywords: apricot, morphological and pomological traits, microsatellite markers

Introduction

Apricot (*Prunus armeniaca* L.) belonging to the Rosaceae family, is cultivated worldwide. As an important fruit in the Northern hemisphere, it represents the third most planted stone fruit species after peach and plum. Among all temperate fruits, apricot ranks as the seventh in terms of worldwide production. Apricots are native to China and Central Asia. They were the *subject* of successive *domestications twice*, one in Western Central Asia (Fergana valley, at the borders of Uzbekistan, Tajikistan, and Kyrgyzstan) and one in China (Vavilov, 1951; Faust et al., 1998). In these regions, apricots are cultivated mainly for fresh market, kernel production, and ornamental use. Apricot is mostly considered as a self-incompatible species, which fruit has no specific aroma (Zhebentyayeva et al., 2012).

Based on morphological and physiological traits, Kostina (1964) recognized four main eco-geographical groups: Central Asian, Irano-Caucasian, European, and Dzhungar-Zailij. Apricots from Central Asia and the Xinjing Province of China are genetically related to wild forms of P. armeniaca and differ from the East Asian apricots which are related to East Asian wild species. From the center of origin, apricot culture spreads to the Irano-Caucasian region, which included Azerbaijan and constituted the secondary center of apricot diversification following the Silk Road (Vavilov, 1951).

Apricots are grown almost in all regions of Azerbaijan, except very humid regions, for a long time. The cultivation of this plant is expanding every year. Different apricot genotypes are grown in regions of Azerbaijan such as Nakhchivan, Terter, Goranboy, Agdash, etc. In 2020, the total fresh apricot production of Azerbaijan was estimated to be 28977,4 metric tons. However, the available information about the morphological characteristics of apricot is limited. These evaluations are based on a wide range of characteristics such as phenological traits, fruit size, and tree vigor.

Genetic variability is known to be the prerequisite for any plant breeding program (Khush, 2002). New fruit cultivars are generally developed on the basis of genetic resources. Essential stages of any breeding program are germplasm collection and characterization, which are mainly performed by describing phenological, pomological, and morphological characteristics. Thus, plant breeding programs have widely used morphological criteria as important markers (Ogasanovic et al. 2007; Karimi et al. 2008). However, DNA-based molecular markers provide a very useful tool for genetic diversity investigation and identification of cultivars. Thus, the use of microsatellites (simple sequence repeats or SSRs) for molecular characterization and genetic diversity evaluation of different crop species has increased, due to their high polymorphism and wide distribution in the genome. Besides, they are highly

reproducible and co-dominant. Microsatellites have been used for genome mapping and cultivar characterization and variability evaluation in apricot (Raji et al., 2014).

We aimed to investigate the morphological and pomological traits and to perform molecular identification by using SSR primer pairs among 17 apricot genotypes from Terter, Goranboy, Agdash regions of Azerbaijan. Our research is an initial step for the national and international germplasm characterization and preservation of these valuable fruit trees for future breeding programs.

Materials and methods

Plant materials

Seventeen apricot cultivars and forms from Genetic Resources Institute of Azerbaijan (AGRI) were used as the study subjects. The plant material consisted of samples from the apricot germplasm collection located at Genetic Resources Institute of Azerbaijan (Table 1). This germplasm was originally collected from cultivars and forms of different cultivation regions of Terter, Goranboy, Agdash in Azerbaijan.

Morphological and Pomological Characterization

The primary selection criterion was based on the fruit and yield attributes of the genotypes (Table 1). Individual genotypes were marked in the field. The data were recorded at the time of fruit maturity during summer (June - July) seasons from 2016 to 2021 and data were pooled for analysis. Total numbers of fruits were counted per plant. Fruits from each genotype were randomly selected and data were collected on fruit length (mm), fruit weight (g), fruit width (mm) and TSS (^{0}Brix) in apricot genotypes. Fruit weight was measured using Sartorius balance with an accuracy of 0.001 g. The length and width of the fruit were measured with a digital Vernier caliper. The measurement of fruit length was made on the polar axis, i.e. between the apex and stylar end. The maximum width of the fruit was measured in the direction perpendicular to the polar axis.

Statistical methods such as principal component analysis and cluster analysis have been employed as powerful options for plant cultivar and accession screenings. As a tool for germplasm description, we have applied principal component analysis to study correlations among variables and establish relationships among cultivars (Asma and Ozturk., 2005). To identify the principal distinguishing characters of the variability the principal component analysis (PCA) was performed (Wani et. al., 2017). This method is commonly applied for the characterization of genetic

resources in such studies. In addition, Pearson correlation coefficients and correspondence analysis were applied to identify a putative redundancy and to discriminate the relevant informative traits (Krichen et al., 2014). The XLSTAT and SPSS software packages were used for the data analysis.

Molecular Characterization

In this study, we used 17 accessions. Before DNA extraction, young leaf tissue samples were stored at −80 ◦C. We applied the CTAB (cetyltrimethyl ammonium bromide) method of Doyle and Doyle (1990) to isolate genomic DNA. Electrophoresis was performed on 1% agarose gel to determine the DNA quality and it was quantified spectrophotometrically at 260 nm. A prepared DNA solution (10 ng/l) was stored at −20 ^{0}C. We chose 10 SSR primer pairs for the study. Fluorescently labeled M13 primer with 6-FAM, NED, PET and VIC were used for SSR fragment analysis and PCR was performed following the method described by Schuelke (2000). In brief, the PCR reaction mix consisted of 2 µl of 10x PCR buffer, 0,6 µl of 50 mM MgCl2, 2 µl of 10 mM dNTP, 0.15 µl of 10 µM of a sequence-specific forward primer with M13 tail at its 5′ end, 0.35 µl of 10 µM of a sequence-specific reverse primer, and 0.20 µl of 10 µM the universal fluorescent-labeled M13 primer, 0.2 µl 5U/µl Taq polymerase, and 3 µl of 25 ng/ µl sample DNA. The total reaction medium was brought up to 20 µl with distilled water. The PCR program consisted of 3 min an initial denaturation at 94 °C followed by 35 cycles of a 30 s denaturation step at 94 °C, a 40 s annealing step at 60 °C, and a 1 min extension step at 72 °C. Additional 7 min extension step was added to PCR program. Seven µL of PCR product were visualized on 3% agarose gel. 10 µL distilled water was added onto the remaining PCR product and stored at −20 °C until to use for fragment analysis. The stored PCR products for each multiplex were mixed. This concentration was further diluted 20 times and 0.75 µL of the diluted aliquot was mixed with 9 µL Hi-Di buffer and 0.25 µL LIZ 500 standard. The mixtures were kept at 95 °C 5 min, and then immediately were put on ice and finally loaded to capillary electrophoresis on a ABI 3500 (applied bio sysetms, Foster City, Calif., USA) located in Erciyes University, Kayseri, Turkey. DNA fragment sizes were determined using GeneMapper 4.1 software (Applied Biyosystems, Foster City, Calif., USA).

The neighbor-joining (NJ) method was used to construct and draw a dendrogram from the genetic similarity matrix using the DARwin 6 software (Perrier & Jacquemoud-Collet, 2006). A genetic similarity matrix based on the proportion of shared alleles was generated and the expected heterozygosity (He), observed heterozygosity (Ho) and polymorphism information content (PIC) were calculated using the PowerMarker V3.25 software (Table 2).

Table 1. List of apricot genotypes and their source city, used in this study.

No.	Cultivar	Source city
1	Zeynebi	Agdash
2	May Natiq	Agdash
3	Ag erik Gulnar	Agdash
4	Ag erik Elchin	Agdash
5	May Goranboy	Agdash
6	Mayovka	Agdash
7	Badami	Agdash
8	Shalakh	Agdash
9	Ordubad eriyi	Terter
10	Ag erik	Terter
11	Badam erik	Terter
12	Girmiziyanag	Terter
13	İrevan eriyi Shalakh	Terter
14	Mayovka	Terter
15	Ag erik Gejyetishen	Goranboy
16	Ag erik Tezyetishen	Goranboy
17	Badam erik	Goranboy

Table 2. Prunus SSR loci that were used in this study for the characterization of apricot genotypes.

Multiplex groups	Locus	Reference	Linkage Group	Position (From. To)
4-Fam	Gol061	Vera-Ruiz et al. (2011)	1	NA
1-Fam	PGS1.03	Soriano et al. (2012)	1	7320588–7320780
3-6-Fam	PGS1.20	Soriano et al. (2012)	1	8458901–8459094
4-Vic	PGS1.21	Soriano et al. (2012)	1	8527745–8527930
4-Ned	PGS1.23	Soriano et al. (2012)	1	8600638–8600792
2-Pet	PGS1.24	Soriano et al. (2012)	1	8668808–8668989
3-Vic	PGS1.252	Soriano et al. (2012)	1	8730677–8730761
2-6-Fam	96P10_SP6	Soriano et al. (2012)	1	8920241–8920383
3-6-Fam	ssrPaCITA5	Lopes et al. (2002)	1	11770142–11770214
4-Fam	ssrPaCITA17	Lopes et al. (2002)	1	NA

Results and discussion

Morphological and Pomological Characterization

Data in Table 3 represent 17 variables for the studied apricot germplasms. Bud break season for these germplasms in this region is generally from late February to mid-March, the full blossom is observed between early March and early April. A 15– 20-day variation in phenological phases was observed during the 6 years of study course. The germplasms Ag erik Gulnar, Ag erik Gecyetishen, Ag erik, and Ordubad eriyi demonstrated later blossom compared with others. The difference in blossoming periods of germplasms under the same geographical conditions might be a result of the total exposure temperature required. Later blossoming is an important factor for the protection from spring frosts in continental climates (Guleryuz 1988; Unal et al., 1999). In this region, late spring frosts end around mid-April and since all the cultivars and forms in germplasm blossomed before early April, they were all under the risk of spring frost damage.

There are large variations in harvest season between apricot cultivars and forms. Most cultivars were harvested between June and July. The earliest apricot cultivars were 'May Natiq', 'Ag erik Elchin', 'May Goranboy', 'Mayovka', 'Mayovka' (Terter). In this study, many cultivars were harvested in late May. The fruits of Zeynebi, Ag erik Gulnar, Badami, Shalakh, Ag erik Tezyetishen, Badam erik, Badam erik (Goranboy), Girmızıyanag, Irevan eriyi (Shalakh) were harvested on mid-June, while Ag erik Gecyetishen, Ordubad eriyi, Ag erik were harvested in late June. In the previous studies, the harvest data for apricot cultivars were in the range of 14 May-26 June in Spain (Ruiz & Egea, 2008), 11 June-10 September in Hungary (Hegedũs et al., 2010), 26 May-25 June in Italy (Lo Bianco et al., 2010). These differences could be due to climatic conditions in Terter, Goranboy, Agdash, where the climate is semi-arid, with hot summers and cold winters. These regions had high day-night temperature changes from February to May (>15°C) and maximum temperatures were >28°C in April and May. Therefore, the apricot cultivars could be early fulfilling degree-day thresholds from full-bloom to be harvested under these conditions. Ruml et al. (2010) indicated that the effect of growing degree-day thresholds on harvest time of apricots is very important for each apricot-producing region. The authors also reported that daily maximum temperatures were the most influential temperature variable for the ripening time of apricots. The fruit size is one of the most important fruit quality traits for fresh apricots. A high degree of variation is an important trait related to fruit size. Regarding the 'Apricot Descriptor' (IPGRI and CEC, 1984), the fruit weight for 'Ordubad eriyi'(100.2 g) , followed by 'Ag erik' (86.3 g), 'Mayovka' (Terter) (86.3 g), and 'Ag erik Gulnar' (71.1 g) were very large (>70 g), 'May Goranboy', 'Badam erik' and 'Ag erik Tezyetishen' were large (61-

70 g), ‘Badami’ and ‘Irevan eriyi’ (Shalakh) were medium-to-large (56-60 g) under ecological conditions of Terter, Goranboy, Agdash. The other cultivars ‘May Natiq’, ‘Mayovka’, ‘Girmiziyanag’, ‘Ag erik Gecyetishen’ and ‘Badam erik’ (Goranboy) (45-50 g) were medium, ‘Shalakh’ (43.1 g) small/medium and ‘Zeynebi’ and ‘Ag erik Elchin’ (10-20 g) small. The fruit weights of 2 accessions were less than 40 g indicating that they are small-fruited. This result is not compatible with the findings of Asma and Ozturk (2005) in Turkish apricots. Large fruits of ‘Ordubad eriyi’ make it a promising cultivar in apricot breeding programs to improve fruit size. In our study most genotypes had desirable fruit sizes. Attractive medium-sized fruits are desired for apricot cultivar breeding (Mratinić et al., 2011).

The fruit length and width was ranged from 11.4 mm to 70.8 mm and from 11.2 to 65.6 mm, respectively. The maximum value for FL was recorded in Ordubad eriyi, followed by Ag erik (63.8 mm), whereas minimum score was in Ag erik Elchin. The highest FWi was obtained in Ordubad eriyi, followed by Mayovka (Terter) (56.2 mm) and Ag erik (54.2 mm). As in previous parameter, the lowest fruit width was also recorded in Ag erik Elchin. The previous studies on apricot also reported a high variability among apricot cultivars regarding fruit-size traits (Badenes et al., 1998; Ruiz & Egea, 2008).

The TSS content is an important quality attribute, influencing notably the fruit taste. The levels of TSS in this study ranged from 10.2 ^{0}Brix (‘Mayovka’) to 20 ^{0}Brix (‘Shalakh’) with a mean of 14.5^{0}Brix, which is greater than the minimum (10 ^{0}Brix) established by the EU (European Union) to market apricots (R-CE No.112/2001). The high values for TSS were noted in genotypes Shalakh (20^{0}Brix) which is characterized by an excellent fruit taste (level of TSS >20^{0}Brix) and quality (Ayanoglu *et al*., 1995). Mayovka had the least amount of TSS. As seen in Table 3, the majority (53%) of the studied apricot cultivars were characterized by yellow skin ground color and flesh color. SGC for 4 cultivars, namely Ag erik Elchin, Badami, Ordubad eriyi, Mayovka (Terter) was orange. The skin ground color of Ag erik Gecyetishen was white and its flesh was cream color. Girmızıyanag had green-yellowish skin ground color.

Table 3. Description of apricot cultivars.

Varieties	**BBS**	**BS**	**HS**	**LFS**	**FW, g**	**FL, mm**	**FWi, mm**	**TSS (°Brix)**	**FC**	**SGC**
Zeynebi	2	2	2	2	11.1	16	14.8	15.1	1	1
May Natiq	1	1	1	1	48.6	46.8	44.8	14	1	1

Ag erik Gulnar	2	2	2	2	71.1	53.4	47.4	15.1	1	1
Ag erik Elchin	1	1	1	1	16.1	11.4	11.2	15.1	2	2
May Goranboy	1	1	1	1	62.8	54.6	52.4	14	1	1
Mayovka	1	1	1	1	49.6	36.6	39.8	10.2	1	1
Badami	2	2	2	2	60.1	52.4	50.8	15.1	2	2
Shalakh	2	2	2	2	43.1	41	35.8	20	1	1
Ordubad eriyi	3	3	3	3	100.2	70.8	65.6	15	2	2
Ag erik	3	3	3	3	86.3	63.8	54.2	13	1	3
Badam erik	2	2	2	2	65.9	53.2	46.1	16	1	1
Girmiziyanag	2	2	2	2	48.8	38.6	34.8	13	1	4
İrevan eriyi (Shalakh)	2	2	2	2	56.2	50.8	46.2	14	1	1
Mayovka (Terter)	1	1	1	1	86.3	58	56.2	13	2	2
Ag erik Gecyetishen	3	3	3	3	48.4	45.8	43.8	16	4	3
Ag erik Tezyetishen	2	2	2	2	68.4	51	44.6	14	1	3
Badam erik (Goranboy)	2	2	2	2	48.9	48.6	36.8	14.5	1	1

Correlation among variables

Several phenological and pomological characteristics were found to be highly correlated (Table 4). As expected, the highest values were recorded between bud break season and blossom season, bud break season and harvest season, bud break season and leaf fall season and blossom season and harvest season, blossom season and leaf fall season, harvest season and leaf fall season with maximum Pearson correlation index ($r=1.00$; $p < 0.01$). Strong positive correlations were also observed for fruit size parameters. Thus, fruits with larger sizes (both FL and FWi) had a higher weight. The Pearson correlation indices of fruit weight with fruit length and width were determined as $r = 0.948$ and $r = 0.941$, respectively. Fruit size traits (FL, FWi, FW) were negatively correlated with the amount of total soluble solids, however, they were not statistically significant. This is in agreement with the results of Badenes et al. (1998) who reported a high correlation ($r=0.87$) between bud break and blossom season and also a correlation (to a lesser extent) between bud break and harvest season ($r=0.79$).

