SEASONALITY OF GENOMIC VARIATION IN *Drosophila melanogaster*

Drosophila melanogaster'de GENOMİK VARYASYONUN MEVSİMSELLİĞİ

ŞENEL SELİN ŞENKAL

ASST. PROF. DR. BANU ŞEBNEM ÖNDER
Supervisor

Hacettepe Üniversitesi
Submitted to Graduate School of Science and Engineering of Hacettepe University as a Partial Fulfillment to the Requirements for the Award of the Degree of Master of Sciences in Biology

2020

Babama…

ABSTRACT

SEASONALITY OF GENOMIC VARIATION IN *Drosophila melanogaster*

Şenel Selin ŞENKAL

Master of Science, Department of Biology
Supervisor: Asst. Prof. Dr. Banu Şebnem ÖNDER
Co- Supervisor: Martin KAPUN
July 2020, 64 pages

Local adaptation is of fundamental importance in evolutionary biology and understanding the genetic basis of adaptation to new environments has gained importance in recent years. One of the most common organisms for these studies is *Drosophila melanogaster*. As is well known, *Drosophila melanogaster* is a cosmopolitan species and is spread almost all around the world. Although its whole genome has been known for many years, this organism has started to be used more in adaptation studies with the development of new generation sequencing technology. Adaptation is the primary mechanism that allows organisms to survive in different environments. Spatial and temporal environmental variation can lead to different selective pressures on populations. The direction of selection by spatial and temporal alteration, and the mechanism behind rapid adaptation are poorly understood. Seasons are one of the important temporal effects on populations in temperate regions, and *Drosophila* may respond through rapid adaptation to temporal changes in the environment. For this purpose, we analyzed genomic variation of D. melanogaster to determine seasonal single nucleotide polymorphisms (SNPs) by using Pool-Seq next generation sequencing method to understand mechanism

underlying rapid adaptation to seasonal changes in this organism. Our results suggest that seasons cause genomic variation in *Drosophila melanogaster*. Tajima's D values were mostly negative for 2014 samples, but we did not see this pattern for other years. Majority of F_{ST} values, the differentiation between the samples from different timepoints of the year, were not high, but at some regions it was as high as 0.45, yet this was not consistent through years. We also calculated allele frequencies and we found 982,000 common SNPs in three year samples which have sharing common positions. Almost half of these SNPs were intronic, 9.4% were exonic, and 8.7% were in the intergenic regions. We found a total of 6516 structural variants such as insertions and deletions. Most of these SNPs were not seasonal however, approximately 3.5% (32,428) of them were seasonally significant. Approximately 72% of these SNPs were in protein coding regions. We also identify genes that contain seasonal SNPs such as *couch potato (cpo), sickie*, and *Insulin-like receptor (InR)* which are involved in crucial signaling pathways in *Drosophila melanogaster*. These results suggest that seasons in temperate regions create a selection pressure on *Drosophila melanogaster* populations and local populations respond to this pressure with rapid adaptation.

Keywords: Seasonality, *Drosophila melanogaster*, Next-generation sequencing, Pool-sequencing.

ÖZET

Drosophila melanogaster'de GENOMİK VARYASYONUN MEVSİMSELLİĞİ

Şenel Selin ŞENKAL

Yüksek Lisans, Biyoloji Bölümü
Tez Danışmanı: Dr. Öğr. Üyesi Banu Şebnem ÖNDER
Eş Danışman: Martin KAPUN
Temmuz 2020, 64 sayfa

Lokal adaptasyon evrimsel biyolojideki temel konulardan biridir, ve yeni çevrelere adaptasyonun genetik mekanizmasını anlamak, son yıllarda artan bir önem kazanmıştır. Bu tip çalışmalar için yaygın olarak kullanılan organizmalardan biri olan *Drosophila melanogaster*, kozmopolit bir tür olup, dünyanın birçok yerine yayılım göstermiştir. Bütün genomu uzun yıllardır bilinmekle birlikte, yeni nesil sekanslama çalışmalarının artmasıyla bu organizma adaptasyon çalışmalarında daha çok kullanılmaya başlanmıştır. Adaptasyon en temel mekanizma olup, organizmaların farklı çevresel koşullarda hayatta kalmasına yardımcı olur. Mekansal veya zamansal çevresel varyasyonlar populasyonlar üzerinde seçilimsel baskıya sebep olur, ancak zamansal ve mekânsal olarak değişikliklerin neden olduğu seçilimin yönü ve bu değişikliklere canlıların hızlı adaptasyonunun mekanizması yeterince anlaşılamamıştır. Mevsimler ılıman kuşakta yaşayan populasyonlar üzerindeki en önemli zamansal etkilerden biridir ve zamansal değişikliklere *Drosophila* hızlı adaptasyon yoluyla cevaplar üretebiliyor olabilir. Bu hızlı adaptasyonun altında yatan sebepleri anlayabilmek ve genomdaki mevsimsel değişkenlerden etkilenen bölgeleri bulabilmek amacıyla Pool-seq metodu ve yeni nesil sekanslama tekniği kullanarak, *Drosophila melanogaster'*in tüm genomunu sekansladık. Elimizdeki sonuçlara göre, bulmuş olduğumuz tek nükleotit polimorfizmlerinin (SNP) yaklaşık

%50'si intron bölgelerinde olup, %9.4'ü ekzonlarda, ve %8.7'si ise intergenik bölgelerdedir. Tajima'nın D'si 2014 yılında tespit edilen bölgelerin neredeyse tümünde negatif çıkarken diğer yıllar için bu durum gözlenmedi. F_{ST}, populasyonun farklı zamanlarda toplanan örnekleri arası farkı anlayabilmek amacıyla hesaplandı, ve sonuçlara göre örnekler arası fark çoğunlukla düşüktü. Bazı bölgelerde farklılaşma 0.45'e kadar çıkmasına rağmen bu yüksek farklılaşmayı aynı bölgeler için diğer yıllarda göremedik. Örneklemler arası alel frekanslarını da hesapladık, ve üç yılın örneklemleri için pozisyonlar bakımından ortak olan 982.000 SNP bulduk. Bu bölgelerin çoğunluğu mevsimsel olmamakla birlikte, yaklaşık %3,5'i (32.428 SNP) mevsimsel olarak anlamlı çıktı. Önemli yolaklarda bulunan *couch potato (cpo), sickie*, ve *Insulin-like receptor (InR),* gibi genlerde mevsimsel SNP'ler tespit edildi. Sonuçlarımız, mevsimlerin ılıman bölgelerde yaşayan *Drosophila melanogaster* populasyonlarında seçilim baskısı oluşturduğunu ve populasyonun bu baskıya hızlı adaptasyon ile yanıt verdiğini göstermektedir.

Anahtar Kelimeler: Mevsimsellik, *Drosophila melanogaster*, Yeni nesil sekanslama, Pool-sekanslama

ACKNOWLEDGMENTS

First, I would like to thank to my advisor Banu Şebnem Önder for her support, patience, great help and encouragement, and to my co-advisor Martin Kapun, for his great help and patience during this process. I started without knowing anything and I came from a different background, however they supply me every information and support during my graduate studies.

I would like to thank TUBITAK for financial support for this thesis under the project number 214Z238 and 116Z238. I also extent my thanks to TUBITAK ULAKBIM, High Performance and Grid Computing Center (TRUBA resources) where the analyses partially performed.

I would like to thank my friends for their support and accepting me with all my weird stuffs. Also I would love to thank to Ekin Demir for her mental support and great friendship during the thesis project and my graduate life, Cansu Aksoy for sharing the same spirit and interest which is a joy, to Ezgi Çobanoğlu for the fun moments, to Seda Coşkun for encouraging me.

I would like to thank to my husband for his understanding, advices, and patience, to my mother for our little talks which is a great thing because she always supported me, made me go further and better and to my brother for his help during these times, and to my-mother in law for helping me with many ways and easing this process for me with many things,

I really want to thank to my father for always believing me, supporting me, giving advices, and helping me especially when I cried or feel bad. I really wish you could be here and see this moment. I am quite sure, you would be proud and encourage me no matter what, even though what I did was not that great. This was who you were. You always supported me for all my undergraduate studies, even it took 7 years to complete. Thank you for making me who I am today. I will always love you.

Thank you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SYMBOLS AND ABBREVATIONS

SNP	Single Nucleotide Polymorphism
NGS	Next-Generation Sequencing
POOL-SEQ	Pool Sequencing
UV	Ultraviolet
DNA	Deoxyribonucleic Acid
IN	Inversion polymorphims
F_{ST}	Fixation Index
MIF	Minor Allele frequency

1. INTRODUCTION

Understanding the genetic basis of adaptation has become an important research area in evolutionary biology in recent years. Advances in sequencing technology enabled us to compare genomes of different populations or even species to understand the mechanisms underlying adaptation and evolution. Genomic variation that has advantages for the organism could be selectively favored, where disadvantaged one could be negatively selected [1]. For example, bacteria that carry penicillin resistant plasmids have transmitted this plasmid from generation to generation because of fitness advantages. When the population is exposed to penicillin, the ones who carry resistant plasmid will survive and the rest of the population will mostly be eliminated [2]. Another classic example comes from the Galapagos finches', where beaks are varied due to their food preferences and environment that they live in. Gray treefrog (Hyla *versicolor*) and Green treefrog (*Hyla cinerea*) are two different species which live in very close habitats however, their colors are different which is an adaptation to their local environment to protect from their predators to survive [3]. *Uta stansburiana* which is known as side blotched lizard populations that live in different habitats have different colors that match the underground and thus enhances their survival probability [4]. These examples amongst many others, show that these traits are results of natural selection and selected for adapting to their local environment, which is changed in spatial and temporal scale, for survival and reproduction.

There are many factors that affect and shape the life of an organism, but seasons are one of main temporal factors in temperate regions, which caused by Earth's tilted rotation. There are many organisms that are cosmopolitan species and could live almost anywhere on the Earth, however temporal changes/selection affects populations in their local environment. One of these cosmopolitan species is, the fruit fly *Drosophila melanogaster*. *D. melanogaster* is one of the best-studied organisms in biology and has been a genetic model since the early 1900s. *D. melanogaster* are easily cultured and have a short generation time and large number of offspring. The *D. melanogaster* complete genome sequence was published in 2000 [5]. *D. melanogaster* is a cosmopolitan species and spread all around the world except poles and deserts. *D. melanogaster* has a high adaptation capacity which allows it to rapidly adapt to environmental changes therefore live in a wide range of habitats. It originated from sub-Saharan Africa and expanded its region to Europe and Asia as a human commensal

approximately 15.000 years ago, then to America and Australia in 1800s [6–10]. Even though this species is native to tropical regions, *D. melanogaster* adapted to temperate regions due to its association with humans [6]. The high adaptation potential helped it survive in different environmental conditions. The great success behind the adaptation to wider range of environments of cosmopolitan species is poorly understood. Therefore, *D. melanogaster* is a powerful model for comparative genomics studies that are aiming to understand the mechanism of local adaptation that is largely unknown.

2. GENERAL INFORMATION:

2.1. The Genetic Basis of Adaptation:

Adaptation is a mechanism that allows organisms to survive in different environments. Organisms may respond to a new environment starting with phenotypic plasticity and could be followed by genetic adaptation [11–13], or could not be followed in some traits [14]. Long living organisms or the ones that only give few offspring in a year, can show phenotypic plasticity which is an adaptation method that organisms gain some tolerance to the environment [15,16]. However, it is possible to observe the effects of environmental changes at the genomic level for organisms that give many offspring in a year [17] and also in long living organisms[1]. Therefore, comparative studies across time (within population) and space (between populations) are useful to understand the evolutionary mechanisms behind adaptation and demographic effects. Previous studies showed that spatial and temporal changes had an effect on organisms [1,18–24]. Geographical factors such as latitude, longitude or altitude, and temporal differences like seasons create different selection pressures on organism and populations may give different responses to environmental changes that could either be seen as change in color, growth and reproduction, body size or life span [4,18,23] or in the genome level with the change in allele frequencies in the genome [19,22,24–27]. These conditions not only affect the survival of the individuals, but also affect the habitat selection, dispersal, or reproduction [20].

Local adaptation is a fundamental process in many studies related to evolutionary biology, population genetics and climate change. Some organisms may survive in different environments but still many of these prefer their local environment. For example, a study with *Carlina vulgaris* indicated that populations grown in their local environment had higher fitness in survival, growth, and reproduction compared to populations transplanted to different regions. They also claimed that these results mostly related with climate in their local environment; due to geographical distance climatic factors have been changed [18]. In a similar study with *Arabidopsis thaliana*, the reciprocal transplantation of two European populations (Sweden and Italy) showed that local populations perform significantly better than the nonlocal populations in 8 out of 10 criteria related to reproduction and survival [28]. A comparative study in Pisgah Lava field in California, USA between dark colored lava-flow population and Off-lava population of side-blotched lizards, *Uta stansburiana*, showed that

both populations have a plasticity to change color when they moved to a new environment. But they have heritable differences in two genes which are involved in melanin regulation. Most of the differentiated sites in these two genes were only found in Pisgah populations, so, they concluded that adaptation which is possibly a result of *de novo* mutations in this population resulted after the formation of lava flow, and could work together with phenotypic plasticity to promote survival in this new environments for Pisgah populations [4].

Many studies suggested that *Drosophila* populations differed at the genomic level due to spatially varying selection [29–32]. One of these studies compared European and African populations of *D. melanogaster* in a 20kb genomic region on the X-chromosome and found eight regions that were specific to the populations of temperate regions which caused amino acid changes. Seven of these were in *Flotillin-2* gene; an insertion and six of the seven were in the intronic region and the other was in the exon region of CG9503 gene. [29]. Begun and Aquadro (1993) showed that African populations were more variable than American populations at the genomic level, that most of the variants were not shared and that there were significant differences at low recombination rated genomic regions between these populations. Another study [31] found that two populations (Zimbabwe and Netherlands) of *D. melanogaster* differed from each other, and also found traces of positive selection in European populations and they claimed that this is consistent with previous hypotheses that European populations were under positive selection. In another study, the expression levels of CG9509, a cis-regulatory region were compared, and the results indicated that this region was expressed less in African populations than non-African populations [32]. These studies suggested that adaptation is a complex mechanism and populations can differ spatially by the pressure which is created by environmental conditions.

2.2. Adaptation to spatial variation:

The environment that we live in, shapes the physiology, morphology, or the life history of an organism. Spatial factors such as altitude, longitude and latitude create different pressures on organisms. For example, skin color of humans changed due to UV radiation, and became darker in the equator region and lighter at higher latitudes. It is thought that this diversity was caused by UV protection in the equator region and UVB requirement for vitamin D synthesis in non-tropical environments [33]. Also, we can observe these coloration processes for many animals such as polar bears, or weasels or foxes that live in the arctic region which is an adaptation method to local environment for camouflage [34]. Other than coloration, there are many traits

affected by spatial changes. Altitudinal populations of copper butterflies (*L. tityrus*) showed variation in several traits and PGI allele frequencies [19], an important enzyme in glycolytic pathway [35], mostly related with cold resistance, whereas heat resistance mostly related with heat shock protein gene (*hsp*) and populations that were collected from higher altitudes were less tolerant to heat stress [36]. In a study, *D.melanogaster* populations from five different continents were compared at the genomic level, and an average of 2,928 structural variants such as insertions, deletions, inversions, and translocations were found among these populations. Additionally, these variations varied in size and location so they claimed that these may have a role in adaptation and genome evolution [26]. Genomic variation of *D. melanogaster* populations varied widely with latitudes [21,37,38], for example high latitudes populations were less polymorphic than low latitude populations [38], and many genes that play an important role in major pathways and *In(3R)Payne* inversion frequency were differentiated with latitudes [21]. Spatial genomic variation varied also with longitudes , for example, *D.melanogaster* populations that were collected from 32 different locations in Europe were compared at the genomic level and results showed that genomic variation and inversion polymorphisms were varied in east to west axis but surprisingly not in north to south axis. However, populations that were collected more than once (summer and fall) in a year were remained, seasonal variations could be observed within populations [24]. Other than whole genome studies, there are also phenotypical and molecular studies for spatial changes. Spatial and temporal variation of diapause in *D. melanogaster* populations from North America were studied by examining the variation in the *couch potato (cpo)* gene which is involved in many biological processes and it is found that the lower expression of *cpo* gene was leading to increment in the expression of diapause and also specific SNP polymorphism was correlated with latitude and showed a parallel clinality in 2009-2010 as in 1997 [39]. In another study, *Payne* inversion polymorphism was examined in *D. melanogaster* populations from Maine (USA) and Florida (USA) and the study showed that 50% of the Florida populations had *Payne* inversion polymorphism however, there was no inversion polymorphism in northern populations [25]. Climate change is one of many reasons for adaptation studies. To understand the underlying mechanism, a group created a model to observe effects of climate change on organisms and found that space had an advantage for selection. However, they claimed that a minor change in the conditions creates a huge pressure on organisms that may affect the viability of it [40].

2.3. Adaptation to temporal variation:

The environmental factors are the reason for spatial variances, therefore we may say that temporal variances also determine the viability of the organism and seasons are one of the main factors for temporal differences in temperate regions and previous studies suggested that temporal factors affects organisms[28,36,41–43]. A study showed that exposure to cold nights for a month caused an increase in brown adipose tissue and activity in humans [44]. Temporal factors are really important factors for the life of an organism. Recent studies suggested that low temperatures might be significant for selection, especially winter and early spring conditions may explain the greatest portion of the genetic variation due to climate[28,42,43]. In a study with *Arabidopsis thaliana*, it was concluded that 15.7% of the genomic variation could be explained with temporal factors, and 16.9% of the genomic variation could be explained with spatial factors which could be the result of isolation by distance. The portion of the spatial factors could be high and similar to climate portions due to isolation by distance[45,46]. However, when the spatial factors were removed from the analyses, it was found that most of the genomic variation could be explained by temperature of the minimum at growing season and precipitation in summer [46]. Another study found that starvation tolerance is positively correlated and desiccation tolerance is negatively correlated with lipid content in *D.ananassae* populations. Also flies that were collected from lower latitudes were more tolerant to starvation and less tolerant to desiccation stress which might be related to nutritional change due to change in temperature [47]. A comparative study of two European populations of cosmopolitan species, *D. melanogaster* and *D. simulans,* and four Afrotropical species, *D. yakuba, D. ananassae, D. iri* and *D. fraburu,* showed that European populations and *D.yakuba* which is a close species to *D. melanogaster* and endemic to Afrotropical region, had a broader range to thermal changes. However, other species that are endemic to tropical (*D. ananassae*, *D. iri* and *D. fraburu*) were cold sensitive and could not grow under 17 ℃ [41]. The study implied that spatial factors may have a role in thermal tolerance, however, space cannot explain everything by itself. Rapid changes could also affect organisms and they can give rapid responses at genomic level. In April of 2011, there was a heat wave and according to the reports across Spain and UK, it was the warmest April (NOAA, 2011), and to observe the effects of this heat-wave, they collected *Drosophila subobscura* samples from two different locations in Spain over two years (2011 and 2012). They found that the genetic constitution of spring 2011 populations was shifted to summer-like populations, and these changes were significantly related with the temperature changes [48]. These studies showed that temporal

changes affect organisms for many traits and temporal factors such as seasons may cause differences between populations.

Seasons are one of the important temporal effects on populations in temperate regions. Previous studies in *Drosophila* species demonstrated that different collection time of populations showed variation in various phenotypic traits such as immune response to infection, reproductive diapause, age distribution during collection time, chill coma recovery time, starvation stress and wing size [49–52] and genomic variation [22,53,54]. A study in North America with *D. melanogaster* showed that specific alleles were higher in spring and, these alleles decreased in fall and this oscillation continued over the years which suggested that seasons act as a selective factor [22]. In another study in North America, *D. melanogaster* populations that were collected from 15 different locations were examined for variation at genomic level and results indicated that seasons have a significant impact on the evolutionary process of *D. melanogaster* [27]. Another study in natural *D. melanogaster* populations showed that immune response to infection changed seasonally and flies that had seasonally changing alleles in *Tep3* region gave different response to infection to pathogens [53].Dobzhansky (1971), wrote that frequency of inversion polymorphism on the 3rd chromosome varied from month to month in two *Drosophila pseudoobscura* populations by referring to his study in 1938 [55]. Dobzhansky and Ayala [56] also supported the idea of seasonality of inversion polymorphism in *D. pseudoobscura* with another study in 1973.

2.4. Sequencing technologies and Pool-Seq method:

The whole genome of *D. melanogaster* was first sequenced in 2000 by using shotgun sequencing technology (Figure 2.1) with lots of cloning procedures and lots of effort [5]. At the end of 1990s, sequencing technology continued to develop. Pyrosequencing, the first of NGS technologies described in 1993 and completed in 2005. The working principle is similar to the “sequencing by synthesis” (SBS) method which came up in 2006 by Illumina. Pyrosequencing measures the release of pyrophosphates and measures the light however high reagent cost and error rate and low coverage are main disadvantages of this technology. SBS is the most popular method and the technique that is used for this thesis. SBS uses nucleotides with fluorescent labeled and terminating end and detect fluorescent after incorporation. Sequence by ligation is another method which uses 16 different octamers, and each consists of 2 bases. Probes ligated with ligase instead of DNA polymerase and the rest of the probe is removed. Disadvantages of this technology is the short reads. Ion Semiconductor sequencing

measures release of H+ ions, and the method is similar to pyrosequencing however high cost and time rate makes it disadvantageous. There are mainly four types of sequencing methods. Common features of these methods are library preparation by amplification and linking DNA to a solid surface via linkers (Figure 2.2). Improvement in the sequencing technology made genomic studies less expensive and faster than before, and also changed our approaches to genomic studies.