Table 4. Correlation matrix among variables studied.

	BBS	BS	HS	LFS	FW	FL	FWi	TSS	FC
BS	1.000**								
HS	1.000**	1.000**							
LFS	1.000**	1.000**	1.000*						
FW	0.337	0.337	0.337	0.337					
FL	0.396	0.396	0.396	0.396	0.948**				
FWi	0.293	0.293	0.293	0.293	0.941**	0.973**			
TSS	0.307	0.307	0.307	0.307	-0.199	-0.080	-0.161		
FC	0.319	0.319	0.319	0.319	0.027	0.019	0.095	0.177	
SGC	0.419	0.419	0.419	0.419	0.203	0.100	0.089	-0.187	0.376

Correlations significant ($p < 0.05$). Abbreviations: TS, tree size; BBS, bud break season; BS, blossom season; HS, harvest season; LFS,

leaf fall season; FW, fruit weight; PW, pit weight; KW, kernel weight; SGC, skin ground color; FC, flesh color; KT, kernel taste; PFR,

flesh/pit ratio; BRIX, total solids soluble; TA, total acidity; Y, yield.

**Correlation is significant at 0.01. Abbreviations: BBS, bud break season; BS, blossom season; HS, harvest season; LFS, leaf fall season; FW, fruit weight; FL, fruit length; FWi, fruit width; TSS, total solids soluble; FC, flesh color; SGC, skin ground color.

Correlations significant ($p < 0.05$). Abbreviations: TS, tree size; BBS, bud break season; BS, blossom season; HS, harvest season; LFS,

leaf fall season; FW, fruit weight; PW, pit weight; KW, kernel weight; SGC, skin ground color; FC, flesh color; KT, kernel taste; PFR,

flesh/pit ratio; BRIX, total solids soluble; TA, total acidity; Y, yield.

Principal component analysis (PCA)

About 100% of the total variance among genotypes was explained by the first seven components (Table 5). PC1 which represents bud break season, blossom season, harvest season, and leaf fall season accounted for about 50.78% of total variance.

PC2 that includes fruit weight, fruit length, and fruit width comprised about 24.28% of overall variance. PC3 consisted of SGC constituted about 11.96% of total variance. Table 6 shows the amount of loading factors for each character in the first five principal components. As observed in PCA, the first three components represented 87.02% of total variance. This value is higher than those of reported by Raji et al. (2014), Asma and Ozturk (2005), and Yilmaz et al. (2012) (54, 70, and 73%, respectively).

Table 5. Eigen values, variance%, and cumulative% of first seven factors contributing to 100% of total variance

PC	Eigenvalue	Variability (%)	Cumulative %
1	5.08	50.78	50.78
2	2.43	24.28	75.06
3	1.20	11.96	87.02
4	0.84	8.44	95.46
5	0.39	3.91	99.37
6	0.05	0.50	99.87
7	0.01	0.13	100.00

Table 6. Zoading factor of variables in the first five principal components (PCs)

Variable	PC_1	PC_2	PC_3	PC_4	PC_5
BBS	0.941	-0.295	-0.064	-0.135	-0.075
BS	0.941	-0.295	-0.064	-0.135	-0.075
HS	0.941	-0.295	-0.064	-0.135	-0.075
LFS	0.941	-0.295	-0.064	-0.135	-0.075
FW, g	0.612	0.762	0.001	0.052	0.093
FL, mm	0.654	0.725	-0.151	0.085	0.050
FWi, mm	0.578	0.782	-0.065	0.184	-0.008
TSS (°Brix)	0.212	-0.505	-0.631	0.402	0.375
FC	0.367	-0.289	0.423	0.746	-0.213
SGC	0.470	-0.148	0.759	-0.095	0.414

Molecular Characterization

Polymorphism and diversity of SSR markers

Microsatellite primer pairs revealed a total of 60 alleles with a mean value of 6 alleles per locus, ranging from 3 alleles (PGS1.252 and ssrPaCITA5) to 9 alleles (PGS1.20 and ssrPaCITA17) (Table 7). Zhang, Q.P et al. (2013) reported a total of 318 alleles and the mean value of 15.14 alleles per locus in 94 apricot cultivars using SSR primers. Gürcan et al. (2015) reported 230 alleles in 239 apricot cultivars using 18 SSR loci. With ten SSR primer pairs, the total number of alleles and the mean value of alleles per locus in 39 apricot cultivar were 53 and 5.3, respectively (Raji et al., 2014). The average number of 7 alleles per locus is also higher than that found by Hormaza et al. (2002) and Sanchez-Perez et al. (2005) in apricot (4.1 and 3.9, respectively). The high value of average alleles per locus (6) confirms that SSRs are very useful markers for the identification of apricot cultivars. Our results showed successful cross-species transferability of SSR primers identified in different Prunus species to study genetic diversity in apricot. The average number of genotypes was 6.7, ranging from 9 in PGS1.20 and ssrPaCITA17 locus to 3 in PGS1.252 locus. The higher number of genotypes represents more power of loci for discriminating genotypes, as were detected in PGS1.20 and ssrPaCITA17. Excepted heterozygosity among loci ranged from 0.44 in ssrPaCITA5 to 0.82 in PGS1.03 with an average of 0.68. The average expected heterozygosity reported for apricot cultivars was 0.79, 0.75, 0.63, and 0.63 (Zhang et al., 2013; Gürcan et al., 2015; Bourguiba et al., 2012; Raji et al., 2014). Observed heterozygosity also ranged from 0.35 in ssrPaCITA5 to 0.94 in ssrPaCITA17, with an average of 0.77. The observed heterozygosity reported by Gürcan et al. (2015) ranged from 0.30 to 0.94 with an average of 0.63 in 239 apricot accessions from different eco-geographical groups. The average value of PIC was 0.63, with a minimum of 0.38 in ssrPaCITA5 locus and a maximum of 0.79 in PGS1.03 loci. The PIC value was similar in apricot (0.586) as reported by Bourguiba et al. (2012).

A NJ dendrogram based on the shared allele distance was used for grouping the apricots. (Fig. 1). The dendrogram generated from the cluster analysis divided 17 genotypes into three main groups (Fig 1). The first group consisted of 8 genotypes (Badami, Ag erik Gulnar, Shalakh, Ag erik Tezyetishen, Badam erik, May Goranboy, Girmiziyanag, Badam erik) that are 47.06% of the total genotypes with the highest fruit length, fruit diameter, fruit weight, medium TSS. Seven genotypes (May Natiq, Zeynebi, Ag erik Gecyetishen, Mayovka, Ag erik Elchin, Shalakh and Mayovka (Terter)) comprised the second group, which contains 41.18% of the total genotypes in this population. It had the low to high fruit weight, low to medium fruit length, fruit width, low to high TSS. The third cluster consisted of two genotypes (Ag erik, Irevan eriyi (Shalakh) characterized by medium to high fruit weight, medium fruit length, fruit width, and TSS.

Table 7. Variability parameters of 10 SSR markers on 17 apricots investigated genotypes.

Primer name	Ng	Na	He	Ho	PIC
Gol061	7	4	0.66	0.65	0.60
PGS1.03	8	7	0.82	0.82	0.79
PGS1.20	9	9	0.75	0.82	0.71
PGS1.21	7	7	0.64	0.94	0.58
PGS1.23	5	5	0.67	0.88	0.62
PGS1.24	7	6	0.70	0.59	0.66
PGS1.252	3	3	0.52	0.88	0.41
96P10_SP6	8	7	0.80	0.82	0.77
ssrPaCITA5	4	3	0.44	0.35	0.38
ssrPaCITA17	9	9	0.78	0.94	0.75
Mean	6,7	6	0.68	0.77	0.63

Ng, the number of genotypes; Na, the number of alleles; He, expected heterozygosity; Ho, observed heterozygosity; PIC, polymorphism information content.

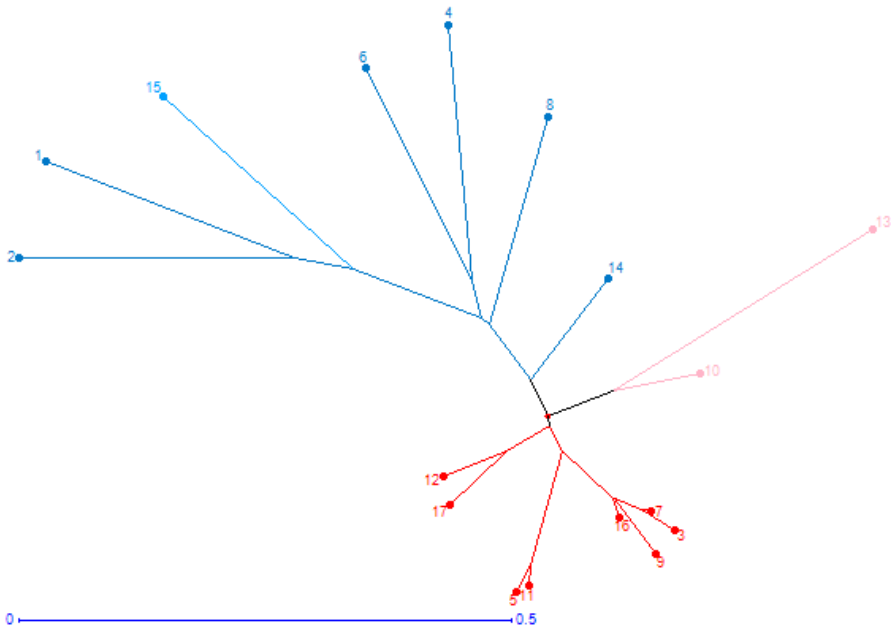


Figure 1. Neighbor-joining dendrogram based on simple matching dissimilarity matrix showing relationships among the 17 apricot accessions. Figure 1 with cluster 1 in red, cluster 2 in blue, and cluster 3 in pink.

Conclusion

For the first time in Azerbaijan, the extant apricot germplasm of Terter, Goranboy, Agdash regions has been evaluated for the purpose of plant sustainable utilization. The research results showed great biodiversity among Azerbaijan apricots. This variation could be used in apricot-breeding programs to improve fruit quality, extend ripening season, and late flowering season. For breeding large fruit-sized cultivars for better marketability and returns to the growers, Azerbaijani apricot cultivars are particularly important due to the large variation in fruit size. Based on this evaluation, 15 genotypes with fruit sizes greater than 40 g were identified. Cultivar "Ordubad eriyi" with the largest fruit (100.2 g) represents a highly precious variety that may be employed to breed large-sized apricots in future breeding programs. SSR molecular markers are of great importance for these breeding programs. In this study, polymorphism of 10 SSR loci from different *Prunus* species was reported. According to the results of the research, *Prunus* SSR loci are highly conserved and can be applied in apricot breeding programs to maximize genetic variability for generating new cultivars for cultivation under Azerbaijan conditions.

References

Agriculture, Erzurum
Agriculture, Erzurum
apricots quality and resistance to spring frosts in erzincan
apricots quality and resistance to spring frosts in erzinc

Asma, B. M. & Ozturk, K. (2005). Analysis of morphological, pomological and yield characteristics of some apricot germplasm in Turkey. Genet. Resour. Crop Ev. 52, 305–313.

Ayanoglu, H., Kaska, N. & Yildiz, A. (1995). Investigations on adaptations of early apricot cultivars in Mediterranean region. Proceedings of the Second National Horticultural Congress, p.159-163, Adana, Turkey.

Badenes, M.L., Martinez-Calvo, J. & Llácer G. (1998). Analysis of apricot germplasm from the European Eco geographical group. Euphytica 102(1):93-99.

Bourguiba, H., Audergon, J., Krichen, L., Trifi-Fara, N., Mamouni, A., Trabelsi, S., D'Onofrio, C., Asma, B., Santoni, S. & Khadari, B. (2012). Loss of genetic diversity as a signature of apricot domestication and diffusion into the Mediterranean Basin. BMC Plant Biol. 2012, 12–49.

Decroocq, S. A., Chague, P., Lambert, G., Roch, J. M., Audergon, F., Geuna, R., Chiozzotto, D., Bassi, L., Dondini, S., Tartarini, J., Salava, B., Krska, F., Palmisano, ˇ I. & Karayiannis, V. (2014). Selecting with markers linked to the PPVres major QTL is not sufficient to predict resistance to Plum pox virus (PPV) in apricot. Tree Genet. Genomes 10, 1161–1170.

Doyle, J. J. & Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. Focus 12, 13–15.

Faust, M., Surányi, D. & Nyujtó, F. (1998). Origin and dissemination of apricot. Hort. Rev. 22, 225–266.

Guleryuz, M. (1988). A study on breeding by selection of wild apricots quality and resistance to spring frosts in erzincan plain. Professor thesis, Ataturk University Faculty of Agriculture, Erzurum

Gürcan, K., Öcal, N., Yılmaz, K. U., Ullah, S., Erdoğan, A. & Zengin, Y. (2015). Evaluation of Turkish apricot germplasm using SSR markers: Genetic diversity assessment and search for Plum pox virus resistance alleles. Scientia Horticulturae, 193

Hegedũs, A., Engel, R., Abrankó, L., Balogh, E., Blázovics, A., Hermán, R., Halász, J., Ercisli, S., Pedryc, A. & Stefanovitsbányai, É. (2010). Antioxidant and antiradical capacities in apricot (*Prunus armeniaca* L.) fruits: Variations from genotypes, years, and analytical methods. J Food Sci 75(9):C722-C730.

Hormaza, J.I. (2002). Molecular characterization and similarity relationships among apricot genotypes using simple sequence repeats. Theor. Appl. Genet. 104, 321–328

IPGRI, CEC. (1984). Revised descriptor list for apricot (*Prunus Armeniaca*). Editors: Guerriero R., Watkins R. International Board for Plant Genetic Resources Commission of European Communities, Committee on Disease Resistance Breeding and use of Genebanks. Rome, Italy.

Karimi, H. R, Zamani, Z., Ebadi, A. & Fatahi, M. R. (2008). Morphological diversity of pistacia species in Iran. Genetic Resources and Crop Evolution 44: 76–81.

Khush, G.S. (2002). Molecular genetics-plant breeder's perspective. (In) Molecular Techniques in Crop Improvement, Jain S M, Brar D S, and B S Ahloowalia (Eds.). pp 1–8 Kluwer.

Kostina, K. F. (1964). "Application the phytogeographical method to apricot classification (in Russian)," in Proceedings (Trudi) of the Nikita Botanical Gardens Vol 24, Moscow.

Krichen, L., Audergon, J. M. & Trifi-Farah, N. (2014). Variability of morphological characters among Tunisian apricot germplasm. Scientia Horticulturae, 179, 328–339.

Lo Bianco, R., Farina, V., Indelicato, S. G, Filizzola F. & Agozzino P. (2010). Fruit physical, chemical and aromatic attributes of early, intermediate and late apricot cultivars. J Sci Food Agric 90(6):1008-1019.

Lopes, M. S., Sefc, K. M., Laimer, M. & da Camara Machado, A. (2002). Identification of microsatellite loci in apricot. Mol. Ecol. Notes 2, 24–26.