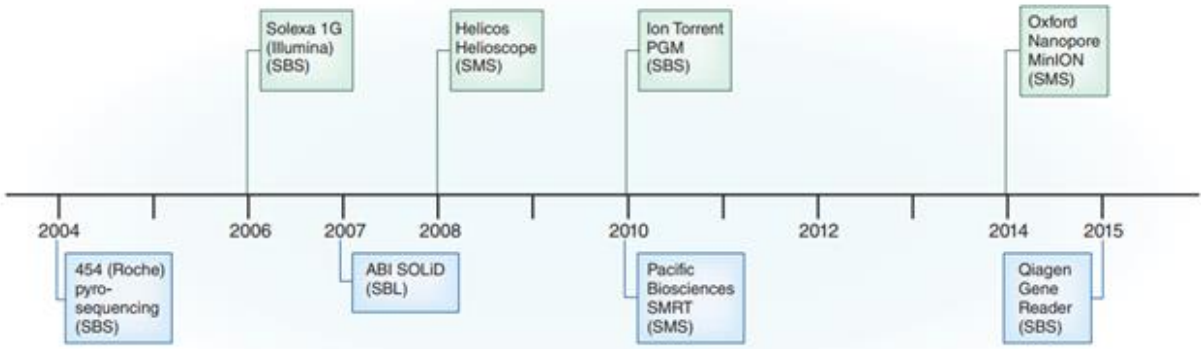

Figure 2. 1:Next Generation Sequencing technologies in years [57]

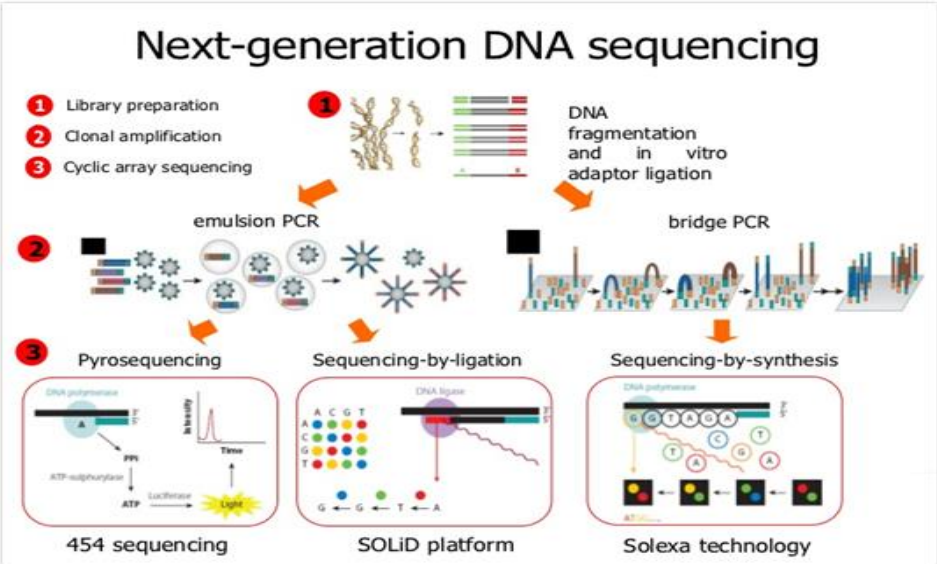

Figure 2. 2.Next Generation Sequencing technical differences. Pyrosequencing is one of first technılogy in the industry, which uses luciferase enzyme to detect nucleotides. Sequencing by ligation is the most popular one and uses flourscent labelled probes to detect

Next Generation Sequencing (NGS) is a slightly newer method that is commonly used in whole genome studies. It is a cheaper and faster method than its predecessors. However, using this technology for single genome sequencing is still expensive and slow. Therefore, Futschik and Schlötterer [57] suggested a method that is called "Pool sequencing" which is sequencing individuals as pools (Figure 2.3). It is a common method that is used in population genetics studies and many others. It is possible to observe common alleles in the population. Yet, this method still has some disadvantages due to copy number variations, it is hard to call low frequency alleles, the method mostly catches high frequency alleles and it is hard to distinguish them from sequencing errors [58]. However, it is still a fast and accurate method for whole genome studies.

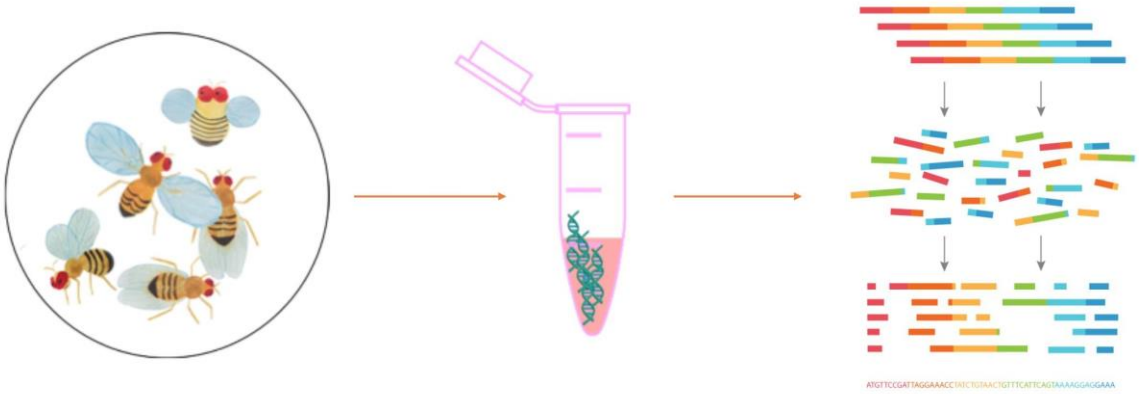

Figure 2. 3. Pool-seq method.

2.5. Statistical analysis for population studies:

Population genetics examines the evolutionary change of organism at the population level. Allele frequencies is one of the major issues when comparing populations at genomic level. Environmental factors may affect phenotype which are encoded by different alleles, or specific alleles could be selected for a better fitness, and it can go up to speciation when gene flow is ceased. There are parameters which helps us to understand the genomic variation in populations or direction of selection. These called Tajima's D, pi, and Watterson's theta. Watterson estimator (θ) used to describe nucleotide ratio of polymorphic sites [59]and π used for pairwise differences also describe as nucleotide diversity between populations [60]. Tajima's D [61]

used for understanding the direction of selection and calculated by using θ and π, it can be either lower or higher than 0, or could be equal to 0. If the Tajima's D is lower than zero that means population expansion or a bottleneck or selection of removing variation, which are decrease the nucleotide diversity, therefore Tajima's D decreases. If it is over zero, that indicates the population narrowing or in balance. Another common thing that is used in population genetics is fixation index which is also known as F_{ST} [62,63], a method to compare populations for genetic differentiation. F_{ST} may vary between 0 and 1 where 0 means no genetic differentiation between populations and 1 indicates a strong differentiation between populations.

3. MATERIALS AND METHODS

3.1. Sample preparation:

The flies (*Drosophila melanogaster)* were collected from Yeşiloz, Ankara in August and October of 2014 2015 and 2016. Yeşilöz is a small village, 100 km away from Ankara, Turkey (Figure 3.1). Yeşiloz is characterized by fruit and vegetable production. Because of this reason the passive transportation of *Drosophila melanogaster* through import fruits and vegetables is considered to be very low. Yeşilöz, is located in the Kirmir valley, and it is characterized by warm and humid climate

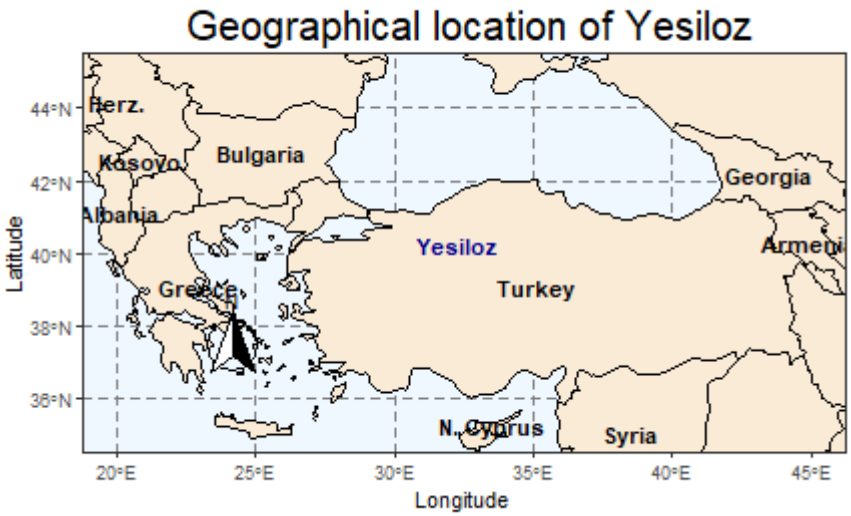

Figure 3. 1.Geographical location of Yeşilöz.

Flies (n = 40) were randomly selected for DNA isolation. The Pool-Seq method was used for DNA isolation. NEBNext Ultra DNA Lib Prep-24 and NEBNext Multiplex Oligos for Illumina-24 for 2014 samples (Table 3.2) were used for library preparation. Each pool (August and October of 2014) was sequenced as paired end on the NextSeq 500 platform by the Genomics Core Facility of the University Pompeu Fabra (UPF; Barcelona, Spain) with a coverage over 50X. Library preparation was done by using NEBNext Ultra II for 2015 and 2016 samples (Table 3.2) and each pool (August and October of 2015 and 2016) was sequenced as paired end on the Illumina HiSeq X platform by the NGX BIO sequencing service (San Francisco, USA) with coverage over 50X. The mean read lengths were 150bp for each sample.

DNA isolation, library preparation and sequencing was done by the **European Drosophila Population Genomics Consortium (DrosEU)**, and sequences were obtained from SRA (https://www.ncbi.nlm.nih.gov) under the project number PRJNA388788.

Table 3. 1.Sampling times and their accession numbers.

Year/months	**2014**	**2015**	**2016**
August	SRR5647749	SRR8439109	SRR8494463
October	SRR5647748	SRR8439107	SRR8494423

3.2. Estimation of genome-wide genetic variation and differentiation:

The raw data was checked with FastQC version 0.11.8 (Babraham Bioinformatics, UK) to avoid overrepresented sequences and bad quality. Adapter sequences of 2016 samples (Illumina Truseq adapter) were trimmed using cutadapt (v2.2)[64], and other sequences did not have adapters therefore adapter trimming did not apply to others. For the alignment, Burrows-Wheeler Alignment (bwa) was used with the parameters mem -M -t 24 (v0.7.17-r1188, [65]) and *Drosophila melanogaster Release 6* (GCF_000001215.4) was used as a reference genome. Mapping was applied with using samtools view (v1.9)[66,67] and aligned files were converted to Binary Alignment Map (BAM) format for further analysis, then sorted and indexed by using samtools sort and index parameters (version 1.9). Duplicates were removed using Picard MarkDuplicates using default parameters. These indexed files and the reference genome were used for creating a mpileup file for each year (2014, 2015, and 2016) using samtools mpileup function. A mpileup file is a pileup file that includes all sorted and indexed bam files and allows us to compare populations with each other. This mpileup files were used as an input for VarScan (v2.4.4,[68]) for SNP calling to create a Variant Calling Formation (vcf) file with the parameters mpileup2snp --output-vcf 1 --min-var-freq 0.01 --p-value 0.01 --min-coverage 15 --min-avg-qual 20. Then the vcf file was annotated using snpEff (v4.3),[69] using the *D. melanogaster* genome (BDGPG.6.86) as a reference. Annotated vcf files were converted to sync files by using VCF2Sync.py script provided by Martin Kapun (https://github.com/capoony/DrosEU_pipeline).

The sync file was used as an input file for PoPoolation (v.1.2.2) [70])and PoPoolation2, programs which were specialized for pooled data. Allele frequencies, F_{ST} values and Fisher's exact test values were calculated using PoPoolation2 (v.1201) [71]. Allele frequencies were calculated using snp-frequency-diff.pl and Fisher's exact test values were calculated for significance of the SNPs using fisher-test with (Table 3.3).

Table 3. 2. The parameters that were used for each script.

Program/year	**2014**	**2015**	**2016**
snp-frequency-diff.pl			
--min-count	6	6	6
--min-coverage	15	40	30
--max-coverage	80	150	110
fst-sliding.pl			
--min-coverage	15	40	30
--max-coverage	80	150	110
--window/step size	1/1	1/1	1/1
--pool-size	40	40	40
fisher-test.pl			
--min-count	6	6	6
--min-coverage	15	40	30
--max-coverage	80	150	110
--window/step size	1/1	1/1	1/1
--pool-size	40	40	40

The fst-sliding.pl output used as an input file for FST.py script provided by Martin Kapun (https://github.com/capoony/DrosEU_pipeline), which calculates F_{ST} values in windows (window-size=10000). These F_{ST} values were used for plotting in R(v.4.0.0)) to see distribution of F_{ST} values on chromosomes. The mpileup file used as an input file for calculation of Tajima's D, Tajima's Pi, and Watterson's Theta for PoPoolation tool with using Variance-sliding.pl script (Table 3.4).

Table 3. 3.Variance-sliding.pl parameters.

Program/year	**2014**	**2015**	**2016**
Variance-sliding.pl			
--window-size	1000	1000	1000
--step-size	500	500	500
--min-count	2	2	2
--min-coverage	15	40	30
--max-coverage	80	150	110
--min-qual	20	20	20
--pool-size(pi/theta/D)	40/40/160	40/40/160	40/40/160

3.3. Genetic differentiation associated with temporal variation

Minor allele frequencies that were obtained using popoolation2 were combined for three years and common positions were detected (~10^6) using *merge* in R (v.4.0.0), and MS Office excel "AND" and "EXACT" functions. We only used autosomes for further analyses which are 2L, 2R, 3L, 3R and 4 and sizes of chromosomes are 23,513,712 bp, 25,286,936 bp, 28,110,227 bp, 32,079,331 bp, and 1,348,131 bp, respectively. This was due to coverage which were much higher than autosomes therefore, we only used autosomes for more reliable results. Also, a study found that X chromosome had no contribution to traits involved in local adaptation [72] and many studies suggested that X-linked selection may differ for haploid-diploid organisms as in haploid-males and diploid-females [73–76]. We did not use the Y chromosome because assemblies based on short-read technology and this technology results with highly fragmented reads, with many gaps, uncertainty, and errors especially for repeat-rich regions, like centromeres, telomeres, or the Y chromosome [77,78]. The most of the Y chromosome newly assembled with a long single-molecule read in *D. melanogaster* [79] but we are not able to include the Y chromosome to our analyses. There was no sharing common positions in mitochondrion genome, therefore it could not be used in the analyses. After common positions were detected (~982,000 SNP), ANCOVA (Analysis of Covariance) was performed to detect correlation of each SNP (each minor allele frequencies of each position for each season) with season and year (anova(lm(SNP~season*year))) in R(v.4.0.0) and the ones that are significant

were selected (p value<0.05). Positions of significant SNPs were compared with the positions of results of the Fisher's exact test to find a sharing common positions, and then these sharing common positions were plotted. ANCOVA was applied to seasonal SNPs for detecting correlations between these SNPs (minor allele frequencies of each position for seasons and climate data in growing season them) with minimum temperature, maximum temperature and precipitation in growing season (Table 2) (anova(lm(SNPs~tmin+tmax+precipitation))) in R(v.4.0.0)]. Minimum temperature, maximum temperature (Table 3.5) and precipitation (Table 3) data was downloaded from WorldClim (https://www.worldclim.org) [80] with using *raster* package in R(v.4.0.0).

Table 3. 4.Minimum and maximum temperatures of each season for years (Data obtained from WorldClim2).

	Temperatures				**Precipitation**	
Year	**MTA**	**MTO**	**MATA**	**MATO**	**APA**	**APOCT**
2014	16.65134	12.58748	32.31490	26.17095	18.544447	53.289421
2015	15.93936	14.07898	31.40292	29.87547	3.563715	18.474703
2016	16.63988	11.98766	32.69441	26.96661	10.899306	27.047443

MTA: Minimum temperature of August, MTO:Minimum temperature of October, MATA: Maximum Temperature of August, MATO: Maximum Temperature of October, APA: Average Precipitation of August, APOCT: Average Precipitation of October.

Common positions, were detected for Tajima's D and nucleotide diversity (pi and theta), were discovered using *merge* and *apply* functions in R version 4.0.0 and plotted using *ggplot2* (Wickham, 2019) for each chromosome arm. Windows of seasonal SNPs were determined, and these positions were plotted with ggplot2 in R version 4.0.0. Fisher's exact test results were plotted by using the *qqman* package ([81]) in R version .4.0.0. F_{ST} values of seasonal SNPs and their means were calculated using *aggregate*. We also identified genes that contains these seasonal SNPs in R(4.0.0). Finally, all values were plotted by using the *ggplot2* package [82] in R for each year and each chromosome arm.

4. RESULTS

The data that was analyzed of 6 samples from Yeşilöz population collected in August and October in 2014 2015 and 2016 (Kapun et al., 2020, see Materials and Methods). Upon quality checks, each SNP file was generated separately for each year to compare seasons within years. The majority of the SNPs (~50%) were mapped to the introns. The distribution of the rest is as follows: 14.03% were downstream variants, 14.58% were upstream variants, 9.48% were exonic, and 8.72% were in the intergenic regions. We have identified a total of 6516 structural variants, of which 3874 insertions and 2642 deletions. 77% of these mutations are found to be silent, whereas 22% are missense. Highest base changes were between G to A and C to T and the transition transversion ratio was 1.10. We used only autosomes for further analyses. Genome wide, the highest numbers of variants were mapped to 2L and 3R. Table 4.1 shows average F_{ST} values between seasons for each year and chromosome. Among the year, F_{ST} values for 2014 were higher than the rest of the years; where the average F_{ST} values of 2014 were between 0.0261 and 0.0317, the average F_{ST} values of 2015 and 2016 were between 0.0158 and 0.0198. Highest F_{ST} values were found for 2014, for chromosome 4. For other years' samples, highest F_{ST} values were found for chromosome 2L and 3L, for 2015 and 2016 samples.

Table 4. 1.Mean F_{ST} values between seasons for chromosome arms and years.

	F_{ST} Values		
Chromosome	**2014**	**2015**	**2016**
2L	0.026097	0.019013	0.017921
2R	0.026977	0.015807	0.018186
3L	0.026436	0.016496	0.019858
3R	0.026758	0.016941	0.017861
4	0.031729	0.013698	0.018987

F_{ST} is a measure of the genetic differentiation. Figure 4.1. showed the distribution of site by site F_{ST} values for all the pairwise comparisons between August and October. The distribution is skewed to the right, where values point out a low genetic differentiation between seasons. But it is important to note the broad distribution of F_{ST} values along the genomes. Which demonstrates that some regions have genetically differ among seasons. As we can see in Figure 4.1, frequency of lower values was much higher in 2014 which indicates that in some regions 2014 samples were seasonally differentiated more than 2015 and 2016 samples, however 2015 and 2016 samples were differentiated within each other in a wider scale of the genome than 2014 samples.

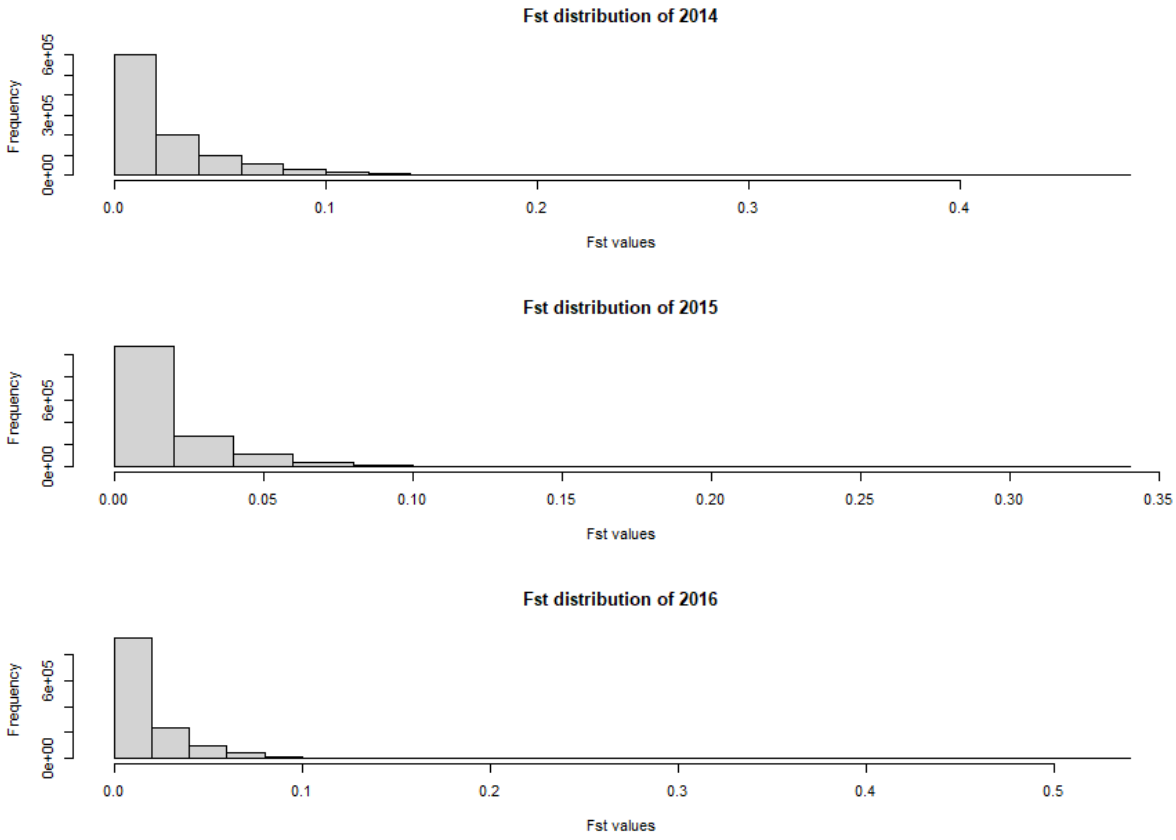

Figure 4. 1.F_{ST} distribution. Distribution of F_{ST} values observed for (a) August vs October 2014, (b) August vs October 2015 and (c) August vs October 2016. Values are calculated for differentiation of each SNP (window size and step size are 1 bp).

To investigate genetic differentiation among seasons and to see it in a broader range, we also calculated F_{ST} values in 10,000 bp window size. As we can see in Figure 4.2., pattern (blue line) seems similar for all years, yet F_{ST} values were lower which indicates that differentiation was lower between samples. However, 2014 samples were differentiated more than other two

samples in some regions. Also, there is something we need to consider. Highly differentiated bases were close to centromeric and telomeric regions on almost every chromosome, and also these high values rise in close regions for 2014 samples. On the other hand, rest of positions had a similar pattern for all years' genomes.

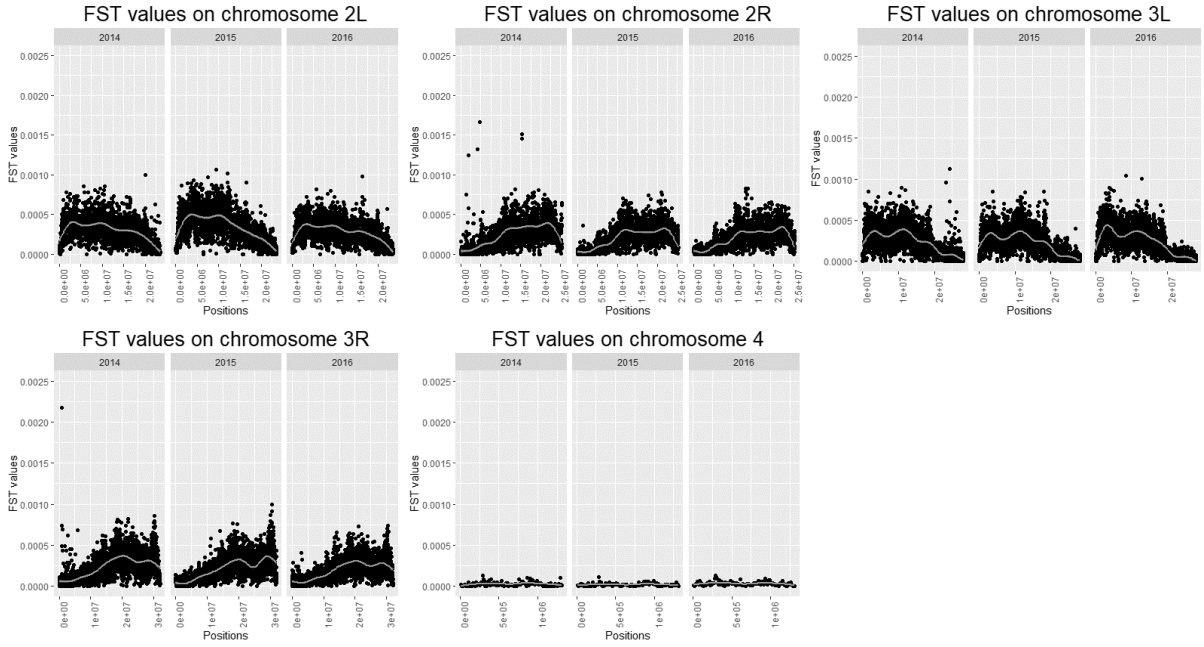

Figure 4. 2.F_{ST} values in 10,000 bp window size for each chromosome arm; 2L, 2R, 3L and 3R respectively. Blue lines: is the geom_smooth() method ggplot2 to observe the patterns of F_{ST} values. The method gam is used for 2L, 2R, 3L and 3R. The method loess is used for chromosome 4.