Mratinić, E., Popovski, B., Milošević, T. & Popovska M. (2011). Analysis of Morphological and Pomological Characteristics of Apricot Germplasm in FYR Macedonia E. J. Agr. Sci. Tech. Vol. 13: 1121-1134

Ogasanovic, D., Plazinic, R., Rankovic, M., Stamenkovic, S. & Milinkovic, V. (2007). Pomological characteristics of new plum cultivars developed in Cacak. Acta Horticulturae 734: 165–8.

Perez-Gonzales, S. (1992). Associations among morphological and phenological characters representing apricot germplasm in Central Mexico. J. Am. Soc. Hort. Sci. 117, 486–490.

Perrier, X. & Jacquemoud-Collet, J. P. (2006). DARwin Software. Available online at: http://darwin.cirad.fr/darwin.

Raji, R., Abbasali, J., Reza, F. & Mohammad Abedini, E. (2014). Investigation of variability of apricot (*Prunus armeniaca* L.) using morphological traits and microsatellite markers. Scientia Horticulturae, 176, 225–231.

Ruiz, D. & Egea, J. (2008). Phenotypic diversity and relationships of fruit quality traits in apricot (*Prunus armeniaca* L.) germplasm. Euphytica, 163:143-158.

Ruml, M., Vuković, A. & Milatović, D. (2010). Evaluation of different methods for determining growing degree-day thresholds in apricot cultivars. Int J Biometerol 54:411-422.

Sanchez-Perez, R., Ruiz, D., Dicenta, F., Egea, J. & Martinez-Gomez, P. (2005). Application of simple sequence repeat (SSR) markers in apricot breeding: molecular characterization, protection, and genetic relationships. Sci. Hortic. 103, 305–315

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments A poor man's approach to genotyping for research and high-throughput diagnostics. Nature Biotechnology 18:1-2.

Soriano, J. M., Domingo, M. L., Zuriaga, E., Romero, C., Zhebentyayeva, T., Abbott, A. & Badenes, M. L. (2012). Identification of simple sequence repeat markers tightly linked to Plum pox virus resistance in apricot. Mol. Breed. 30, 1017–1026.

Unal, M. S., Sahin, M., Olmez, H., Celik, B., Asma, B. M. & Bas, M. (1999). The Breeding of Late Flowering and Resistance to Late Spring Frosts Apricots through Crossing (First Phase). Tagem/IY/96–06–02–014, Fruit Research Institute, Malatya.

Vavilov, N. I. (1951). The origin, variation, immunity and breeding of cultivated plants. Soil Sci. 72:482.

Vera-Ruiz, E. M., Soriano, J. M., Romero, C., Zhebentyayeva, T., Terol, J., Zuriaga, E., Llácer, G., Abbott, A.G. & Badenes, M.L. (2011). Narrowing down the apricot Plum pox virus resistance locus and comparative analy. Mol. Plant Pathol. 12, 535–547

Wani A. A., Zargar, S. A.; Malik, A. H., Kashtwari, M., Nazir, M., Khuroo, A. A., Ahmad, F. D. & Tanveer A. (2017). Assessment of variability in morphological characters of apricot germplasm of Kashmir, India. Scientia Horticulturae, 225, 630–637.

Yilmaz, K. U., Paydas Kargi, S. & Kafkas, S. (2012). Morphological diversity of the Turkish apricot (*Prunus armeniaca* L.) germplasm in the Iran of Caucasian eco-geographical group. Turk. J. Agric. For. 36, 688–694.

Zhang, Q. P., Liu, D. C., Liu, S., Liu, N., Wei, X., Zhang, A. M., & Liu, W.S. (2013). Genetic diversity and relationships of common apricot (*Prunus armeniaca* L.) in China based on simple sequence repeat (SSR) markers. Genetic Resources and Crop Evolution, 61(2),

Zhebentyayeva, T. N., Ledbetter, C., Burgos, L. & Llácer, G. (2012). "Apricots," in Handbook of Plant Breeding. Fruit Breeding, Vol. 8, eds M. L. Badenes and D. H. Byrne (New York, NY: Springer), 415–458.

Khazar Journal of Science and Technology Volume 5 №2 2021, 49-70

DOI: 10.5782/2520-6133.2021.5.2.49

Brain Tumor Segmentation Methods based on MRI images: Review Paper

Behnam Kaini Kalejahi[1*], Sebalan Danishvar[2], Jala Guluzade[1]

[1]*Faculty of Engineering and Applied Sciences, Khazar University, Azerbaijan*

[2]*College of Engineering, Design and Physical Sciences*

[*]Corresponding author: bkiani@khazar.org

Abstract

Statistically, incidence rate of brain tumors for women is 26.55 per 100,000 and this rate for men is 22.37 per 100,000 on average. The most dangerous occurring type of these tumors are known as Gliomas. The form of cancerous tumors so-called Glioblastomas are so aggressive that patients between ages 40 to 64 have only a 5.3% chance with a 5-year survival rate. In addition, it mostly depends on treatment course procedures since 331 to 529 is median survival time that shows how this class is commonly severe form of brain cancer. Unfortunately, a mean expenditure of glioblastoma costs 100,000$. Due to high mortality rates, gliomas and glioblastomas should be determined and diagnosed accurately to follow early stages of those cases. However, a method which is suitable to diagnose a course of treatment and screen deterministic features including location, spread and volume is multimodality magnetic resonance imaging for gliomas. The tumor segmentation process is determined through the ability to advance in computer vision. More precisely, CNN (convolutional neural networks) demonstrates stable and effective outcomes similar to other automated methods in terms of tumor segmentation algorithms. However, I will present all methods separately to specify effectiveness and accuracy of segmentation of tumor. Also, most commonly known techniques based on GANs (generative adversarial networks) have an advantage in some domains to analyze nature of manual segmentations..

Keywords: Generative adversarial networks, Brain Tumor, Segmentation, Medical Images. **Introduction**

Today, three major categories can explain brain tumor segmentation based on different principles and degree of human interaction requirements. These include semi-automatic, fully automatic and manual segmentations (Pham et al., 2000). Firstly, semi-automatic brain tumor segmentation consists of software, interaction and user. The realization of tumor segmentation algorithms is a target area for software computing. The interaction covers the adjustment of segmentation information between the software and the user. For user case, it provides feedback response and visual information for software computing, but it requires to input some parameters before processes. Moreover, three processes of semi-automatic segmentation include feedback response, evaluation and initialization. One of the disadvantages of this segmentation category is to obtain same user at different times or different results from different experts. However, semi-automatic segmentation shows better results in comparison to manual segmentation. On the other hand, fully automatic brain tumor segmentation algorithm is a combination of prior knowledge and artificial intelligence. It is more likely to stimulate human intelligence to develop machine learning algorithms, but this method helps the computer analyze brain tumor segmentation without human interaction.

The manual brain tumor segmentation paints the regions of anatomic structure by using various labels and it draws the boundaries of brain tumor (Pham et al., 2000). In this case, anatomy as a representation of brain tumor images should be studied by brain tumor experts, but most of the time manual segmentation yields poor results due to error-prone and time-consuming issues. This is because the more brain tumor images in the clinic are emerging, the more errors occur for the experts. The solutions are semi-automatic and fully automatic segmentation methods which address such problems directly in an advanced way. In contact, these two segmentation methods exhibit partial-volume effects and irregular boundaries with discontinuities for tumor images. Currently, alternative three categories are proposed through MRI-based methods for brain tumor segmentation in order to solve challenges faced by semi-automatic and fully automatic brain tumor segmentation images.

1. Segmentation Methods

2.1. Classification and clustering methods

Due to the practice of radiology by radiologists, making accurate decisions and learning patterns from empirical data or learning complex relationship can be studied by providing machine learning which simplifies the diagnosis and anaylsis for medical images (Duda et al., 2012). In medical practice, unsupervised, semi-supervised and supervised learning are essential categories of classification and

clustering methods based on different principles and utilization of labels of training samples (Bishop et al., 2006), (Mitchell et al., 2006). Unsupervised learning algorithm contains no label information and only one set of observations for each sample (Alpaydin et al., 2004). Obviously, latent variables and a set of unobserved variables cause such features and observations. The main objective of the unsupervised learning is to reveal the latent variables and to determine relationship between samples between samples and behind the observations respectively (Hastie et al., 2005). Hence, clustering algorithm explains unsupervised learning effectively. Semi-supervised learning explores a combination of unsupervised and supervised learning algorithms (Chapelle et al., 2006). Due to having high costly labeling of data and being inapplicable for some applications semi-supervised learning was targeting the development of its algorithms (Christakou et al., 2005). Indeed, it has an advantage in terms of using unlabeled and labeled data in the training process. In addition, supervised learning algorithms study two major parts which include output observations (called as effects) or labels and input observation (called as causes) or features. Supervised learning shows a functional relationship which is a set of numerical coefficients and equations from training data. This data generalizes given procedures to testing data so that classification algorithm effectively explains supervised learning as a representative method. Classification or clustering methods in brain tumor segmentation include k-means, Markov Random Fields (MRF), Support Vector Machines (SVM), Artificial Neural Networks (ANN), Fuzzy C-means (FCM) and Atlas-based.

2.2. FCM algorithms

Pattern recognition is commonly used area for FCM method which corresponds to each cluster center between the data point and the cluster on the basis of distance issue, but before it assigns membership to each data point. Getting high possibility of membership towards the cluster center depends how nearer the data is to the cluster center. Getting encouraging results of MR data, getting satisfactory results better than k-means algorithm and better for overlapped data set, assigning membership of data points to not only one cluster center in which most data point exclusively belongs to one cluster center not similar to k-means are significant advantages of FCM algorithm.

Brain tumor segmentation is divided into tissue classes which generate segmentation images in order to demonstrate neuropathological and neuroanatomic issues by generating contrast information from raw MR image data. Necrotic core, active cells and unsupervised FCM clustering algorithm used edema are included in these classes. Furthermore, the integration of multispectral histogram analysis and knowledge-based methods enables MRI images to determine segmentation of brain

tumor. Fuzzy clustering as a knowledge-based method is an alternative procedure for MRI images of brain tumor segmentations. It is also used to build the tumor shape based on 3-D connected components. in addition to fuzzy knowledge and modified seeded region growing, a segmentation method which is so-called Fuzzy Knowledge-based Seeded Region Growing (FKSRG) shows effective segmentation results for multispectral MR images compared to segmentation of functional MRU with Brain Automated Segmentation Tool (Lin et al., 2012).

The FCM is considered as an iterative algorithm which is very time-consuming clustering method. In fact, Fast Generalized FCM (FGFCM) and Bias-Corrected FCM (BCFCM) algorithms are major solutions to decrease possible execution time in advance. FGFCM clustering algorithms are considered as robust FCM framework for segmentation of brain tumor. While BCFCM is a time-efficient algorithm in terms of providing brain images with good quality segmentation. In this method, supporting virtual brain endoscopy is a process obtained effectively by this algorithm in order to better analyze brain tumor segmentation.

A modified FCM-based method determines fast and accurate segmentation aiming to decrease sensitivity issue of standard FCM algorithm. For that reason, this method was proposed for mixed noises including impulse, intensity non-uniformity and Gaussian noises. Moreover, this method uses context based dependent filtering technique to better realize gray and spatial level distances. In this case, first step is extraction of a scalar feature value through neighborhood of each pixel. The next procedure is an observation of enhanced FCM algorithm regarding histogram-based approach connected with clustering process. Most researchers study a neighborhood attraction concept depending on features and relative location of pixels in order to develop the performance of FCM algorithm. However, segmentation results depend on degree of attraction in which determination of this process is difficult to obtain in most cases. The Genetic Algorithm (GAs) and Particle Swarm Optimization (PSO) have been introduced to reach optimal solutions. PSO has no challenges in finding exact solutions while GAs shows low performance for this issue. While the determination of the optimum value for degree of attraction is practically achieved through the combination of PSO and GAs. In addition to segmentation of brain tumor, the combination of fuzzy c-means and k-means algorithms are essential to determine the stage and size of tumor accurately. In contrast to the manual segmentation, this combination offers reproducibility and accuracy with tumor tissue segmentation that reduces time for the improvement of segmentation.

2.3. Atlas-based algorithms

Atlas-based algorithm is used to register various images, guide segmentation of brain tissue, and restrict tumor location. It is also used for generative classification models and three major steps are included in this algorithm. First of all, the atlas and the patient are added to global correspondence by assistance of an affine registration. Second, a template is introduced for brain tumor regarding seeding of a synthetic tumor into brain atlas. Third, brain tumor growth and optical flow principles explain seeded atlas deformation (Cuadra et al., 2004). Furthermore, researchers examine the tissue model through defining probabilistic information and imposing spatial constraints depending on atlases. Moreover, Expectation Maximization (EM) method is introduced to modify an atlas from different MRI modalities knowing that the information about tumor location is linked to the modified atlas with patient-specific information. This procedure enables a probabilistic tissue model to be employed and brain tumor to be segmented. One of advantages of atlas-based methods is about better integration between domain knowledge and consideration of atlas-based segmentation. While it is challenging to account the variability of such prior information. In fact, lesion growth prior model shows radial expansion of lesion which comes from the starting point for brain atlas deformation. Obviously, this process results in better segmentation results of brain tumor by analyzing large space-occupying tumors. In a population, brain atlases contain averaging pre-segmented images which are equally constructed and these methods experience lower segmentation guidance capability and local inter-subject structural variability. In order to solve such problems a multi-region-multi-reference framework is a best alternative to consider for atlas-based neonatal segmentation of brain tumor. As a result of using a spatial regularization and generative probabilistic model, the combination of a latent brain tumor atlas and healthy brain atlas is an accurate determination of brain segmentation from multi-sequence images. The Figure 1 demonstrates the results of T1 and T2 brain tumor segmentation as a part of atlas-based segmentation of magnetic resonance brain images. The localization suggests multi-modal segmentation optic pathway gliomas to be used for classification with probabilistic tissue model based on brain atlas at a recent time (Weizman et al., 2012). The effectiveness and practicability of atlas-based methods have a close relationship with the precise atlas. In Table 1, Talairach-Tournoux, Whole Brain, BrainWeb, and Brodmann are examples for atlases described as follows:

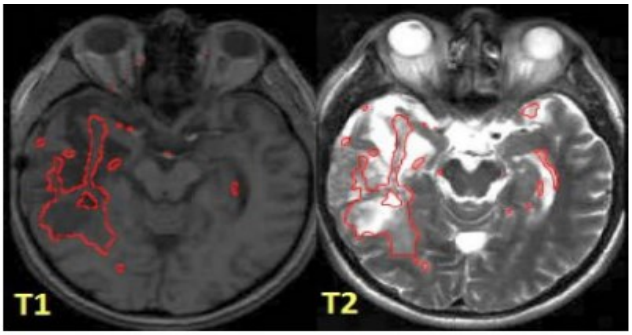


Figure 1. The results of T1 and T2 brain tumor segmentation

Table 1. Talairach-Tournoux, Whole Brain, BrainWeb, and Brodmann as an example

Name	Representation	References
Brodmann	The first brain atlas	Brodmann, 1909
Talairach-Tournoux	Construct a three-dimensional coordinate to provide a standard space	Talairach & Tournoux, 1988
BrainWeb	Widely used in the brain MRI images analysis	Evans et al., 1991
Whole Brain	Used in neurosurgery at Harvard Medical School	Shenton et al., 1995

2.4. MRF algorithms

These algorithms describe the integration of spatial information with classification or clustering process. Overlapping and effect of noise are possible issues that have been reduced by adding MRF in clustering methods. If the neighbor of labeled region is the same, MRF will determine this process by the fact that the region is strongly labeled (non-brain tumor or brain tumor) (Tran et al., 2005). Sequence data is labeled and segmented through building probabilistic models using Conditional Random Fields (CRF). Both MRF and CRF provide high accuracy for segmentation of brain tumor results by representing complex dependencies among data sets. GMM as an example of the mixture model can model different tissues including Necrotic Core (NC), Edema (E), GM, WM, Active Cells (AC) and CSF. This model uses Iterated Condition Modes (ICM) algorithm to train the MRF. Each tissue can be segmented by different models of different tissues. A multi-layer MRF framework can easily detect brain abnormalities so that such layers include input, structural coherence, region intensities, and spatial locations (Gering et al., 2002). Moreover, it is clear

that a change in high-level classification depends on a given voxel which is correlated with strong similarities shared by the attributes of lower-level layers. Spatial accuracy-weighted Hidden Markov random field and Expectation maximization (SHE) provides better quality of tumor segmentation in terms of enhanced-tumor and automated tumor segmentation (Nie et al., 2009). In clinical applications, high-resolution images are commonly determined together with low-resolution sequences. The process of tumor segmentation follows multi-channel MR images using different resolutions through incorporation of the optimization procedure of the Hidden MRF (HMRF) with the spatial interpolation accuracy of low-resolution images proposed by SHE (Bauer et al., 2011). Consequently, SHE algorithm presents more accuracy for the results of tumor segmentation. In case of an automatic method, brain tissues are segmented based on non-rigid registration of an average atlas which is combined with a biomechanically justified tumor growth model. It aims to detect causality of tumor mass-effect in a way to simulate soft-tissue deformations. Correspondence between the patient image and the atlas is the process provided by the tumor growth model which is considered and formulated as mesh-free MRF energy minimization problem before registration step. Compared to other approaches, tumor growth model is fast, simple and non-parametric due to maintaining similar accuracy. An automated hierarchical probabilistic framework supported through the use of an adapted MRF framework and multi-window Gabor filters enable brain tumors to be segmented from multispectral brain MRI (Subbanna et al., 2013). BraTS database in the framework assists segmentation of brain tumor as edema and non-edema. The Figure 2 shows how labels of algorithm correspond to labels of expert closely.