The files that had minor allele frequencies which were obtained by Popoolation2 (v.1201, Kofler *et al.*,2011) were compared with each other and common SNPs (positions) were detected for each year. Overall, there were ~983,000 SNP that were common for samples of all 3 years and 32,428 of them were seasonally significant (Table 4.2.). Seasonal significance was calculated with ANCOVA test by comparison of each SNP with seasons as factor (y=SNP ~ season*year) and p-values were obtained (Figure 4.3.). Majority of the p-values were between 0.01-0.05, yet there were some SNPs which were highly significant for seasonal changes ($p<0.0001$). However, only two of them were on genes mRpS2 and CG13532, and others were in intergenic regions. Seasonally significant SNPs were mainly on chromosome 2L and the lowest number of SNPs were on chromosome 4, the smallest of autosomes. 5701 and 3057 of these seasonal SNPs were in the common inversion *In(2L)t* and *In(3R)Payne* regions on chromosome 2L and 3R, respectively.

Table 4. 2.Total number of seasonally significant SNPs per chromosome.

Chromosome	**2L**	**2R**	**3L**	**3R**	**4**
Count of SNPs	9527	6406	8459	7994	42

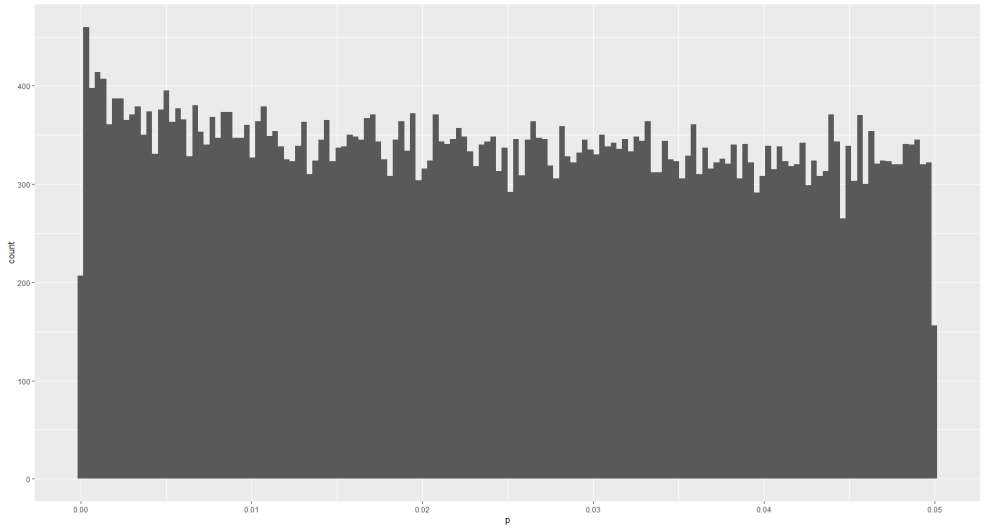

Figure 4. 3.The p-value distribution of seasonally significant SNPs that is obtained by ANCOVA test.

The -log(p-values) were calculated using Fisher's exact test script in Popoolation2 to understand the significance of SNPs and these were calculated merely comparing seasons within years. Following the detection of seasonally significant SNPs, we solely compared positions of Fisher's exact test results with them, yet we did not seek significance of -log(p-values), initially. As we can see, the Manhattan plot in Figure 4.4., 2015 and 2016 have more significant values than 2014. Even though their differentiation was lower than 2014 samples, the variation was more significant between seasons, and this was true for many of these SNPs within years.

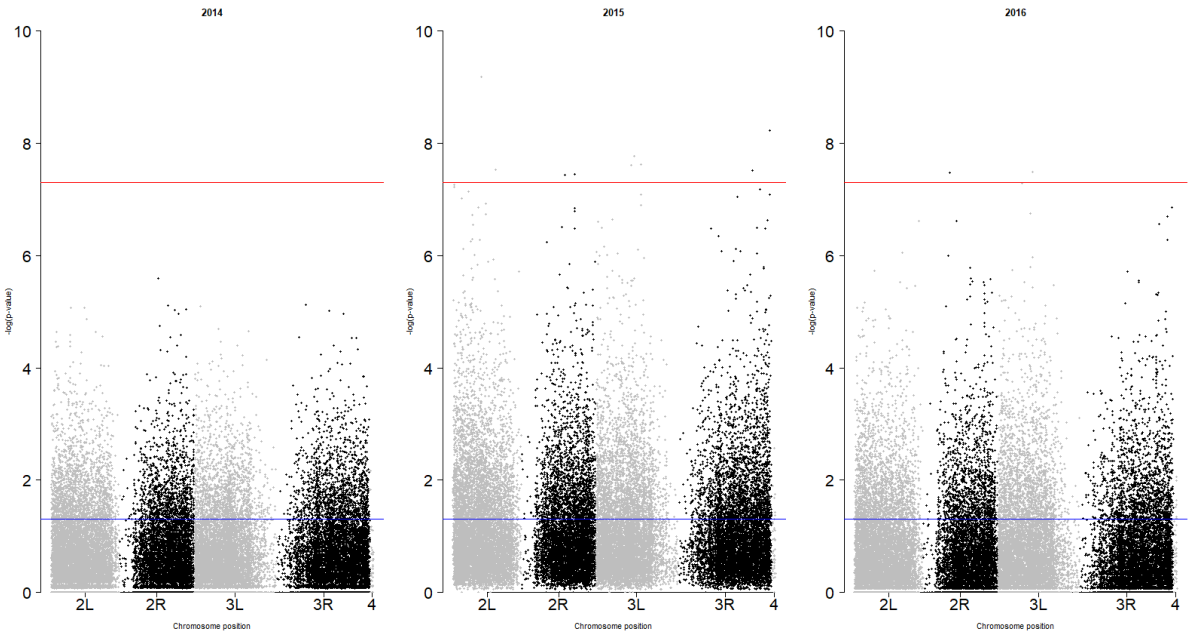

Figure 4. 4.Manhattan plots of seasonally significant SNPs. Blue line is a threshold of p = 0.05, and the red line is a threshold of p = 0.001 x 10-4 of Fisher's exact test values.

Afterwards, we named common and significant SNPs ($p < 0.05$) for all samples by comparing Fisher's exact test values within each other. Based on the results of Fisher's exact test the majority of these SNPs did not seem significant within years, however, this contradiction might be the result of comparing allele frequencies only within years as compared in two different samples. However, a broader comparison of positions between Fisher's exact test results and seasonally significant SNPs, the number of seasonal SNPs decreased more. Yet, we know that these SNPs were significant for seasonal changes regardless of years, but also allele frequency changes were significant within years regardless of season. As we can see in Figure 4.5, minor allele frequencies of seasonal SNPs were oscillating (Figure 4.5.), where some of them were increasing and others were decreasing from fall to summer and these were consistent through years. Some of these SNPs had drastic changes but some had smaller changes in allele frequencies yet changes in minor allele frequencies were significant within years.

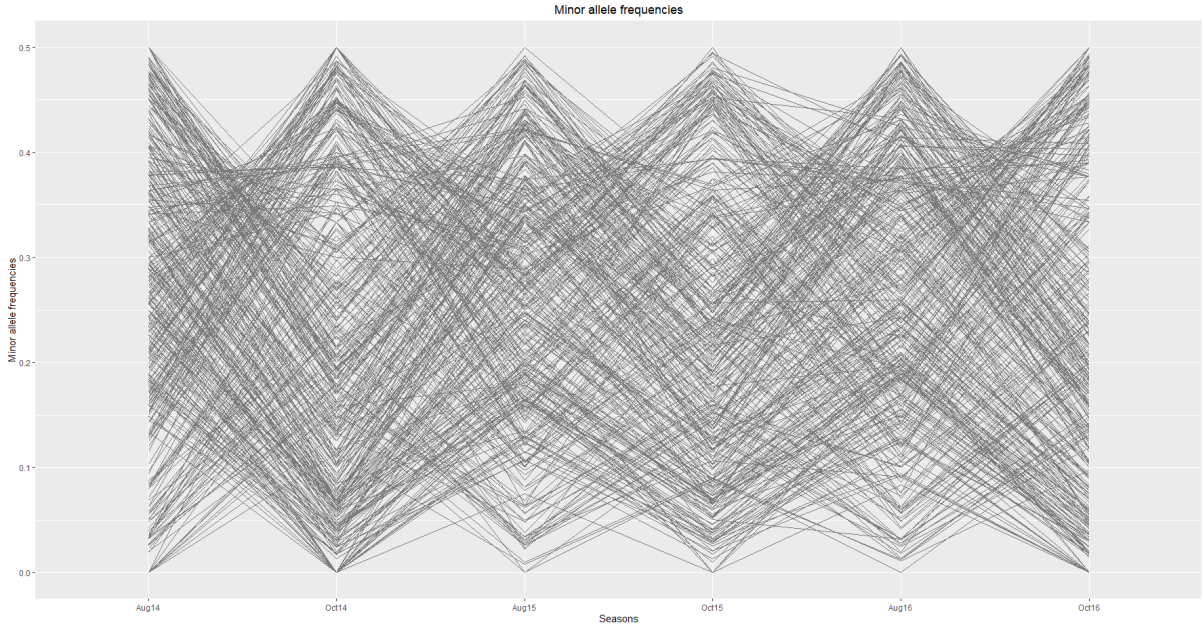

Figure 4. 5. Minor allele frequency distribution of seasonal SNPs. The positions are detected by comparing Fisher's exact test within each other and with seasonally significant SNPs.

Table 4.3. shows the number of significant SNPs that were calculated by comparison of seasonal SNPs with environmental factors with ANCOVA of, 5,623 of these seasonally significant SNPs which were significantly related for changes in the minimum temperature of the growing season. There were only 14 SNPs that were significant changes in the maximum temperature and none of these SNPs was related to precipitation of the growing season (See Material and Methods). Most of these SNPs were located the left arm of the second chromosome, 2L.

Table 4. 3.Total number of SNPs that are significantly related for changes in minimum temperature.

Chromosome	**2L**	**2R**	**3L**	**3R**	**4**
Count of SNPs	1702	1044	1567	1298	12

F_{ST} values were calculated using changes in allele frequencies (window and step size 1bp). We compared positions of F_{ST} results with all seasonally significant SNPs and identified F_{ST} values of these positions (Figure 4.6). The differentiation between 2014 samples were still higher than other years. Earlier results showed that chromosome 4 was the highest mean F_{ST} value of 2014 samples and still it is (Table 4.4). Yet in the Figure 4.6, some regions on chromosome 4 seem

to have high F_{ST} values, but those were not as high as the other chromosomes, and there were only 42 seasonally significant SNPs. Chromosomes 2L and 3L seem to have higher F_{ST} values than other chromosomes but not just for 2014 samples, other years' samples as well.