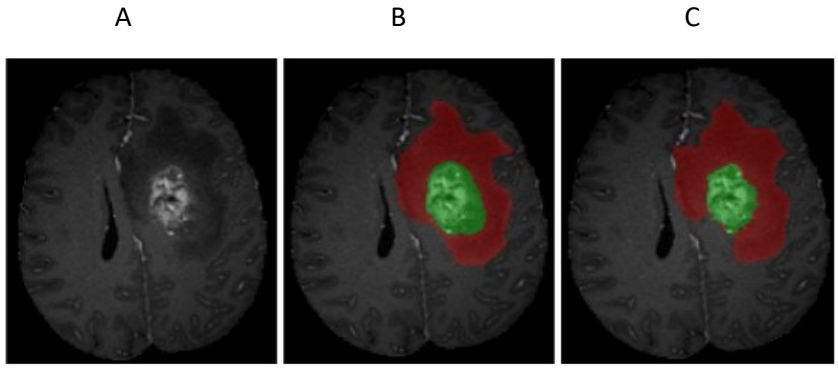


Figure 2. a) unlabeled TIC slice; b) expert labelling; c) algorithm labels (red – edema, green – non-edema)

2.5. SVM algorithms

Vladimir N. Vapnik invented the original version of SVM algorithm, but Cortes and Vapnik in 1993 studied current standard incarnation. To deal with supervised classification issues, a parametrically kernel-based method was proposed as SVM and brain tumor segmentation was a commonly known field for SVM algorithms. One-class SVM examines the ability to learn the nonlinear distribution of image data which uses no prior knowledge (Zhou et al., 2013). It is also applicable to achieve better segmentation results by following an implicit learning kernel and automatic procedure of SVM parameters training. These results support the extraction of brain tumors for better segmentation results compared to fuzzy clustering method. The researchers were willing to build voxel-wise intensity-based feature vectors via a high number of MRI modalities classified by SVM (Verma et al., 2008). The healthy tissues and also sub-compartments of healthy and tumor regions are segmented by this method, but similar approach based on SVM used a lower number of modalities and segmented one tumor region. The feature selection with kernel class was introduced to improve this method and it showed better results (Ruan et al., 2011). In order to segment the brain tumor from multi-sequence MRI images a fusion process and a multi-kernel based SVM in collaboration with a feature selection was offered as an alternative. Ameliorating the contour of tumor region (use of the distance and the maximum likelihood measures) and classification of tumor region (use of a multi-kernel SVM) are major two steps of a multi-kernel based SVM integrated with fusion and feature selection processes. In addition to classifying the tumor region, this step focuses on the performance of multi-image sources and multi-results. The accuracy and diminution of total error are expected results obtained through this method compared to traditional version of single kernel SVM. A fully automatic method was also essential to define segmentation of brain tissue which used multispectral intensities for SVM classification as a combination (Bauer et al., 2011). This method textured with hierarchical approach (specifically, subsequent hierarchical regularization) based on CRF to get acceleration and robustness. Thus, better results and accuracy can be achieved by different levels of regularization at various stages. The Figure 3 denoted potentially useful effects of SVM algorithms for MRI images in segmentation of brain tumor. The Table 2 explains the relatively good methods and their presentations of MRI-based brain tumor segmentation.

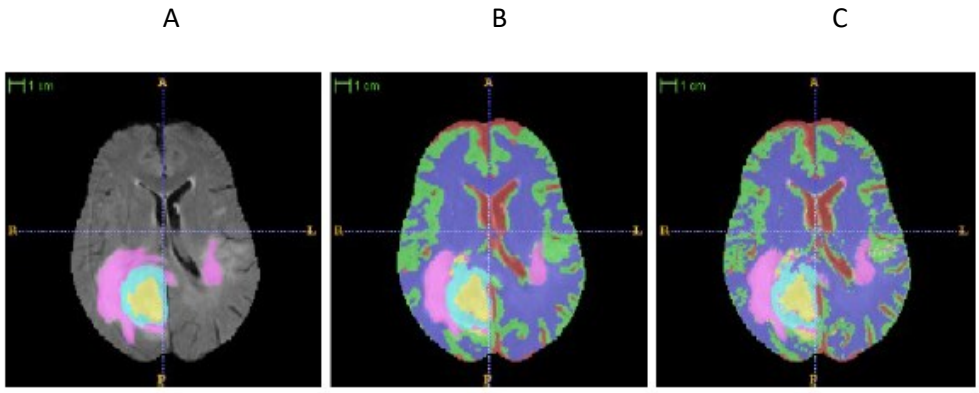


Figure 3. manual segmentation. a) hierarchical SVM-classification with CRF-regularization; b) non-hierarchical SVM-classification without regularization

Table 2. MRI-based brain tumor segmentation

Method	**Presentation**	**References**
Combination of k-means and fuzzy c-means	Better accuracy and reproducibility	Gupta et al., 2013
FKSRG	Lower over- and under-segmentation	Lin et al., 2012
Multi-region + multi-reference framework	Higher tissue overlaps rates and lower standard deviations	Shi et al., 2010
Generative probabilistic model + spatial regularization	Improvement over the traditional multivariate tumor segmentation (25 glioma)	Menze et al., 2010
Probabilistic model + localization	More robust applied to monitor disease progression	Weizman et al., 2012
Non-rigid registration + atlas + MRF	Multivariate tumor segmentation	Menze et al., 2010
SVM + CRF	10 multispectral patient datasets more detail segmentation low computation times	Bauer et al., 2011
Decision Forests + tissue-specific Gaussian mixture models	Segmenting the individual tissue types simultaneously such as AC, NC, E, etc.	Zikic et al., 2012
SVM + Kernel feature selection	Good results tested in T1w, T2w and T1c, low computation time	Zhang et al., 2011

2.6. Deformable model methods

Parametric and geometric deformable models are major components of model-based segmentation methods that can handle the issue of the appearance of 3-D MRI data

with respect to extraction of boundary elements which belong to same structure and integration of those elements into a consistent and coherent structure model (Lorenzo et al., 2013). In most cases, such issues are challenging to be segmented through simple methods compared with combination of SVM classification techniques which demonstrate high accuracy and diminution of total error. The capability of segmenting images of anatomic structures helps to determine resistance of deformable models. This segmentation exploits constraints about size, shape and location of anatomic structures which stem from image data with a prior knowledge. Deformable models are highly adjustable for the variability of biological structures in terms of various individuals. In addition to deformable models, they are more likely to assist clinicians and medical researchers through intuitive interaction mechanisms in order to determine necessary model-based image interpretation task.

2.7. Parametric deformable models

These models were so-called active contour models and snakes in some periods. After introduction of snakes in 1988, they are used to locate object contours including appropriate initialization in practice. The snakes are capable and sensitive to detect the boundary of brain tumors that is highly significant as a step of brain tumor segmentation in parametric deformable models. According to studies done about the resolution of the boundary, the snake technique shows more effective results compared to conventional edge detection including Canny, Sobel and Laplacian algorithms. However, the snake function in homogeneous regions is positively obtained while it is zero at the edges. Therefore, the improvement of brain tumor segmentation results on T1 brain tumor MRI was achieved using the balloon model and the Gradient Vector Flow (GVF). GCF aimed to analyze inability and short capture range which stem from the traditional snakes to track concavity of boundary. Moreover, the spatial relations as refinement step enable a parametric deformable model to estimate boundaries of any type of brain tumors accurately which is on T1 MRI (Gooya et al., 2011). The growth of snakes capture range can be defined by the balloon model apart from the parametric deformable model. The combination of deformable registration and segmentation of brain scans was proposed based on Expectation Maximization (EM) algorithm to a normal atlas in a way to explain incorporation of atlas seeding and a glioma growth model. This process studies modified atlas which represents the normal atlas into one with edema and a tumor. In addition, utilization of the posterior probability estimation of different tissue labels and registration into the patient space are essential characteristics of the modified atlas. The Expectation Maximization algorithm is highly optimistic to refine the posterior probabilities of tissue labels, the tumor growth model parameters and the estimates of registration parameters. It is also necessary to note that manual location of initial position of the parametric

deformable model demonstrates avoidance of converging to wrong boundaries if it is close to desired boundary.

2.8. Geometric deformable models

Geometric Deformable Models (GDM) is sometimes known as level sets which improves topological changes for merging of contours and splitting processes. In fact, these procedures are more challenging to be handled naturally in terms of topological changes when using segmentation of 3D MRI data through parametric deformable models. In most cases, segmentation methods of brain tumor cannot be easily achieved in practice when dealing with regularly shaped objects. On the other hand, the issue stems from improvement of initialization of parametric active contours and symmetrical placement of initial contour with respect to boundaries of interest. Level-set snakes were highly preferable to gain an advantage compared to mathematical morphology and conventional statistical classification, because snakes experience careful initialization which have constant propagation and leak through missing boundary parts. A knowledge-based segmentation algorithm combines level-set snakes and pixel-intensity distribution that present more precise boundaries. Some researchers examined a deformable model using a Charged Fluid Framework (CFF) to aim brain tumor segmentation for a certain period of time (Prastawa et al., 2004). However, CFF was extended and modified for brain tumor segmentation by proposing the Charged Fluid Model (CFM) (Chang & Valentino et al., 2006). Brain tumor can be segmented in a variable level set formulation by proposing a region-based active contour model (Li et al., 2008). This model suggested that the image intensities were approximated on two sides of contour by two fitting functions originated from data fitting energy. A regulation term as a part of the level set formulation shows derivation of a curve evolution which potentially targets energy minimization. The level set regularization term preserved regularity of level set function in a way to eliminate expensive re-initialization. In this case, the progress of level set function depends on accurate computation of the level set regularization term. A few researchers determined a local clustering criterion function which is specified for intensities considered in a neighborhood of each point (Li et al., 2011). However, a local intensity clustering property was also studied through brain tumor and other images with respect to intensity inhomogeneities. Integration of the local clustering criterion supports an energy functional over the neighborhood center in order to convert to formulation of the level set. Estimation of bias field and level set evolution with an interleaved process led to minimization of the energy (Hamamci et al., 2012). Combination of tumor segmentation and level set including tumor probability was achieved by a tumor-cut algorithm which resulted spatial smoothness.

To sum up, FCM, ANN and MRF are commonly used algorithms in terms of deformable model analysis. This method can aim the accuracy of brain tumor segmentation by incorporating two or more algorithms. Therefore, brain tumor segmentation has a direct effect on medical image analysis for surgical planning that is most important issue in term of validity of segmentation process. As a part of tumor segmentation, commonly applied evaluation standards include Dice Similarity Coefficient (DSC) and Jaccard coefficient which range from 0 indicating no overlap to 1 indicating perfect overlap (Crum et al., 2006). Moreover, probabilistic brain tumor segmentation was analyzed through three following validation metrics; Mutual Information (MI), Dice Similarity Coefficient (DSC) and the Receiver Operating Characteristic curve (ROC) (Archip et al., 2007). These methods aim to sustain tumor monitoring, a preliminary judgment on diagnosis as well as the physician with therapy planning.

2.9. Trilinear Interpolation Algorithm Techniques for 3D MRI Brain Image

There are many diagnoses and detection imaging techniques for treatment of potential risks caused by brain tumor diseases (Athertya & Poonguzhali, 2012). Computed Tomography (CT), Scanner, Position Emission Tomography (PET) and Magnetic Resonance Imaging (MRI) are commonly recognized diagnostic imaging tools to be practiced for tumor segmentation. In recent year, researches show that the process of obtaining 3D images from 2D medical images becomes increasingly necessary to determine an appropriate identity and regional development of tumors. Moreover, Machine Cube supports 2D CT image constructed 3D spine image or 3D surface of knee (Patel & Mehta, 2012). In this method, data blocks are divided into cubes that were made up of eight adjacent voxels. By using the triangular mesh material surfaces were constructed from there eight adjacent voxels. As a result, simple construction operations, 3D image production with high resolution and fast calculation are advantageous procedures experienced through this method. If the large number of 2D image data is processed, the calculation process slows down. In most cases, noise occurred due to images captured from sensor should be reduced by pre-processing 2D images in order to construct the 3D image with high quality of improvement. Therefore, a mean-unsharp filter is able to increase filtered noise and high frequency components. Before low-level image separation the intensity values of grayscale image may process the growth method of MRI images (Kumaran & Thimmiaraja, 2014). In this case, low-contrast images can be easily changed to higher-contrast images. Furthermore, making the tumor boundary depends on how morphological operations are utilized in order to stretch and fill a possible object for segmentation process. In this section, we will determine Otsu method for constructing 3D image with a support of segmentation of 2D images by finding the gray level threshold values.

On the other hand, segmentation method can be jointly connected with the algorithm of regional development with respect to similarity of adjacent pixels related to the nuclear point. Choosing the error and nuclear initial point which depend on between neighboring pixels and nuclear point can be defined as a result of combination of algorithm (Ho et al., 2016). Practically, the acquisition of 2D medical image may process noise on it, but Otsu technique as a threshold of traditional method segment some of the areas required for 3D image development. According to some research studies done on Otsu method, it may not often demonstrate better results. Instead, dividing 2D image into many layers by Otsu method with multilevel may provide better and efficient results. Pixels in 2D image are combined with region development algorithm to process segmentation of the image in which pixels have same regions (Zhang et al., 2016). The process may continue until the 2D image reaches coordinate axes (x, y, z) after segmentation. The calculation of approximate value of a point allows to construct the 2D image surface between two consecutive layers which are represented in the spatial domain. In fact, the linear interpolation approves this construction of 2D image surface before calculation of a value of the point (Hagan et al., 2009).