Table 4. 4.Mean F_{ST} of common positions for each chromosome arms, between seasons for every year.

	F_{ST} Values		
Chromosome	**2014**	**2015**	**2016**
2L	0.02447324	0.01793305	0.01626782
2R	0.02438678	0.01481126	0.01724404
3L	0.02492350	0.01502989	0.01805080
3R	0.02426471	0.01449722	0.01607840
4	0.02684138	0.01288048	0.01712163

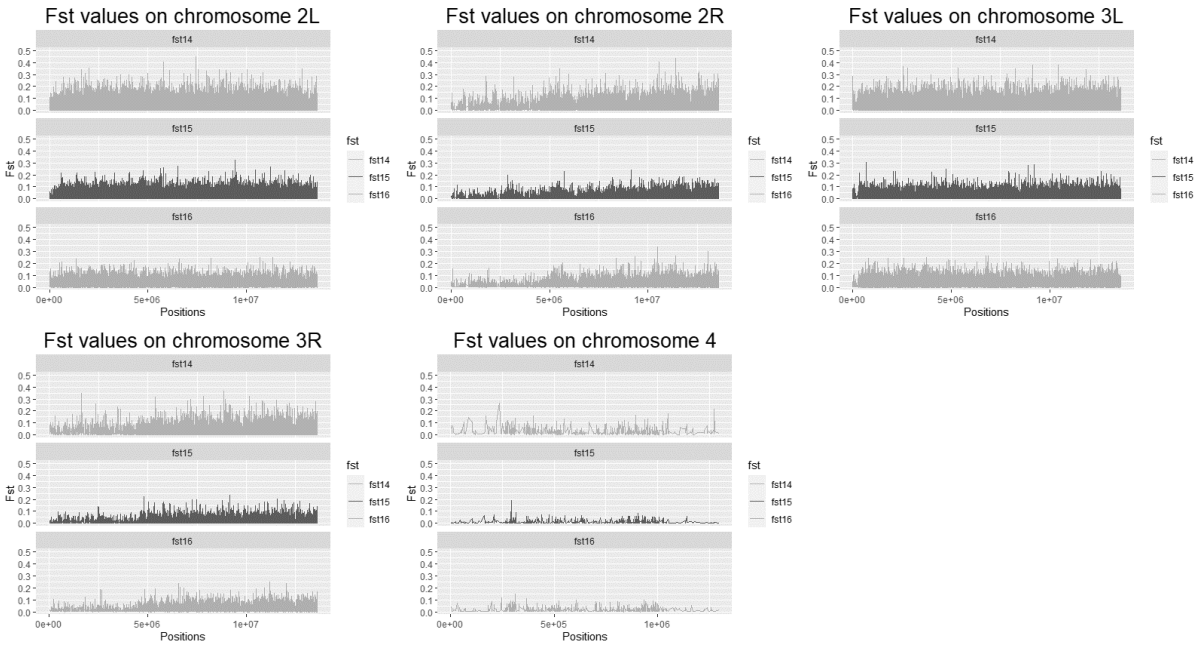

Figure 4. 6.F_{ST} values of seasonally significant SNPs for each chromosome arm. Red lines for 2014, green lines for 2015 and blue lines for 2016 samples.

The Tajima's D values were calculated for each year in windows (window size 1000 bp and step size 500 bp). Therefore, we found the windows of seasonal SNPs. As we can see in Figure 4.7., the 2014 population mostly has negative Tajima's D values for almost all of the windows of these SNPs and 2014 samples were mostly differentiated in many other results, as well, but we did not do any further analyses to understand the cause of this issue.

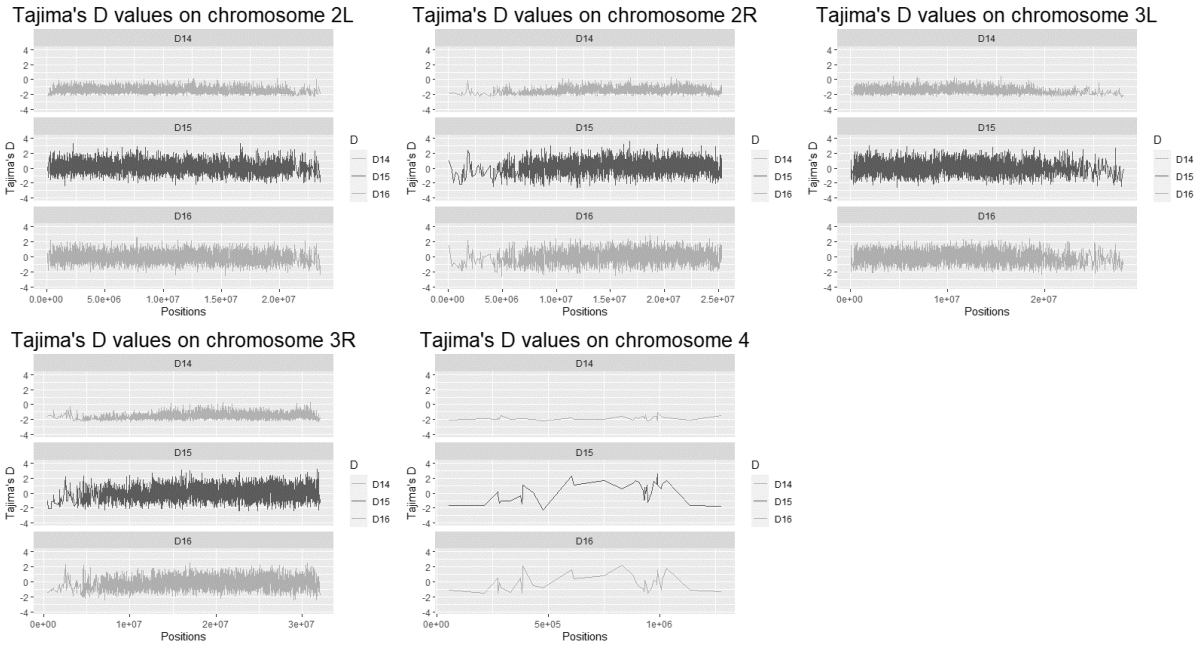

Figure 4. 7.Tajima's D values for each chromosome arm.

Lastly, we found protein coding genes that holds seasonally significant SNPs from the positions that we obtained from ANCOVA analysis, 71.41% of seasonally significant SNPs were in the protein coding genes. Number of protein coding genes on chromosome 2L (1377) might be lower than the chromosome 3R (1424), still it has the highest number of SNPs that were in protein coding genes (~6570 and 5740 in 2L and 3R, respectively).

Some of these genes seem highly polymorphic that contains many of seasonal SNPs (Figure 4.8.) such as *sickie (sick)* which has 87 of seasonal SNPs, *Phosphodiesterase 1c (Pde1c)*, *bruno1 (bru1)* or *toucan (toc)* on chromosome 2L, *muscleblind (mbl)*, *slowpoke2 (SLO2)*, *luna*, *Fish-lips (fili)* and *plexus (px)* on chromosome 2R, *bruno 3 (bru3)* on chromosome 3L that has 97 of seasonal SNPs. The chromosome 3R has especially many seasonally significant SNP containing genes such as *crossveinless c (cv-c)*, *fruitless (fru)*, *Shal K+ channel interacting protein (SKIP)*, *Dystrophin (Dys)*, *headcase (hdc)* and *couch potato (cpo)*. *Couch potato (cpo)* gene was previously identified as clinal and seasonal gene and it is also related with expression of diapause [45]. Molecular function of *cv-c* is lipid binding and many others. It also plays a role in sleep homeostasis and nuclear chromosome segregation mechanisms and it has 69 known alleles. *Fru* plays a vital role in the reproduction, male mating, and courtship behavior, and has 126 known alleles. *SKIP* plays a role in smell sense. *Sick* gene mediates the immune deficiency signaling pathway and involves defense against Gram-negative bacterium. *Bru1* and *Bru3* are RNA binding proteins where the product of *bru1* is necessary for gametogenesis, muscle organization, and developmental patterning, and *bru3* is involved in mRNA splicing. We also found seasonal SNPs in *Insulin-like receptor (InR)*, *Thioester-containing protein 2 (Tep2)* and *Thioester-containing protein 3 (Tep3)*. Furthermore, seasonal SNPs also found in *chico*, *hsp83*, *forkhead box, sub-group O (foxo)*, *ribosomal protein S6 kinase* (*s6k)*, *Lnk*, and *Tor* which are the members of Insulin-like Receptor Signaling Pathway of *D. melanogaster*.

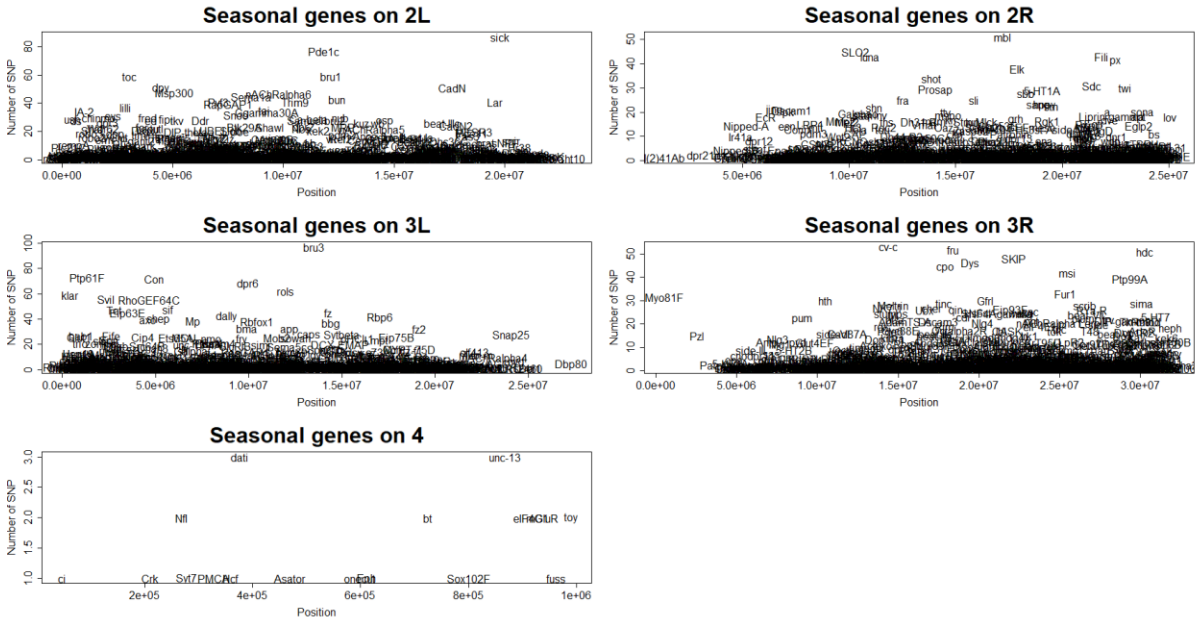

Figure 4. 8. Genes that contain seasonally significant SNPs. The number of SNPs (y axis) are plotted against the chromosome position (x axis).

5. DISCUSSION

Organisms that live in temperate regions could exposed to many different environmental fluctuations related to seasonal changes in temperature, humidity, sunlight, nutrients, etc... Cosmopolitan species such as *D. melanogaster* adapts to these environmental conditions and could live almost anywhere around the Earth but its close species such as *D. yakuba*, which is endemic to tropical Africa and in some traits, it can show a similar habits as *D. melanogaster*, yet cannot spread around the world as *D. melanogaster*. So, what is behind this high degree of adaptability in some organisms? There are many studies compared populations in spatial changes such as altitude, latitude and longitude [25,38,83,84]. As changes in location, temporal variations also have an impact on life, especially the regions that have harsh winter conditions [14,25,43,85]. Comparing spatial and temporal variation in populations is a widely used method to understand the molecular basis of adaptation and evolutionary change in organisms. To evaluate the molecular basis of adaptation in the context of the impact of temporal changes we used pool-seq whole genome data of a *D. melanogaster* population that was collected from Yeşilöz, Ankara (32.26 E, 40.231 N) in two seasons of three consecutive years. The city of Ankara may have harsh winters, but the climate of Yeşilöz differs from Ankara, where the climate is warmer and more humid, yet seasonal changes occur in the village. Previous studies with this population suggested some seasonal signs for some phenotypes such as heat, cold and starvation tolerance, where heat tolerance decreases, and cold tolerance increase in inbred lines that are collected in colder months. However, the same inbred lines showed different results for male and female in response to starvation tolerance [86]. Where females' starvation tolerance increased from spring to fall, and the converse was the case for males. This differentiation might be related to fat content because the population is preparing to get into diapause and this might be the result of increase in fat storage in females [87]. Previous studies with inbred lines from Yeşilöz populations supported this idea, where body weight and fat content were statistically higher in October lines against August lines [86]. When this is the case, can we see seasonal genetic variation? Although seasonal genetic variations have been documented in several studies [22,24,27,50,53,86,88–90] even so our understanding about seasonal adaptation is limited. To improve our understanding of seasonal adaptation, we aimed to analyze genome-wide seasonal genetic variation in *D. melanogaster* population that were collected across seasonal time.

There are few studies comparing populations within each other for seasonal changes and they found that some alleles vary seasonally in temperate *D. melanogaster* populations [22,27,85].

Our results also suggested that seasonal changes affect the genotype of the *D. melanogaster* population in a temperate region. Each of our samples has over 2 million SNPs that were different from the reference genome. When we compared these samples for common positions, there were ~983,000 SNPs. It might not be for every SNP on the genome but approximate 3.5% of the autosomal SNPs were consistently affected by seasonal changes over the years and also, seasonal changes in allele frequencies were mostly significant with changes in minimum temperature as suggested in one of previous studies [46]. It is possible that minimum temperature can act as a selective force in nature and shifts allele frequencies over seasons in coherence over years.

In general, the comparisons between genomes show low F_{ST} values of seasonally significant SNPs. However, few of these SNPs differed significantly between genomes (F_{ST}~0.5) but these differentiations were not maintained through years. These unbalanced differentiation between years, might be a result of rapid adaptation due to environmental changes. The highest mean F_{ST} for 2015 was consistent with other results because most of our SNPs were on chromosome 2L, and we have many SNPs that were related with minimum temperature on chromosome 3L, so 2016 results were also consistent with our findings. On the other hand, the highest mean F_{ST} for 2014 samples was on chromosome 4, which is surprising because we did not observe many SNPs for chromosome 4 in further analyses. This might be specific to 2014 samples, or due to the small size of chromosome 4, the proportion of SNPs might result as this. Because we also found that 2014 populations had mostly negative Tajima's D values which indicates to population expansion or a bottleneck, however we did not see a consistency of negative Tajima's D values for other years. We can only speculate that the population might be exposed to an alteration in environmental factors such as temperature changing, or usage of pesticides, or predator effect or competition with other species. It might also be related with microbiota or nutrients; however, we did not have enough information about that. According to the National Oceanic and Atmospheric Administration (NOAA) the warmest of these three years was 2016, and the coolest was 2014, however according to a research done by NASA in 2015, 2014 was the warmest year until then, hence population might be affected from rapid changes in heat. According to WorldClim2 data, 2014, especially the growing season for October samples, was the most humid season of all these years. We have not find any relationship between seasonal SNPs and humidity data, but, 2014 samples could give a quick response to these environmental changes and might develop a rapid adaptation as seen in Spain populations faced with heat waves in 2011 [53,85]. Rapid changes such as heat waves could affect organisms at the genome level, especially in some organisms that give many offspring during the year [17]. The linkage

disequilibrium (LD) is the association of alleles which decreases the recombination between them. Low nucleotide diversity in 2014 populations and low F_{ST} values of all of years' samples might be also related with linkage disequilibrium because some of these regions were close and also some of those were in inversion regions. It is known that recombination is very low in inversion regions due to linkage disequilibrium [22]. In our previous studies, we found that some inversion polymorphisms are seasonal in the Yeşilöz population [91]. The frequency of these inversion polymorphisms were not that high in the inbred lines yet some individuals in our population carry these inversions and frequency of these were seasonally altering [91]. We did not do further analyses for this matter but according to our results, 2014 samples experienced some issues that cause reduction in the population diversity.

The genomic positions of seasonally varied SNPs in the Yeşilöz population were found in 5,396 different genes, and 71.41% of these SNPs were in gene regions. Some of these genes such as *Bru1* and *Bru3* on chromosomes 2L and 3L, respectively are RNA binding proteins. Other genes that had seasonally significant SNPs in 2L belongs to *Tep2* and *Tep3* genes with two and one seasonal SNPs respectively, and in a study they found that flies that had seasonal SNPs in *Tep3* gene gave different immune responses to infection [53]. We also found another gene on chromosome 2L, *sick*, that contains 87 seasonal SNPs and it is involved in immune defense against Gram-negative bacterium.

The right arm of the third chromosome had many genes that contain seasonal SNPs. The *cv-c* gene, which is involved in lipid metabolism, and *fru* that is important for male courtship behavior have many seasonal SNPs within. Some of SNPs were in *cpo* which is a gene that is known to be clinal and associated with diapause [85,92]. There are also other studies suggesting the seasonality of the *cpo* [39,85]. Our results showed that 44 SNPs in *cpo* significantly vary between seasons, whereas Rodrigues and colleagues (2020), found 14 SNPs in *cpo* that are seasonally significant [93]. We also found that 4 of 44 seasonal SNPs in *cpo* are highly significant for seasonal changes ($p<0.01$), and allele frequencies of 3 of them are increasing and one of them is decreasing from summer to fall. In the same study, Rodrigues and colleagues [93] found that the *InR* gene is not significant for seasonal changes but we found 3 SNPs in the *InR* gene region that are seasonally significant. This differentiation in number of seasonal SNP in genes might be the result of usage of different method while naming genes or studying with different populations. We also found genes like *chico, hsp83, foxo, s6k, Lnk, tor,* and *InR* with seasonal SNPs that are the members of Insulin-like Receptor Signaling Pathway of *D. melanogaster*.

Previous studies with the inbred lines of Yeşilöz population, showed that inversions *In(2L)t* have seasonal cycles, especially populations mostly carry homozygous *In(2L)t* during spring however, it decreased and heterozygotic *In(2L)t* increased during the fall [91]. *In(2L)t* seems like the most carried cosmopolitan inversion for the Yeşilöz populations. Another study with our population showed that *In(2L)t* and *In(3L)P* had seasonality; frequencies of this cosmopolitan inversions increased during the summer and decreased during fall [86]. Some of the genes that were identified in the current study were in these inversion regions, altering frequency of inversion polymorphism may also affect seasonal changes in allele frequencies due to lack of recombination in inversion regions because of strong linkage disequilibrium [22]. Rapid environmental changes alter the selective pressure in seasonal environments. Understanding the evolutionary impact of seasonality on organism is a fundamental interest in evolutionary biology. Our results suggested that genetic polymorphism contribute rapidly in response to seasonality through cyclic changes in allele frequencies, in genes associated with adaptation. What could be there to do further? We propound many positions and genes that are likely to be affected by seasonal changes. Number of genes were remarkably high; this might be the result of using different technique for identifying genes. Applying other types of analyses, the number of positions and genes that contains these positions might be decrease. However, in some regions, we believe that seasonality is strong such as *cpo*. Therefore, these SNPs and genes might observed for many more years to see if these positions are still affected by seasonal changes for a longer period because we only compared 3 years' samples. Signaling pathways that these seasonal genes involved could be studied and changes could be observed at protein level, especially *Bru1* and *Bru3* since they are RNA binding proteins, or genes that are involved in Insulin signaling pathway. The similar study could be done with female flies, because we only used male flies for this study. Moreover, there are some studies suggested that selection of genes might be different for male and females [73–76]. Although most of these studies were about X-chromosome but when we examined the starvation data from previous studies with the inbred lines of Yeşilöz population, we saw that starvation tolerance increased in females but decreased in males from spring to fall [91]. Also, some of our candidate SNPs were related with male behavior, so studying genome-wide for both sexes may give different results in this manner. Further investigation will/might be required to understand reasons behind low Tajima's D values of 2014 samples, because it might the result of rapid adaptation.

6. REFERENCES:

[1] L.B. Barreiro, G. Laval, H. Quach, E. Patin, L. Quintana-Murci, Natural selection has driven population differentiation in modern humans, Nat. Genet. 40 (2008) 340–345. https://doi.org/10.1038/ng.78.

[2] K.E. Holsinger, Natural Selection*, in: Encycl. Genet., Elsevier, 2001: pp. 1291–1297. https://doi.org/10.1006/rwgn.2001.1161.

[3] J.T.C. Roger Conant, A field guide to reptiles & amphibians : eastern and central North America, Houghton Mifflin, 1998.

[4] A. Corl, K. Bi, C. Luke, A.S. Challa, A.J. Stern, B. Sinervo, R. Nielsen, The Genetic Basis of Adaptation following Plastic Changes in Coloration in a Novel Environment, Curr. Biol. 28 (2018) 2970-2977.e7. https://doi.org/10.1016/j.cub.2018.06.075.

[5] M.D. Adams, S.E. Celniker, R.A. Holt, C.A. Evans, J.D. Gocayne, P.G. Amanatides, S.E. Scherer, P.W. Li, R.A. Hoskins, R.F. Galle, R.A. George, S.E. Lewis, S. Richards, M. Ashburner, S.N. Henderson, G.G. Sutton, J.R. Wortman, M.D. Yandell, Q. Zhang, L.X. Chen, R.C. Brandon, Y.H.C. Rogers, R.G. Blazej, M. Champe, B.D. Pfeiffer, K.H. Wan, C. Doyle, E.G. Baxter, G. Helt, C.R. Nelson, G.L. Gabor Miklos, J.F. Abril, A. Agbayani, H.J. An, C. Andrews-Pfannkoch, D. Baldwin, R.M. Ballew, A. Basu, J. Baxendale, L. Bayraktaroglu, E.M. Beasley, K.Y. Beeson, P. V. Benos, B.P. Berman, D. Bhandari, S. Bolshakov, D. Borkova, M.R. Botchan, J. Bouck, P. Brokstein, P. Brottier, K.C. Burtis, D.A. Busam, H. Butler, E. Cadieu, A. Center, I. Chandra, J. Michael Cherry, S. Cawley, C. Dahlke, L.B. Davenport, P. Davies, B. de Pablos, A. Delcher, Z. Deng, A. Deslattes Mays, I. Dew, S.M. Dietz, K. Dodson, L.E. Doup, M. Downes, S. Dugan-Rocha, B.C. Dunkov, P. Dunn, K.J. Durbin, C.C. Evangelista, C. Ferraz, S. Ferriera, W. Fleischmann, C. Fosler, A.E. Gabrielian, N.S. Garg, W.M. Gelbart, K. Glasser, A. Glodek, F. Gong, J. Harley Gorrell, Z. Gu, P. Guan, M. Harris, N.L. Harris, D. Harvey, T.J. Heiman, J.R. Hernandez, J. Houck, D. Hostin, K.A. Houston, T.J. Howland, M.H. Wei, C. Ibegwam, M. Jalali, F. Kalush, G.H. Karpen, Z. Ke, J.A. Kennison, K.A. Ketchum, B.E. Kimmel, C.D. Kodira, C. Kraft, S. Kravitz, D. Kulp, Z. Lai, P. Lasko, Y. Lei, A.A. Levitsky, J. Li, Z. Li, Y. Liang, X. Lin, X. Liu, B. Mattei, T.C. McIntosh, M.P. McLeod, D. McPherson, G. Merkulov, N. V. Milshina, C.