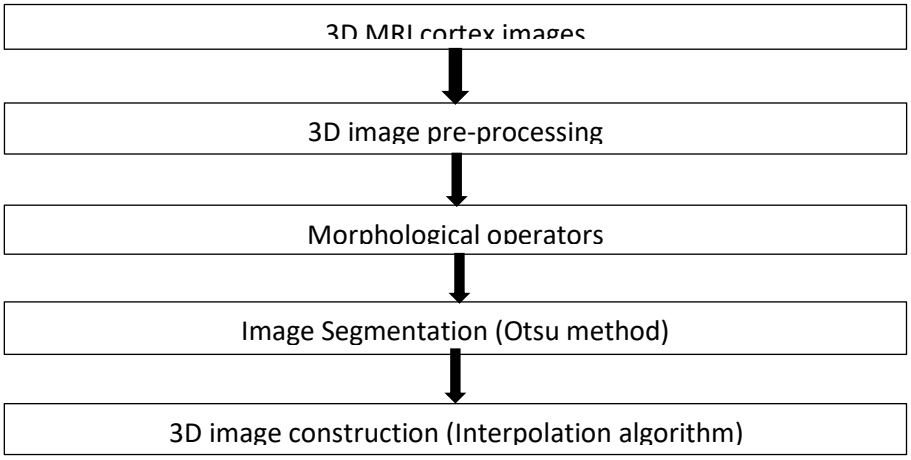


Figure 4**.** Otsu method or image segmentation

Figure 4. 3D image construction

Binh Duong General Hospital provided dataset for construction of 3D image by considering 44 2D MRI brain images with 256x192 pixels. There are some major following steps to describe the construction of 3D image from 2D MRI human cortex images. First step is to re-process 2D image including image enhancement and noise rejection. Second step is to employ morphological operator and eliminate pixels in 2D images around object boundaries. Third step is called Otsu method or image

segmentation where segmented and pre-processed images for 3D construction separate the brain area from the cortex. Final step is to construct 2D images after the segmentation to obtain 3D image by using a trilinear interpolation algorithm (also known as 3D image construction) figure 4.

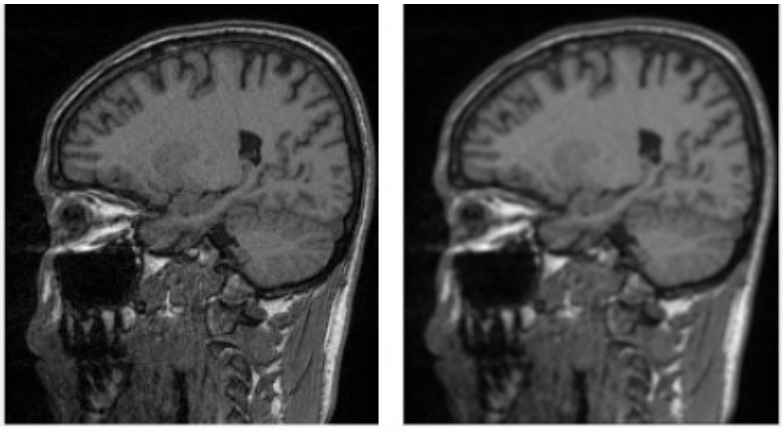
2. Image pre-processing

Image pre-processing follows an average filter process to smooth 2D MRI images with noise in order to construct 3D image. The equation below describes elimination of noise by applying the average filter that is convoluted with each 2D MRI image.

$$r(x,y) = \frac{1}{ab} \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x+s, y+t)$$

- r(x,y) is output image after the process of filtering;
- w(s,t) is axb filter window.

In figure A, the average filter with a 3-by-3 kernel after convolution is given as the result of 2D MRI cortex image $\frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.



5- 1 5- 2

Figure 5. 5-1original human cortex image; 5-2 cortex image after the mean filter

After separation of complete brain region form the cortex image, the convoluted image provides smoother processing results in terms of using mean filter compared with using the median filter.

6-1 6-2

Figure 6. 6-1 image after using the mean filter; 6-2 image after using the median filter

In Figure C, not only enhancement of image with high frequency components but removal of low frequency components may be realized through application of unsharp filter with 3x3 size.

$$\frac{1}{2}\begin{pmatrix} -1 & 0 & -1 \\ 0 & 6 & 0 \\ -1 & 0 & -1 \end{pmatrix}$$

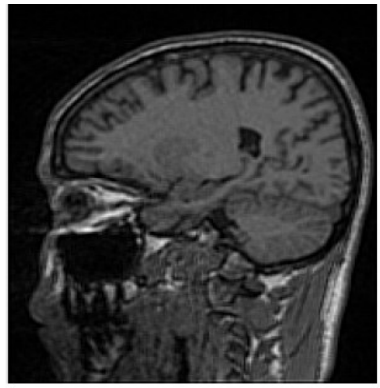


Figure 7. 2D MRI cortex image after the unsharp filter

Moreover, image enhancement is a solution to make object better on higher contrast in case of low contrast after image filtering. Spreading the pixel values in the image is a significant procedure to expect transformation of the higher contrast image when proposing a histogram equalization algorithm (Kaur & Rani, 2016). This algorithm uses a special equation to enhance the image as follows:

$$p_r(k) = \frac{n_k}{MN}$$

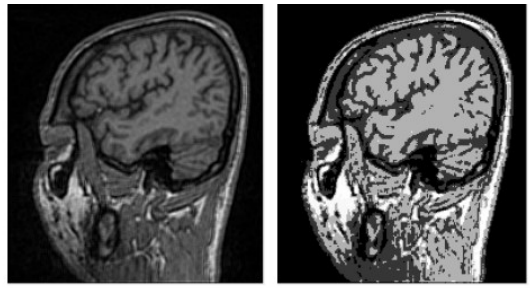
- nk is the number of pixels observed at the kth gray level in terms of the input image;
- MN is total pixels of the image;
- Pr(k) denotes the probability density function in the image connected with the kth gray level values.

Apart from probability density function (PDF), the calculation of the output expression of the image is defined:

$$s = (L-1)\sum_{j=0}^{k} p_r(j)$$

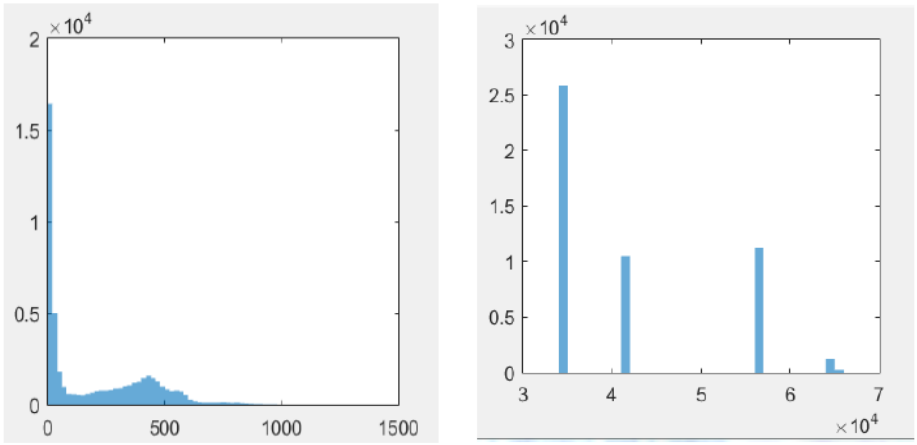
- L denotes the number of gray levels in the image;
- S(x,y) denotes the number of the pixels with respect to output image at the kth gray level.

In Figure D, first image represents before enhancement issue which uses unsharp filtering process. While, second image clearly shows after enhancement procedure. Nevertheless, the histogram equalization stimulates unsharp image with high frequency components to be enhanced for creation of higher contrast. Hence, figure 1 and 2 demonstrates representation of images where image before enhancement experiences the large number of pixels with distribution of being closed to the zero point. However, image after enhancement offers the histogram equalization with pixel values spreading on gray level axis.



8-1 8-2

Figure 8. 8-1 image using the unsharp filter; 8-2 image after enhancement



9-1 9-2

Figure 9. 9-1 image before enhancement; 9-2 image after enhancement using the histogram equalization

4. Morphological Operation

The enhanced image continues the process for image imperfection after filtering 2D MRI cortex image, but this filtering removes noise and enhances the image. As a result of the image enhancement, a morphological algorithm removes a few unnecessary parts around objects for the image. Only potentially important object is remained for 3D image construction by morphological algorithm. In addition, Dilation and Erosion are common operations of this algorithm in which the dilation operator considers structures and shapes of 2D image enhancement. At this moment, the morphological operation removes the undesired parts and combines boundaries around the objects (Sahar et al., 2016). On the other hand, the morphological image determines the convolution between a kernel and the input image for image calculation as given below:

$$m(x,y) = max\begin{cases} s(x-i,y-j) + h(i,j) \\ |(x-i),(y-j) \in D_s, (i,j) \in D_h \end{cases}$$

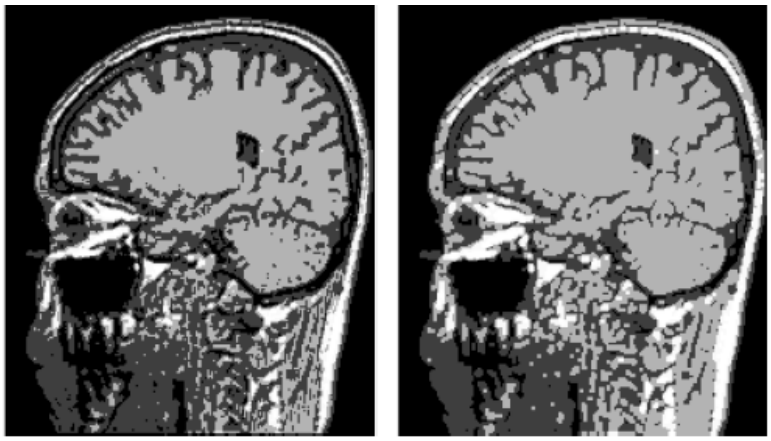
- h(i,j) is a kernel with the size of 2x2 matrix
- Ds and Dh are the domains of s image and k kernel.
- S(x-i,y-j) denotes output images.
- M(x,y) denotes input images.

In figure 10, the matrix of the 2x2 kernel is structured as follows:

1	1
1	1

Figure 10. matrix of the 2x2

In figure 11-2, the dilation operator in morphological algorithm enlarges the boundaries of object regions when producing the 2D MRI cortex image. For this reason, image after morphological operation combines the boundaries of object regions and provides better results compared with the enhanced image given as in figure 11-1.



11-1 11- 2

Figure 11. 11-1 representation of the enhanced image; 11-2 Image after the morphological operation

References

Archip, N., Jolesz, F. A. & Warfield, **S. K.** (2007). A validation framework for brain tumor segmentation, Academic Radiology, vol. 14, no. 10, pp. 1242-1251.

Athertya, J. & Poonguzhali, S. (2012)."3D CT Image Reconstruction of the Vertebral Column", International Conference on Recent Trends in Information Technology, pp. 81 - 84.

Bauer, S., Nolte, L. P. & Reyes, M. (2011). Segmentation of brain tumor images based on atlas-registration combined with a markov-random-field lesion growth model, in

Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, IEEE, pp. 2018-2021.

Bauer, S., Nolte, L. P. & Reyes, M. (2011). Fully-automatic segmentation of brain tumor images using support vector machine classification in combination with hierarchical conditional random field regularization, in Medical Image Computing and Computer-Assisted Intervention-MICCAI 2011. Springer, pp. 354-361.

Bishop, C. M. (2006). Pattern Recognition and Machine Learning, vol. 1. Springer New York.

Brodmann, K. (1909**).** Vergleichende Lokalisationslehre der Gro hirnrinde. Springer.

Cai, W., **Chen, S. & Zhang, D.** (2007). Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation, Pattern Recognition, vol. 40, no. 3, pp. 825-838.

Cai, H, **Verma, R, Ou, Y., Lee, S. K., Melhem, E. R. & Davatzikos, C.** (2007). Probabilistic segmentation of brain tumors based on multi-modality magnetic resonance images, in Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on, IEEE, pp. 600-603.

Chapelle, O., **Schoellkopf, B. & Zien, A.** (2006). Semi-Supervised Learning, vol. 2. MIT Press Cambridge.

Chan, T. F. & Vese, L. A. (2001) Active contours without edges, Image Processing, IEEE Transactions on, vol. 10, no. 2, pp. 266-277.

Chang, H. H. & Valentino, D. J. (2006). Image segmentation using a charged fluid method, Journal of Electronic Imaging, vol. 15, no. 2, pp. 023011-023011.

Christakou, C., **Lefakis, L., Vrettos, S. & Stafylopatis, A.** (2005). A movie recommender system based on semi-supervised clustering, in Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on, IEEE, vol. 2, pp. 897-903.

Cuadra, M. B, **Pollo. C, Bardera, A, Cuisenaire, O, Villemure, J. G. & Thiran, J. (**2004**).** Atlas-based segmentation of pathological mr brain images using a model of lesion growth, Medical Imaging, IEEE Transactions on, vol. 23, no. 10, pp. 1301-1314.

Crum, W. R, Camara, O. & Hill, D. L. (2006). Generalized overlap measures for evaluation and validation in medical image analysis, Medical Imaging, IEEE Transactions on, vol. 25, no. 11, pp. 1451-1461.

Duda, R. O., **Hart, P. E. & Stork, D. G.** (2012). Pattern Classification. John Wiley & Sons.

Evans. A, Marrett, S., Torrescorzo, J., Ku. S. & Collins. L. (1991). Mri-pet correlation in three dimensions using a volume of Interest (voi) atlas, Journal of Cerebral Blood Flow & Metabolism, vol. 11, pp. A69-A78.

Fletcher-Heath, L. M, **Hall, L. O., Goldgof, D. B. & Murtagh, F. R.** (2001). Automatic segmentation of non-enhancing brain tumors in magnetic resonance images, Artificial Intelligence in Medicine, vol. 21, no. 1, pp. 43-63.

Forouzanfar, M., **Forghani, N. & Teshnehlab, M. (**2010**).** Parameter optimization of improved fuzzy c-means clustering algorithm for brain mr image segmentation, Engineering Applications of Artificial Intelligence, vol. 23, no. 2, pp. 160-168.

Garc´ıa-Lorenzo, D., **Francis, S., Narayanan, S., Arnold. D. L. & Collins, D. L.** (2013). Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging, Medical Image Analysis, vol. 17, no. 1, pp. 1-18.

Gupta, M. P. & Shringirishi, M. M. (2013) Implementation of brain tumor segmentation in brain mr images using k-means clustering and fuzzy c-means algorithm, International Journal of Computers & Technology, vol. 5, no. 1, pp. 54-59.

Gooya, A., **Pohl, K. M., Bilello, M., Biros, G. & Davatzikos, C.** (2011). Joint segmentation and deformable registration of brain scans guided by a tumor growth model, in Medical Image Computing and Computer-Assisted Intervention-MICCAI 2011. Springer, pp. 532-540.

Hagan, R. (2009). "Numerical Methods for Isosurface Volume Rendering", Virginia Polytechnic.

Hamamci, A., Kucuk, N., Karaman, K., Engin, K. & Unal, G., Tumor-cut. (2012). Segmentation of brain tumors on contrast enhanced mr images for radiosurgery applications, Medical Imaging, IEEE Transactions on, vol. 31, no. 3, pp. 790-804.

Hastie, T., **Tibshirani, R., Friedman, J. & Franklin**, **J.** (2005). The elements of statistical learning: Data mining, inference and prediction, The Mathematical Intelligencer, vol. 27, no. 2, pp. 83-85.

Ho, Y., **Lin, W., Tsai, C., Lee, C. & Lin, Chih, Y.** (2016). "Automatic Brain Extraction for T1 – Weighted Magnetic Resonance Images Using Region Growing", International Conference on Bioinformatics and Bioengineering, pp. 250 - 253.

Kaur, H. & Rani, J. (2016). "MRI brain image enhancement using Histogram equalization Techniques", International Conference on Wireless Communications, Signal Processing and Networking, pp. 770 - 773.

Kumaran, N. S., & Thimmiaraja, H. (2014). "Histogram Equalization for Image Enhancement Using MRI brain images", WCCCT, pp. 80 - 83.

Li, C., **Kao, C. Y., Gore, J. C. & Ding, Z.** (2008). Minimization of region-scalable fitting energy for image segmentation, Image Processing, IEEE Transactions on, vol. 17, no. 10, pp. 1940-1949.

Li, C., Huang, R., Ding, Z., Gatenby, J., Metaxas, D. N. & Gore, J. C. (2011). A level set method for image segmentation in the presence of intensity inhomogeneities with application to mri, Image Processing, IEEE Transactions on, vol. 20, no. 7, pp. 2007-2016.

Lin, G. C., **Wang, W. J., Kang, C.C. & Wang, C. M**. (2012). Multispectral mr images segmentation based on fuzzy knowledge and modified seeded region growing, Magnetic Resonance Imaging, vol. 30, no. 2, pp. 230-246.

Menze, B. H., **Van Leemput, K., Lashkari, D., Weber, M. A., Ayache, N. & Golland, P.** (2010). A generative model for brain tumor segmentation in multi-modal images, in Medical Image Computing and Computer-Assisted Intervention- MICCAI 2010. Springer, pp. 151-159.

Mitchell, T. M. (2006). The discipline of machine learning, Carnegie Mellon University, School of Computer Science.

Nie, J., **Xue, Z., Liu, T., Young, G. S., Setayesh, K., Guo, L. & Wong, S. T.** (2009). Automated brain tumor segmentation using spatial accuracy-weighted hidden markov random field, Computerized Medical Imaging and Graphics, vol. 33, no. 6, pp. 431-441.

Patel, A. & Mehta, K. (2012). "3D Modeling and Rendering of 2D Medical Image", International Conference on Communication Systems and Network Technologies, pp. 149 - 152.

Pham, D. L., **Xu, C. & Prince, J. L. (**2000**).** Current methods in medical image segmentation 1, Annual Review of Biomedical Engineering, vol. 2, no. 1, pp. 315-337.

Prastawa, M., **Bullitt, E., Ho. S. & Gerig, G.** (2004). A brain tumor segmentation framework based on outlier detection, Medical Image Analysis, vol. 8, no. 3, pp. 275-283.

Ruan, S., Lebonvallet, S., Merabet, A. & Constans, J. (2007). Tumor segmentation from a multispectral mri images by using support vector machine classification, in Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on, IEEE, pp. 1236-1239.

Ruan, S., **Zhang, N., Liao, Q. & Zhu, Y.** (2011). Image fusion for following-up brain tumor evolution, in Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, IEEE, pp. 281-284.

Sahar, M., **Nugroho, H. A., Tianur, A. I. & Choridah, L.** (2016). "Automated Detection of Breast Cancer Lesions Using Adaptive Thresholding and Morphological Operation", International Conference on Information Technology Systems and Innovation, pp.1 - 4.

Shiç F., **Yap, P.T., Fan, Y., Gilmore, J. H., Lin, W. & Shen, D.** (2010). Construction of multi-region-multi-reference atlases for neonatal brain mri segmentation, Neuroimage, vol. 51, no. 2, pp. 684-693.

Shenton, M., Kikinis, R., McCarley, W., Saiviroonporn, P., Hokama, H., Robatino, A., Metcalf, D., Wible, C., Portas, C. & Iosifescu, D. (1995). Harvard brain atlas: A teaching and visualization tool, in Proceedings of the 1995 Biomedical Visualization, pp. 10-17.

Stelldinger, P., Latecki, L. J. & Siqueira, M. (2007**).** "Topological Equivalence between a 3D Object and the Reconstruction of Its Digital Image", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 126 - 140.

Subbanna, N. K., **Precup, D., Collins, D. L. & Arbel, T.** (2013). Hierarchical probabilistic gabor and mrf segmentation of brain tumours in mri volumes, in Medical Image Computing and Computer-Assisted Intervention-MICCAI 2013. Springer, pp. 751-758.

Szilagyi, L., **Benyo, Z., Szil´agyi, S. M. & Adam**, **H.** (2003). Mr brain image segmentation using an enhanced fuzzy cmeans algorithm, in Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE, IEEE, vol. 1, pp. 724-726.

Szilagyi, L., **Szil´agyi, S. M. & Benyo, Z.** (2007). A modified fuzzy c-means algorithm for mr brain image segmentation, in Image Analysis and Recognition. Springer, pp. 866-877.

Talairach, J. & Tournoux, P. (1988). Co-Planar Stereotaxic Atlas of the Human Brain: 3-dimensional proportation System an Approach to Cerebral Imaging. New York, USA: Thieme Medical Publishers.

Tran, T. N., **Wehrens, R. & Buydens, L.** (2005). Clustering multispectral images: A tutorial, Chemometrics and Intelligent Laboratory Systems, vol. 77, no. 1, pp. 3-17.

Verma, R., **Zacharaki, E. I., Ou, Y., Cai, H., Chawla, S., Lee, S. K., Melhem, E. R., Wolf, R. & Davatzikos, C.** (2008). Multiparametric tissue characterization of brain neoplasms and their recurrence using pattern classification of mr images, Academic Radiology, vol. 15, no. 8, pp. 966-977.

Weizmann, L., **Ben Sira, L., Joskowicz, L., Constantini, S., Precel, R., Shofty, B. & Ben Bashat, D.** (2012). Automatic segmentation, internal classification, and follow-up of optic pathway gliomas in mri, Medical Image Analysis, vol. 16, no. 1, pp. 177-188.

Zhang, N., Ruan, S., Lebonvallet, S., Liao, Q. & Zhu, Y. (2009). Multi-kernel SVM based classification for brain tumor segmentation of MRI multi-sequence, in Image Processing (ICIP), 2009 16th IEEE International Conference on, IEEE, pp. 3373-3376.

Zhang, N., **Ruan, S., Lebonvallet. S., Liao, Q. & Zhu, Y.** (2011). Kernel feature selection to fuse multi-spectral MRI images for brain tumor segmentation, Computer Vision and Image Understanding, vol. 115, no. 2, pp. 256-269.

Zhang, X., Li, X., Li, H. & Feng, Y. (2016). "A Semi – Automatic Brain Tumor Segmentation Algorithm", International Conference on Multimedia and Expo, pp. 1 - 6.

Zhou, J., Chan, K., Chong, V. & Krishnan, S. (2006). Extraction of brain tumor from mr images using one-class support vector machine, in Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the, IEEE, pp. 6411-6414.

Zikic, D., **Glocker, B., Konukoglu, E., Criminisi, A., Demiralp, C., Shotton, J., Thomas, O., Das, T., Jena, R. & Price, S.** (2012). Decision forests for tissue-specific segmentation of high-grade gliomas in multi-channel mr, in Medical Image Computing and Computer-Assisted Intervention-MICCAI 2012. Springer, pp. 369-376.

Khazar Journal of Science and Technology Volume 5 №2 2021, 71-76

DOI: 10.5782/2520-6133.2021.5.2.71

Investigation of SARS-Cov-2 Infection in Domestic Animals

Qiyasaddin Jalladov[1], Aytan Hajiyeva[1], Sabina Mammadova[1], Chichak Aliyeva[1,2], Shabnam Mammadova[1], Shumshad Rustamli[1], Kamala Aliyeva[1], Rana Kangarli[1], Fakhranda Alizade[1], Narmin Akhundova[1]

[1]Scientific Research and Risk Assessment Division, Scientific Research Center Department
Azerbaijan Food Safety Institute, Azerbaijan
[1,2]Life Sciences department, Khazar University
[*]Corresponding author: ciceksuleymanova3@gmail.com

Abstract

Concerned with the COVID-19 pandemic, the study of this disease in animals has got a great scientific importance in clarifying the information about the source and circulation of the infection. The study aimed to investigate the source of infection of domestic animals (dog, cat, cattle, sheep, goat and poultry) with SARS-CoV-2, as well as to identify susceptible animal species and ways of transmission of the virus. Observations were made on the animals selected for the study, from which nasopharyngeal and oropharyngeal smears were taken for PCR, and blood samples were taken for enzyme-linked immunosorbent assay (ELISA).

The experimental part of the study was carried out in veterinary clinics, animal shelters and farms. Dogs and cats are kept in animal shelters and examined in veterinary clinics, as well as cattle, sheep, goats and poultry grown on various farms, were involved in the study.

Antibodies to SARS-CoV-2 were detected in 11 of 645 samples taken from animals whose clinical signs of COVID-19 disease were initially observed or whose owners were exposed to the disease.

Based on the results of the study, monitoring the dynamics of the spread of SARS-CoV-2 among animals is of great scientific and practical importance in preventing this process.

Keywords: Coronavirus disease 2019 (COVID-19), severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), investigation, antibody, domestic animals.

Introduction

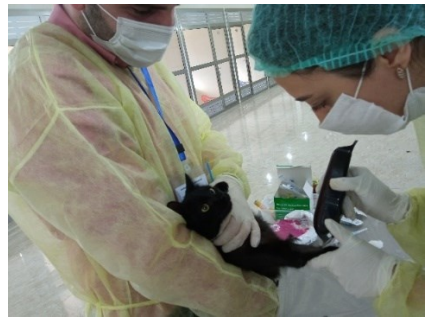
SARS-CoV-2 infection, caused the COVID-19 pandemic in the world and atypical pneumonia in humans, was first detected in late 2019 in the Chinese city of Wuhan and was named 2019-nCoV (Yu et al., 2020). The genome structure of the causative agent is homologous to 50% MERS-CoV, 79% SARS-CoV and 88% BtRsCoV, so it belongs to the second type of acute respiratory syndrome viruses (Pal et al., 2020). According to recent data, SARS-CoV-2 is classified as a highly mutagenic virus belonging to the Beta- and Deltacoronavirus groups of the *Orthocoronavirinae* subfamily of the *Coronaviridae* family (Zhou et al., 2020).

Although the exact source of the causative agent of infection is not known, it is believed that SARS-CoV-2 was originated in wild animals and, later, transmitted to humans (Ahmed et al., 2020). According to preliminary data, the spread of the disease was caused by bats sold at animal markets in the Chinese city of Wuhan (Mackenzie & Smith, 2020). In the later stages of the pandemic, rodents were also reported to be infected with Sars-CoV-2 in various parts of the world. Rodents are thought to have played an intermediate role in transmitting the virus from bats to humans (Yuan et al., 2020).

Although there are no solid scientific findings on the mechanism of transmission of SARS-CoV-2 from animals to humans or vice versa, in some countries, owners of dogs and cats infected with SARS-CoV-2 have been found to have the disease (Sit et al., 2020; Calvet et al., 2021). In the United States, there have also been reports of the disease being transmitted to zoo animals through contact with an employee infected with SARS-CoV-2 (USDA, 2021). In late April 2020, Netherlands reported the first case of SARS-CoV-2 infection in commercially farmed mink. Additional reports of SARS-CoV-2 infected farmed mink came later in 2020 from Denmark (June), Spain (July), the U.S. (August), Italy (August), Sweden (October), France (November), Greece (November), Lithuania (November), and Canada (December) (AVMA, 2021).

Material and methods

Dogs and cats are kept in animal shelters and examined in veterinary clinics, as well as cattle, sheep, goats and poultry grown on various farms, were involved in the study. At the initial stage of the study, blood samples were taken from animals (nasopharyngeal and oropharyngeal smears) for molecular genetic (PCR) examinations, and blood from peripheral veins for enzyme-linked immunosorbent assay (ELISA) examinations.



Samples were taken from 211 dogs, 136 cats, 19 cattle, 268 ruminants (sheep, goats) and 11 chickens. PCR and ELISA tests (virus in swab samples, antibodies in blood samples against the virus) were performed on samples taken from dogs and cats, and ELISA tests (antibodies in blood samples) were performed on samples taken from other animals. Blood samples were collected via leg venipuncture and sera were separated and stored at −20°C until further processing. All samples were collected under full personal-protective equipment.

Molecular genetic analyzes were performed using the BIO-RAD CFX96 Real-Time device. Extraction process done by the QIAamp® Viral RNA Mini kit. During the Extraction process, 140µl of the sample was taken at the initial stage and 160µl of RNA was isolated. Purification carried through Oasig™ lyophilized OneStep 2X RT-qPCR Master Mix kit, 5µl of extracted RNA was added to 20µl of Master mix. In the last stage, 25µl of the mixture was placed on the device for reading and recording.

Enzyme-linked immunosorbent assay (ELISA) analyzes were performed using the Thermo Multiskan FC device using the ID Screen SARS-CoV-2 Double Antigen Multi-species kit. Double antigen ELISA for the detection of antibodies directed against the nucleocapsid of SARS-CoV-2 in animal serum or plasma. Before the start of analysis, all kit reagents were kept at room temperature (21°C± 5°C) and homogenized by vortexing. Firstly, 25µl Dilution buffer was added to each well. Secondly, an equal amount of Negative and Positive control was added properly to wells A1 and B1, C1 and D1. Then, 25µl sample was added to each remaining wells. The plate was covered and placed in a thermostat for 45 minutes of incubation. The plate was emptied and washed 3 times with 300µl wash solution. Then 300µl conjugate was added to each well. The plate was covered and kept at room temperature for 30 minutes. Again the plate was emptied and washed 3 times with 300µl wash solution. In the next step, 100µl of substrate solution was added. The plate was covered and kept in the dark condition at room temperature for 20 minutes. In the end, 100µl of stop solution was added. The results were read and recorded at 450 nm.

Result and discussion

It should be noted that no SARS-CoV-2 agent was detected in the smear samples as a result of the research. No antibodies to SARS-CoV-2 were detected in immunosorbent assays analysis in blood samples from cattle, sheep, goats, and poultry. However, blood samples taken from 7 dogs and 4 cats with symptoms of the disease (runny nose, cough, diarrhea, etc.) revealed the presence of antibodies against the virus. The results demonstrate that the disease is limited among dogs and cats. Information on the type, age, residential area (owner/stray), and the amount (titer) of specific antibodies in the blood of dogs and cats, for which antibodies to SARS-CoV-2 have been detected, are presented in Tables 1 and 2.

Table 1. Amount (titer) of specific antibodies in the blood of dogs

Dog N^{o}	Gender of animal	Age	Sampling place	Titer (%)
1	Male	10 months	Animal shelter	143,58
2	Female	1 year	Animal shelter	157,18
3	Female	2 years	Animal shelter	64,24
4	Female	4 years	Vet. clinic	61,00
5	Female	3,5 months	Vet. clinic	237,86
6	Male	2 years	Vet. clinic	60,23
7	Male	1 years	Vet. clinic	98,97

The table 1 provides information on dogs with ≥60% antibodies in their blood. As can be seen, the number of dogs with positive results is 7(Dog numbers are conditional). Of these, 62.24% of titers were found in the 3rd dog, 61.00% in the 4th dog, and 60.23% in the 6th dog. The antibody titer is 98.97 in the 7th dog. The 1st, 2nd and 5th dogs were found to have the highest titers, which are 143.58%, 157.18% and 237.86%, respectively. (≥ 60% positive) *

The table 2 provides information on cats with ≥60% antibodies in their blood. As can be seen, the number of dogs with positive results is 4(Cat numbers are conditional). Of these, 104.39% of titers were found in the 1st cat, 162.19% in the 2nd cat, 226.58% in the 3rd and 153.63% in the 4th dog. (≥ 60% positive) *

347 samples were examined by PCR and the result was negative. This means that no active patients were recorded in the sampled dogs and cats.

* According to the instructions of the ID Screen SARS-CoV-2 Double Antigen Multi-species set, the result is considered positive if the antibody titer is ≥ 60%.

Table 2. Table 1. Amount (titer) of specific antibodies in the blood of cats

Cat Nº	Gender of animal	Age	Sampling place	Titer (%)
1	Female	2 years	Vet. clinic	104,39
2	Female	3 years	Vet. clinic	162,19
3	Male	1,5 months	Vet. clinic	226,58
4	Female	8 months	Vet. clinic	153,63

The results of the study show that dogs and cats are more susceptible to the disease among experimental animals, and the incidence of the disease among dogs is higher than among cats. Thus, the incidence was 3.32% among dogs and 2.94% among cats.

At the same time, the incidence of SARS-CoV-2 infection in females, both dogs and cats, is high, and the incidence of disease in domesticated animals is predominant (73%). Given the direct contact of dogs with antibodies in the blood of animals living in shelters, it can be assumed that there is a risk of transmission of the coronavirus from animal to animal.

The level of antibody titer, which is important in the formation of immunodeficiency, is inversely proportional to the age of the animals in the dog population, and the amount of titer produced decreases with age. In cats, however, it can be concluded that the antibody titer in the blood is not related to the age of the animal.

During the collection of survey data on infected animals, it was determined that there were people infected with COVID-19 or with clinical symptoms of the disease among the owners of the animals or the staff serving them in the shelters. This, in turn, indicates that the human factor plays a role in the circulation of the pathogen in the transmission of the pathogen to animals.

Our results including 7 dogs and 4 cats with symptoms of the disease (runny nose, cough, diarrhea, etc.) revealed the presence of antibodies against the virus.

Since no antibodies to SARS-CoV-2 or the pathogen have been found in cattle and ruminants (cattle, sheep and goats), as well as poultry, it can be concluded that these animals are insensitive to the disease. Cases of SARS-CoV-2 infection have been reported in domestic dogs and cats. Thus, the formation of specific antibodies in the blood of domestic animals was observed by immunosorbent assays. Exposure of animals to SARS-CoV-2 infection increases the likelihood of contact with their infected owners or infected caregivers. Therefore, while contacting with animals during the COVID-19 pandemic, people should strictly follow the sanitary and hygienic rules and take into account the possibility of cross-transmission of zoonotic

SARS-CoV-2. Research should also be continued to investigate the possibility of the virus circulating among animals and transmitting it from animals to humans.

More studies are needed to investigate the transmission route of SARS-CoV-2 from humans to cats and dogs. Finally, it is imperative that further studies be quickly carried out in order to better establish the risk of contamination of pets from humans, as well as the risk that infected pets would have as a source of infection for humans. Importantly, immediate action should be implemented to keep a suitable distance between humans and pet animals such as cats and dogs, and strict hygiene and quarantine measures should also be carried out for these animals.

References

Ahmad T., Khan M., Musa T.H., Nasir S., Hui, J., Bonilla-Aldana, D.K. & Rodriguez-Morales., A.J. (2020). COVID-19: Zoonotic aspects, 1-4.

AVMA. (2021). SARS-CoV-2 in animals, 1-6.

Calvet, G. A., Pereira, S.A., Ogrzewalska, M., Pauvolid-Corrêa, A., Resende, P.C., Tassinari, W.S., et al. (2021). Investigation of SARS-CoV-2 infection in dogs and cats of humans diagnosed with COVID-19 in Rio de Janeiro, Brazil, 4-8.

Mackenzie J.S. & Smith. D.W. (2020). COVID-19: a novel zoonotic disease caused by a coronavirus from China, 2-6

Pal, M., Berhanu, G., Desalegn, C. & Kandi. V. (2020). Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2), 1-14

Shchelkanov, M.Yu., Popov, A.Yu., Dedkov, V.G., Akimkin, V.G. & Maleev, V.V. (2020). The history of the study and modern classification of coronaviruses (nidovirales: coronaviridae), 221–246.

Sit, T.H.C., Brackman, C. J., Ip, S. M., Tam, K. W. S., Law, P. Y.T., et al. (2020). Infection of dogs with SARS-CoV-2, 586: 776–778

USDA. (2020-2021). List of SARS-CoV-2 Cases in the United States, 1-9

Yuan, S., Jiang, Si. C. & Li. Zi. L. (2020). Analysis of Possible Intermediate Hosts of the New Coronavirus SARS-CoV-2, 2-5

Zhou Z., Qiu, Ye. & Xingyi Ge. (2020). The taxonomy, host range and pathogenicity of coronaviruses and other viruses, 3-90.

Khazar Journal of Science and Technology Volume 5 №2 2021, 77-82

DOI: 10.5782/2520-6133.2021.5.2.77

The Effects of Water Stress on the Growth of Corn and the Activity Dynamics of NADP–Dependent Isocitrate Dehydrogenase Enzyme in the Leaf and Root T issues

Naila Aliyeva, Ziyeddin Mamedov
Department Biophisic and Biochemystry, Faculty of Biology, Baku State University, Azerbaijan
Corresponding Author: naila.aliyeva.bk.2018@gmail.com

Abstract

Developmental dynamics of the NADP-dependent isocitrate dehydrogenase enzyme has been studied (NADP-ICDH, EC 1.1.1.42) as the reducing potential of the cell for the development of corn sprouts through water stress conditions (water shortage and over water stress) which has a great role in forming NADPH pool. It was determined that the activity of NADP-ICDH enzyme increases both in root and leaf tissues significantly for the effects of water stress and over water stress induces the activity of enzyme more sharply than water shortage stress. The activity degree of enzyme in all three variants manifests itself higher in leaf tissues than root tissues. It seems NADP is required for the corn seedlings in order to eliminate stress complications and NADP-ICDH enzyme is directly involved in its synthesis.

Keywords: corn seedlings, water stress, over water, water shortage

Introduction

Water stress is manifested as water shortage and over water. This case includes main abiotic stress-based group that makes a serious ecological problem for the developmental productivity of plants. Water shortage is more widespread in nature than over water. One of the main reasons of water shortage is global heating. Plenty of water is used due to a 2 C^0 drop in temperature during a year. On the other hand water demand increases for the grows of population number year –after- year/ water demand speeds up year after the year due mainly to population growth increase. Climatologists predict that the number of the people under water stress will increase

twice for water demand in 2050 compared to 2010's. Nowadays water usage in agriculture accounts 70% of general water demand on the Earth. As well as 60 % increase in food demand will increase water usage too till 2050 year (Boretti et al., 2019). Salinization of suitable lands on the Earth causes the reduction of water storage capacity of the soil that results disruption of the plant's water supply. Sub-optimal water supply is disrupted for the physiology and biochemical changes in a plant during the water shortage (Jain et al., 2019). The main response to the water shortage of plants is aimed to protect it by reducing transpiration. Further, the synthesis of the abscisic acid increases and this causes the closing of mouthpieces (Munemasa et al., 2013). Although this and such other mechanisms can protect a plant preventing water loss on the first stage of the stress, generally metabolism in plants gets weak due to the weakening of photosynthesis in the next growing level of a plant and water shortage causes the destruction of a plant.

On the other side, water evaporation causes over rain under high temperature in tropical and subtropical countries and this leads to over water stress in plants as a result of overflow of water balance (Ashraf., 2012). In addition, over water stress is both a stress factor as the result of regular rain and overflow or over-irrigation, and hence this effect can be more fatal than water shortage for plants. Thus, diffusion gets weak at the result of low oxygen in the soil during the overflow and toxic substances are formed in roots under anaerobic conditions. Leaf withering and/or root rot, in a word, underdeveloped plants appear in this case.

So, both forms of water stress make/cause obstacles in the growing of plants as well as corn plants, especially in the level of sprout (Song et al., 2019). Plants activate the defense mechanism in order to be protected from the influence of stress. The main based component of this is to make the enzymes NADPH pool which forms reducing potential of cells. NADPH is considered as one of the widespread components in nature that possesses high energy, forms basis of reducing potential of cells, plays a great role in biosynthesis process, and has a central place in metabolism (Corpas et al., 2014). NADP-dependent isocitrate dehydrogenase (NADP-ICDH, EC 1.1.1.42) enzyme that makes NADPH in living organism whose activity has been determined in corn sprouts, playing an important role in the defense of plants from stress. This enzyme splits isocitrate into α-ketoglutarate and CO_2 gas. The reaction is observed by forming NADPH. There are both mitochondrial and cytoplasmic forms of the enzyme in eukaryotic cells. The enzyme acts on non-photosynthetic cells, or on the darkest stage of metabolism of photosynthetic cells (Leterrier et al., 2012).

It should be noted that ferredoxin -NADP-reductase is considered the main enzyme making NADPH, forming its pool on the light stage of photosynthetic cells. NADP metabolites which realize reducing coenzyme function and enzymes which

make out its synthesis in plants have a great role in their defense and adaptation to extreme conditions (Ying et al., 2008). Taking all into consideration, activity dynamics of NADP-ICDH enzyme are related to the development of corn sprouts in extreme conditions which have been created by water stress and are aimed to be studied in this research work/ and fall into the core of this study.

Materials and methods

Object of experiments and germination of seeds

Experiments have been done on the *Pride (Gurur)* genotype of corn (*Zea mays* L.) that was brought from the Institute of Genetic Resources of the Azerbaijan National Academy of Sciences. Corn seeds have been disinfected in a 3% solution of hydrogen peroxide for five minutes in order to get rid of pathogens. Then they were soaked in Petri Dish by washing in distilled water during a day and the next day they were replaced in a plant growing device with phytotron principle (MLR Plant Growth Chamber) by planting in vegetation containers under 24-27 C^0 temperature. After the observation of sprouts on the soil, the controlled variants were/variant was irrigated by distilled water once on the third day of the experiment. The second variant, created for the aim of over water, was irrigated twice a day,but the third variant firstly was irrigated once every two days in order to create drought, after six days it was irrigated once every three days, but the irrigation process was stopped through the next days. The biometric indicator of the plant and the calculation of the activity of NADP-dependet ICDH enzyme were registered once every four days.

Making of enzyme preparation

100 mM Tris-HCl buffer (pH 9) that contains 5mM $MgCl_2$, 2 mM EDTA, 14 mM merchaptoethanol, 5% polyvinylpirolidone (PVP), 1% polyethylene glycol (PEG) is used as extract solution. Tissue: extract solution was taken in a ratio of 1 gr:4 ml. The obtained homogenate was firstly centrifuged 1000 gr, then 9000 gr, after it was centrifuged fast. The supernatant part of the solution was taken and it was used as enzyme preparation.

The indication of the activity of NADP-ICDH enzyme

The activity was calculated at 340 nm wavelength according to NADP reduction speed under 24 C^0 temperature on the spectrophotometer of MRS (Israel) by the spectrophotometric method. 50 mM Tris-HCl (pH 8,2) which contains 2,5 mM $MgCl_2$, 2 mM-D, L- isocitrate and 0,5 mM NADP and its solution were used as a reaction environment. Enzyme preparation was added to this solution and the

reaction started by adding NADP to the reaction environment and the measurement repeated four or five times.

Results and discussion

It is known that there are serious abiotic stress factors that reduce the growing and productivity of plants, so water stress has got an important place among such kinds of cases. Although the drought, as the result of water shortage, is widespread, water stress manifesting as over water problem causes the incomplete growing and destruction of plants too. If we take into account that ecological factors have laxative effects on all the creatures/ living creatures in nature, including plants, year-by-year, it is necessary to eliminate abiotic stress for future and water stress too. Meanwhile, studying how to make defense mechanism in order to create stress resistant sorts is one of the important issues. If we consider that environmental change is regular, then this task stays actual. Bringing this fact into considering this fact, we have tried to study the influence of the defense system of corn plants in the condition of water stress which is one of the factors appearing with the influence of environmental changes.

Biometric indication of the growing of the root and trunk sprouts of the corn, grown under the artificial water condition, has been introduced in the following table as a control variant.

Table 1. Growing indications of corn sprouts under the condition of control and water stress (by sm).

Variants	4 days'	8 days'	12 days'
Control	leaf 6.2±03 root 3.1±02	leaf 9.4±04 root 4.5±03	leaf 13.6±04 root 5.9±03
Water shortage stress	leaf 4.6±01 root 2.5±01	leaf 6.1±01 root 3.2±01	leaf 8.1±02 root 4.1±01
Over water stress	leaf 3.9±01 root 2.1±01	leaf 5.0±01 root 2.9±01	leaf 6.3±01 root 3.3±01

As seen in the table, the leaves of the sprouts under the control variant have been increased 2,1 times compared to the fourth day for 12 days, but the growing of the root system has been increased 1,9 times. Both the growing of the leaves and the root system have been slow down significantly under the influence of the both forms of water stress. So, similar figures for the plant's leaves constitute 1,7 times, but this figure is 1,6 times for the root system. It means that informal water supply of the plant in the both forms of water stress has disrupted the functions in the tissues of

the leaf and root system by effecting the metobolism of the plant, and as a result, it caused to reduce the biometric indications of them.

The next table describes the dynamics of the activity of NADP-dependent isocytrate dehydrogenase enzyme on the homogenate, prepared from the root and leaf tissues of corn plants in the condition of control and water stress.

Variants	First	4 days'	8 days'	12 days'
Control leaf				
root	86.1 ± 3.2	97.6 ± 3.1	109.7 ±3.5	122.5 ± 4.1
	78.5 ± 2.3	89.0 ± 3.8	99.7 ± 3.4	110.8 ± 3.8
Water				
shortage leaf	-	103.5 ± 4.1	124.3 ± 3.8	139.5 ± 4.0
root	-	99,6 ± 4.3	111.3 ± 4.2	124.1 ± 4.1
Over water				
leaf root	-	114.7 ± 2.9	137.6 ± 3.3	147.8 ± 4.1
	-	101,2 ± 4.1	117,1 ± 3.6	128.2 ± 3.7

Table 2. The description of the activity dynamics of NADP_ICDH enzyme in the tissues of the leaf and root systems in the condition of control and water stress related to the growth of corn sprouts.

As it can be seen/As shown in the table, the activity of NADP-ICDH enzyme, related to the corn sprouts, is increasing significantly both in the cells of the leaf and root systems. Both forms of water stress cause the next induction of the activity of the enzyme. We can say that/We can conclude that according to the table, over water stress activates the enzyme more strongly than water shortage stress in both tissue cells. The activity degree of the enzyme in the leaf tissues shows itself more than root tissues in three variants. It seems that NADPH is required for corn sprouts to eliminate the complications of stress conditions and NADP-ICDH enzyme participates in its synthesis directly.

Conclusion

The activity dynamics of NADP-dependent isocitrate-dehydrogenase enzyme in the root and leaf tissues during the growth period of corn plants in the condition of control, over water and water shortage and the biometric indicators of the plant have been observed. It is determined that the activity dynamics of NADP-dependent isocitrate- dehydrogenase enzyme is followed by increasing sequence with the influence of water stress both in the root and leaf cells. The activity of enzyme in the leaf tissues is higher than root tissues. The influence of stress is negative on the

growth of the biometric indicators of corn plants, as well. The growing of the plant's leaves is noticeably slower than the roots in both forms of water stress compared to the control condition.

References

Arora, A., Sairam, R.K. & Srivastava, G.K. (2002). Oxidative stress and antioxidative sistems in plants, Division of Plant Phisiology, 82, 1227-1238.

Ashraf, M.A. (2012). Waterlogging stress in plants, African Journal of Agricultural Research, 7 (13), 1976-1981.

Boretti, A. & Rosa, L. (2019). Reassessing the projections of the World Water Development Report, Clean Water 2, 15.

Corpas, F.J. & Barroso, J. B. (2014). NADPH-generating dehydrogenases: their role in the mechanism of protection against nitro-oxidative stress induced by adverse environmental conditions. Environmental Science, v 2, 1-5.

Galvez, S. & Gadal, P. (1995). On the function of the NADP-dependent isocitrate dehydrogenase isoenzymes in living organisms. Plant Science, 105, 1-14.

Gill, S.S., Anjum, N.A., Hasannuzman, M., Gill, R., Trivedi, D.K., Ahmad, I., Pereira, E. & Tuteja, N. (2013). Glutathione and glutathione reductase: a boon in disguise for plant abiotic stress defense operations. Plant Physiology Biochemistry, 70, 204–212.

Jain, M., Kataria, S., Hirve, M. & Prajapati R. (2019). Water Deficit Stress Effects and Responses in Maize, Plant Abiotic Stress Tolerance, 129-151.

Kumar, S. (2020). Abiotic Stresses and Their Effects on Plant Growth, Yield and Nutritional Quality of Agricultural Produce, İnternational Journal of Food Science and Agriculture, 4(4), p.367-378.

Leterrier, M., Barroso, J.B., Valderrama, R., Palma, J.M. & Corpas, F. J. (2012). NADP-dependent isocitrate dehydrogenase (NADP-ICDH) from Arabidopsis roots contributes in the mechanism of defense against the nitro oxidative stress induced by salinity Scientific World Journal, 12, 694-740.

Munemasa, Sh., Muroyama, D., Nakamura, Y., Mori, I. & Murata, Y. (2013). Regulation of reactive oxygen species-mediated abscisic acid signaling in guard cells and drought tolerance by glutathione, frontiers in Plant Science, vol. 4, 472.

Sagi, M. & Fluhr, R. (2006). Production of reactive oxygen species by plant NADPH oxidases. Plant Physiology, 141, 336–340.

Song, L., Jin, J. & He, J. (2019). Effects of Severe Water Stress on Maize Growth Processes in the Field, Sustainability ,11, (18).

Ying, W. (2008). NAD+ /NADH and NADP+/NADPH in cellular functions and cell death: regulation and biological consequences. Antioxidative. Redox Signal. 10, 179–206.

Khazar Journal of Science and Technology Volume 5 №2 2021, 83-89
 DOI: 10.5782/2520-6133.2021.5.2.83

Applicatıon of the Hydroponic Green Fodder Technology in Poultry Breedıng and Maintenance of The Broiler in as Provided by Zoogygienic Conditions

Siala Rusramova, Mirzammad Hasanov
Veterinary Scientific Research Institute
Corresponding author siala.rustamova@gmail.com

Abstract

The article sets out as a goal the ensuring in a sustainable form of innovative development in the designing of feeding strategies within the framework of the project "Sustainable development of poultry farming and the creation of a value chain for food production" in farms of the republic. Also, for rational and proper feeding of birds, the use of the Hydroponic Green Fodder technology, the presence of high-calorie protein (raw protein) in the feed, as well as amino acids, a small amount of cellulose, carotene, calcium and phosphorus, create favorable conditions for the development of chickens.

Keywords: Hydroponics, rational feeding, technology, raw protein,infection.

Introduction

The development of breeding and feeding strategies under the “Sustainable Development of Poultry Farming and Creation of Value Added Chain for Production of Foodstuffs” Project, funded by the Government of Azerbaijan and have been implementing since December 2018 under the FAO-Azerbaijan Partnership Program to ensure production in line with market demand, was set as a goal to develop in a sustainable manner in/through the functioning of added value chain for poultry (Mitrofanov, 2010).

Since poultry farming is one of the fast-growing, widespread and profitable sectors of animal husbandry, the government's concern for the efficient use of this industry has increased significantly. Thus, in the field of poultry farming, new farms and individual farms have been created (Zayas, 2013). Relevant work has been carried

out in the direction of increasing poultry production using hydroponic and other innovative technologies. The increase in financial support for the development of the poultry farming from year to year, the commissioning of production sectors involved in the production and packaging of poultry meat and the marketing organization have played an irreplaceable role in the development of this sector.

The development of poultry farming, as well as an increase in the output of poultry farming products, largely depends on the fact that the chickens selected for raising of broilers are healthy, lively, well developed and of the same weight. That is, if there is a weight of 55-60 grams of eggs laid in the incubator, the weight of one-day-old chicks is 35-40 grams. Healthy chickens hatch on time and in droves, grow quickly, and mortality decreases. Healthy chicks are selected in 6-8 hours after hatching, and the retarded chicks from growth are culled (Mikailov, 2014).

For the proper feeding of birds, the Hydroponic Green Fodder technology is one of the key factors in ensuring the intensive growth of chicks. Because of the presence in the feed allowance consisting of high-calorific protein (raw protein), as well as amino acids, a small amount of cellulose, carotene, calcium and phosphorus, creates favorable conditions for the development of chickens (Mikailov et al., 2017).

Innovation of Hydroponic Green Fodder technology

The introduction of innovations in farmers' poultry units plays an important role in the modernization and intensification of production (Shafiyeva, 2005). Therefore, there is a need to apply an innovation mechanism for improvement the feed supply in farmers' poultry units. From this perspective, the innovation of Hydroponic Green Food technology for improvement the feed supply in poultry farms can meet the needs of not only small farm enterprises, engaged in poultry farming, but also of large complexes. From this point of view, there are no analogues to the Hydroponic Green Fodder technology. Using this method, in agriculture, regardless of weather conditions, it is possible to produce green fodder of steady quality, continuing during 365 days a year. Grains sprout at an above-zero temperature of 18° C, and micro-macro vitamins are transferred to the plant along with water. So long as, green fodder is rich in high-calorific protein (raw protein), as well as amino acids, carotene, calcium, phosphorus and vitamin E, which leads to increased productivity indicator in poultry. The feed obtained by this method improves digestion in birds and stimulates weight gain.

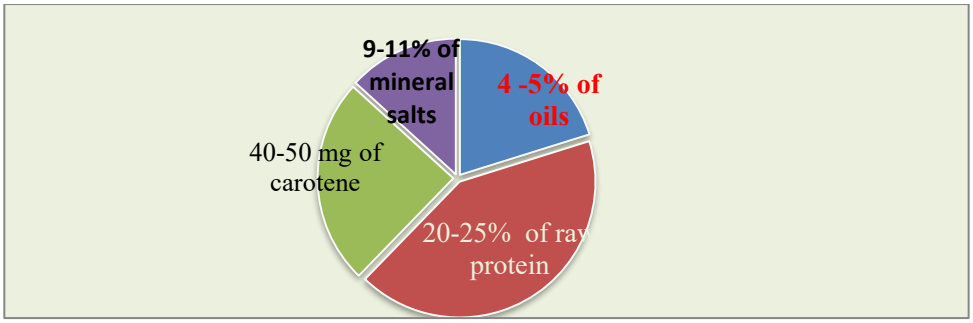
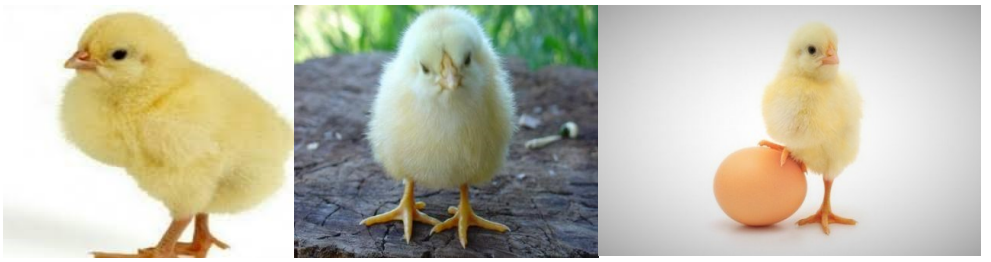


Figure 1. The amount of substances in the composition of 1 kg of green fodder

Birds eat up this food with great appetite. It should be considered that 1 kg of green fodder contains 20-25% of raw protein, 4-5% of fat, 35-50% of nitrogen-free extraactive substances, 9-11% of mineral salts and 10-15% of cellulose. Plus, 1 kg of green fodder contains 40-50 mg of carotene, which directly affects the productivity of birds.

As a result of the application of the Hydroponic Green Fodder technology in family poultry farms, in 2019, the total number of eggs obtained from each chicken increased by 11.6%, the weight gain (200 g) in each chicken (in broiler) intended for slaughter - increased by 13.3%, the number of chicks leaving the incubator increased by 12.2%. As a result of the use of Hydroponic Green Fodder in the feed allowance, the efficiency of the farm has increased and the costs of feed for each bird have decreased. The complete organization of the sprouting of bird feed by the hydroponic method takes 5-6 days. Thus, from 1 kg of wheat within 5-6 days, 4-5 kg of green mass is obtained. It should be noted that wheat is first washed in special containers, this clears the grains from poisonous spores of fungi and dust, and then the grains for sprouting are laid out in each container at a rate of 1–1.2 kg per one container. Starting from the second day, seedlings are obtained. On the third day, the root system of the grain begins to branch out and develop vegetatively. On/Through the fourth to fifth days, along with the rapid development of both roots and green mass, the amount of raw protein (protein), starch (sugar), mineral salts, etc. is increased in the composition of the feed. Therefore, from an environmental point of view, with

regard to the full practical value of pure fodder, it is used to meet the needs of poultry for raw protein, water carbon, calcium, phosphorus and carotene. Birds eat food along with the roots, and the main nutritional value is in the root system. Continuous feeding with hydroponic green fodder ensures the health of broilers, daily weight gain, and the development of quality poultry meat production.



In order to ensure the innovative development of broilers, the proper formulation of the feed allowance of chickens creates the opportunity for rapid growth of the broiler and hence the presence of a high percentage of healthy rearing. Demand for feed depends on the age and body weight of the broiler.

For the first five days, chicks are fed on boiled eggs, wheat bran, bottom milk, fresh anthroponic green fodder, crushed shells, and so on. Eggs should be boiled in clean water, then crushed together with the shells and given as feed. It is also recommended to add additional 2-3% of feed yeast to feed. In this case, the yeast increases the amount of vitamins in the feed by increasing/raising the number of cells and bacteria. The normal metabolism in young birds is largely depended on the good care of chicks. Thus, the soaked feed mixture is fed to 30-day-old chickens 3-4 times a day (one ration of hydroponic green fodder), and then 2-3 times. Chicks up to ten days of age are fed every two hours, and then up to 30 days of age every 3 hours. For each head of broiler in the first week, 12-15 g are given, 21-20 g in the second week, 40-45 g in the third week, 60-65 g in the fourth week, 75-85 g in the fifth week, 90-100 grams in the sixth to seventh weeks and after the eighth week 100-110 grams of feed are given. The live weight of a broiler fed in this way is estimated 1500-1600 g and more for seven to eight weeks with the consumption of 2-2.5 kg of a feed unit per 1 kg of live weight.

By a long-term feeding with compound feed, for each ton of it, 10-15 grams of antibiotics should be added. The antibiotics given at this dose have a stimulating effect on the rapid growth of broilers and prevent many intestinal diseases. On the other hand, long-term use of antibiotics with high doses retards the growth of broilers, deteriorates the composition of some vitamins and gives poor results. All antibiotics should be removed from the feed ration 8-10 days before slaughter. After

49-63 days, broilers are transferred to the slaughtering and processing shop. For 6-8 hours prior to the slaughter, broilers are kept without feeding.

Breeding of broilers on bedding

During the rearing on the bedding, 4-5 batches of chickens intended for meat can be raised in one premise per year. After breeding of each batch, the premise is cleaned, disinfected, the bedding is laid anew and the next batch is accepted. The following technology is applied on a thick bedding: 0.7-1 kg of slaked lime is laid on each plot of a square meter, corn stalks in 5-7 cm of thickness, peat, tree bark or chopped straw are laid on it. Until the end of the breeding period, 1.5-2.0 kg of bedding is used per bird. During the breeding period, 10-12 chickens are kept on each square meter in winter, and 9-10 chickens in summer.

Effect of aeration on breeding of broilers

In broiler development, air exchange is of great importance. With insufficient oxygen in the premise, the chickens intended for meat eat less than usual, respiratory diseases develop among them, and the expected weight gain is not achieved. Therefore, ventilation must be regulated in such a way that for each kg. of live weight there was produced 1.5 cubic meters of air exchange. In addition, when the premise, where the broilers are bred, is lit for 18 hours in the first week during the day and

then gradually reduced to 14 hours. Good results can be obtained in increasing the weight of meat chickens. Since chicks are unable to regulate their body temperatures from day one, temperature fluctuations inside the premise will have a sharp effect on them. The presence of a sharp air flow inside the premise causes such abnormalities in birds as diarrhea, delayed feathering, and deterioration in digestion. Cold and wet weather causes kidney damage, wet and cold causes severe diarrhea, and dust, ammonia and carbon dioxide leads to the occurrence of various diseases of the respiratory organs. At normal heat levels, chicks are in brisk condition, their feathers are smooth, they readily eat feed, drink enough water and this ensures the rapid weight gain. Control over the quality of water and feed used in poultry farms is of great importance. Water pollution with organic substances also leads to dangerous consequences. Poor water quality leads to corrosion, tanning and becoming moldy of some means in poultry farming as well. Water hardness depends on the amount of calcium containing in it. The presence of more than 4 mg of calcium in one liter of water indicates its hardness. In order to provide birds with good quality of water, oxidizing agents that are not harmful to birds should be added to the water composition. For this purpose/to do this, citric acid and chlorine are most commonly used. In this case, the total number of microorganisms in the water decreases, the pollution with organic substances comes down, and the accumulation of some metals reduces. Drinking containers should always contain clean water. In order to protect chickens from gastrointestinal diseases, potassium hypermanganate is added to the water in a ratio of 1: 10,000.

Result and discussion

1. Development of breeding and feeding strategies within the framework of the "Sustainable development of poultry farming and creation of value added chain for production of foodstuffs".

2. Observing appropriate principles in the direction of increasing the production of poultry meat in farm enterprises and individual farms using/taking advantage of hydroponics and other innovative technologies.

3. To organize intensive growth of poultry, it is proposed using the hydroponic green fodder technology for full-quality and rational feeding of broiler and keeping in accordance with zoohygienic conditions.

4. When rearing a broiler, air exchange and breeding on bedding the regulation in accordance with zoohygigenic standards should be controlled.

5. Entry routes to poultry-houses, hatchery houses, slaughtering shop and other premises should be equipped with disinfectant facilities, built in 1.5 m long and 15 cm deep, which should be regularly moistened with disinfectant solutions.

6. For disinfection of automobile tires, new disinfection barriers/means should be organized and a 3% formaldehyde solution or chlorinated lime solution with 1% active chlorine in its composition, 5% creolin or xylonaphtha solution should be added there/should be added at the entrance.

References

Mitrofanov, N.S. (2010). Technology of products from poultry meat, Moscow.

Mikailov, M.S. (2014). "Bacterial contamination of hatchery facilities and air of premises of poultry farming" Journal of Veterinary Medicine, No.6, p.15.

Mikailov, M.S., Abbasov, S.B., Hasanaliyev, N.Kh. & Hasanov, M.S. (2017). "Infection with conditional pathogenic microbes of poultry and decorative birds." Scientific works of the Institute of Microbiology of the National Academy of Sciences of Azerbaijan 2017, volume 15. No. 1, p. 122.

Shafiyeva, N.V. (2005). Avian influenza. Background information. Baku, 2005, p.11.

Zayas, Y.F. (2013). Quality of meat and meat products, Moscow.

Information to Contributors

Khazar Journal of Science and Technology focuses on the results of original research projects in Mathematics, Physics, Chemistry, Earth science and life science as well as Engineering and Technology.

The journal is published in English.

The Editorial Board does not accept articles published or submitted for publication elsewhere.

Information material, such as notifications about conferences, books, jubilees, etc. may also be published under the information heading. Articles and other materials published in the Khazar Journal of Science and Technology represent the opinions of the author(s) and should not be considered to reflect the opinion of the Editorial Board.

Specifications

Manuscripts should not exceed 25 published pages (approximately 10,000 words). Those of a length exceeding 25 pages, nut containing very important results, can be published by the consent of the Editorial Board. Articles may be submitted by e-mail. Please submit your manuscript to the following addresses:

Submission-kjsat@khazar.org and submission.kjsat@gmail.com

The first page should include the title of the article, the names and primary affiliations of the author(s), an abstract of 50-250 words, and 4-7 keywords. All figures, photographs, tables, or drawings should be numbered.

References must be in one of the following styles (depending on the subject area): AMS, LaTeX, ACS, AIP, Vancouver system, etc. entries in the reference list should be put in alphabetical order or in order of citation. Footnotes in the text should be avoided if possible.

The author(s) should provide subject classification codes (AMS subject classification numbers, the Physics and Astronomy Classification Scheme, ACM computing Classification System or Library of Congress Classification).

The Khazar Journal of Science and Technology follows the COPE's guidelines in dealing with allegations of research misconduct. Khazar Journal of Science and Technology follows the conflict of interest policy of the PNAS (Proceedings of the National Academy of sciences, USA).