Mobarry, J. Morris, A. Moshrefi, S.M. Mount, M. Moy, B. Murphy, L. Murphy, D.M. Muzny, D.L. Nelson, D.R. Nelson, K.A. Nelson, K. Nixon, D.R. Nusskern, J.M. Pacleb, M. Palazzolo, G.S. Pittman, S. Pan, J. Pollard, V. Puri, M.G. Reese, K. Reinert, K. Remington, R.D.C. Saunders, F. Scheeler, H. Shen, B. Christopher Shue, I. Siden-Kiamos, M. Simpson, M.P. Skupski, T. Smith, E. Spier, A.C. Spradling, M. Stapleton, R. Strong, E. Sun, R. Svirskas, C. Tector, R. Turner, E. Venter, A.H. Wang, X. Wang, Z.Y. Wang, D.A. Wassarman, G.M. Weinstock, J. Weissenbach, S.M. Williams, T. Woodage, K.C. Worley, D. Wu, S. Yang, Q. Alison Yao, J. Ye, R.F. Yeh, J.S. Zaveri, M. Zhan, G. Zhang, Q. Zhao, L. Zheng, X.H. Zheng, F.N. Zhong, W. Zhong, X. Zhou, S. Zhu, X. Zhu, H.O. Smith, R.A. Gibbs, E.W. Myers, G.M. Rubin, J. Craig Venter, The genome sequence of Drosophila melanogaster, Science (80-.). 287 (2000) 2185–2195. https://doi.org/10.1126/science.287.5461.2185.

[6] J.R. David, P. Capy, Genetic variation of Drosophila melanogaster natural populations, Trends Genet. 4 (1988) 106–111. https://doi.org/10.1016/0168-9525(88)90098-4.

[7] D. Lachaise, M.-L. Cariou, J.R. David, F. Lemeunier, L. Tsacas, M. Ashburner, Historical Biogeography of the Drosophila melanogaster Species Subgroup, in: Evol. Biol., 1988: pp. 159–225. https://doi.org/10.1007/978-1-4613-0931-4_4.

[8] K. Thornton, P. Andolfatto, Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of Drosophila melanogaster, Genetics. 172 (2006) 1607–1619. https://doi.org/10.1534/genetics.105.048223.

[9] W. Stephan, H. Li, The recent demographic and adaptive history of Drosophila melanogaster, Heredity (Edinb). 98 (2007) 65–68. https://doi.org/10.1038/sj.hdy.6800901.

[10] S.J.Y. Laurent, A. Werzner, L. Excoffier, W. Stephan, Approximate Bayesian Analysis of Drosophila melanogaster polymorphism data reveals a recent colonization of Southeast Asia, Mol. Biol. Evol. 28 (2011) 2041–2051. https://doi.org/10.1093/molbev/msr031.

[11] A.D. Bradshaw, Evolutionary Significance of Phenotypic Plasticity in Plants, Adv. Genet. 13 (1965) 115–155. https://doi.org/10.1016/S0065-2660(08)60048-6.

[12] M.J. West-Eberhard, Phenotypic plasticity and the origins of diversity, Annu. Rev. Ecol. Syst. Vol. 20. (1989) 249–278. https://doi.org/10.1146/annurev.es.20.110189.001341.

[13] M. Pigliucci, K. Hayden, Phenotypic plasticity is the major determinant of changes in phenotypic integration in Arabidopsis, New Phytol. 152 (2001) 419–430. https://doi.org/10.1046/j.0028-646X.2001.00275.x.

[14] H.M. Stone, P.A. Erickson, A.O. Bergland, Phenotypic plasticity, but not genetic adaptation, underlies seasonal variation in the cold hardening response of {Drosophila} melanogaster, BioRxiv. (2019) 691741. https://doi.org/10.1101/691741.

[15] D.L. Denlinger, Regulation of Diapause, Annu. Rev. Entomol. (2002). https://doi.org/10.1146/annurev.ento.47.091201.145137.

[16] V. Koštál, Eco-physiological phases of insect diapause, J. Insect Physiol. (2006). https://doi.org/10.1016/j.jinsphys.2005.09.008.

[17] J. Gillespie, Polymorphism in random environments, Theor. Popul. Biol. (1973). https://doi.org/10.1016/0040-5809(73)90028-2.

[18] U. Becker, G. Colling, P. Dostal, A. Jakobsson, D. Matthies, Local adaptation in the monocarpic perennial Carlina vulgaris at different spatial scales across Europe, Oecologia. 150 (2006) 506–518. https://doi.org/10.1007/s00442-006-0534-9.

[19] I. Karl, T. Schmitt, K. Fischer, Genetic differentiation between alpine and lowland populations of a butterfly is related to PGI enzyme genotype, Ecography (Cop.). 32 (2009) 488–496. https://doi.org/10.1111/j.1600-0587.2008.05660.x.

[20] C.L. Mott, Environmental Constraints to the Geographic Expansion of Plant and Animal Species, Nat. Educ. Knowl. 3 (2010) 72. https://www.nature.com/scitable/knowledge/library/environmental-constraints-to-the-geographic-expansion-of-13236052/ (accessed November 29, 2019).

[21] D.K. Fabian, M. Kapun, V. Nolte, R. Kofler, P.S. Schmidt, C. Schlötterer, T. Flatt, Genome-wide patterns of latitudinal differentiation among populations of Drosophila melanogaster from North America, Mol. Ecol. 21 (2012) 4748–4769. https://doi.org/10.1111/j.1365-294X.2012.05731.x.

[22] A.O. Bergland, E.L. Behrman, K.R. O'Brien, P.S. Schmidt, D.A. Petrov, Genomic Evidence of Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in Drosophila, PLoS Genet. 10 (2014). https://doi.org/10.1371/journal.pgen.1004775.

[23] D.K. Fabian, J.B. Lack, V. Mathur, C. Schlötterer, P.S. Schmidt, J.E. Pool, T. Flatt, Spatially varying selection shapes life history clines among populations of Drosophila melanogaster from sub-Saharan Africa, J. Evol. Biol. 28 (2015) 826–840. https://doi.org/10.1111/jeb.12607.

[24] M. Kapun, M.G. Barrón, F. Staubach, D.J. Obbard, R.A.W. Wiberg, J. Vieira, C. Goubert, O. Rota-Stabelli, M. Kankare, M. Bogaerts-Márquez, A. Haudry, L. Waidele, I. Kozeretska, E.G. Pasyukova, V. Loeschcke, M. Pascual, C.P. Vieira, S. Serga, C. Montchamp-Moreau, J. Abbott, P. Gibert, D. Porcelli, N. Posnien, A. Sánchez-Gracia, S. Grath, É. Sucena, A.O. Bergland, M.P.G. Guerreiro, B.S. Onder, E. Argyridou, L. Guio, M.F. Schou, B. Deplancke, C. Vieira, M.G. Ritchie, B.J. Zwaan, E. Tauber, D.J. Orengo, E. Puerma, M. Aguadé, P.S. Schmidt, J. Parsch, A.J. Betancourt, T. Flatt, J. González, Genomic analysis of European Drosophila melanogaster populations reveals longitudinal structure, continent-wide selection, and previously unknown DNA viruses, Mol. Biol. Evol. (2020). https://doi.org/10.1093/molbev/msaa120.

[25] E. Durmaz, C. Benson, M. Kapun, P. Schmidt, T. Flatt, An inversion supergene in Drosophila underpins latitudinal clines in survival traits, J. Evol. Biol. 31 (2018) 1354–1364. https://doi.org/10.1111/jeb.13310.

[26] E. Long, C. Evans, J. Chaston, J.A. Udall, Genomic structural variations within five continental populations of drosophila melanogaster, G3 Genes, Genomes, Genet. 8 (2018) 3247–3253. https://doi.org/10.1534/g3.118.200631.

[27] H.E. Machado, A.O. Bergland, R. Taylor, S. Tilk, E. Behrman, K. Dyer, D.K. Fabian, T. Flatt, J. González, T.L. Karasov, I. Kozeretska, B.P. Lazzaro, T.J. Merritt, J.E. Pool, K. O'Brien, S. Rajpurohit, P.R. Roy, S.W. Schaeffer, S. Serga, P. Schmidt, D.A. Petrov, Broad geographic sampling reveals predictable, pervasive, and strong seasonal adaptation in Drosophila, BioRxiv. (2019) 337543. https://doi.org/10.1101/337543.

[28] J. Ågren, D.W. Schemske, Reciprocal transplants demonstrate strong adaptive differentiation of the model organism Arabidopsis thaliana in its native range, New

Phytol. 194 (2012) 1112–1122. https://doi.org/10.1111/j.1469-8137.2012.04112.x.

[29] A. Werzner, P. Pavlidis, L. Ometto, W. Stephan, S. Laurent, Selective Sweep in the Flotillin-2 Region of European Drosophila melanogaster, PLoS One. 8 (2013). https://doi.org/10.1371/journal.pone.0056629.

[30] D.J. Begun, C.F. Aquadro, African and North American populations of Drosophila melanogaster are very different at the DNA level, Nature. 365 (1993) 548–550. https://doi.org/10.1038/365548a0.

[31] S. Glinka, L. Ometto, S. Mousset, W. Stephan, D. De Lorenzo, Demography and Natural Selection Have Shaped Genetic Variation in Drosophila melanogaster: A Multi-locus Approach, Genetics. 165 (2003) 1269–1278.

[32] S.S. Saminadin-Peter, C. Kemkemer, P. Pavlidis, J. Parsch, Selective sweep of a cis-regulatory sequence in a non-African population of Drosophila melanogaster, Mol. Biol. Evol. 29 (2012) 1167–1174. https://doi.org/10.1093/molbev/msr284.

[33] N.G. Jablonski, G. Chaplin, Human skin pigmentation as an adaptation to UV radiation, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 8962–8968. https://doi.org/10.1073/pnas.0914628107.

[34] T. Caro, T.N. Sherratt, M. Stevens, The ecology of multiple colour defences, Evol. Ecol. 30 (2016) 797–809. https://doi.org/10.1007/s10682-016-9854-3.

[35] W.B. Watt, Bioenergetics and Evolutionary Genetics: Opportunities for New Synthesis, Am. Nat. 125 (1985) 118–143. https://doi.org/10.1086/284331.

[36] K. Fischer, I. Karl, Exploring plastic and genetic responses to temperature variation using copper butterflies, Clim. Res. 43 (2010) 17–30. https://doi.org/10.3354/cr00892.

[37] D.K. Fabian, J.B. Lack, V. Mathur, C. Schlötterer, P.S. Schmidt, J.E. Pool, T. Flatt, Spatially varying selection shapes life history clines among populations of Drosophila melanogaster from sub-Saharan Africa, J. Evol. Biol. 28 (2015) 826–840. https://doi.org/10.1111/jeb.12607.

[38] J.A. Reinhardt, B. Kolaczkowski, C.D. Jones, D.J. Begun, A.D. Kern, Parallel geographic variation in Drosophila melanogaster, Genetics. 197 (2014) 361–373.

https://doi.org/10.1534/genetics.114.161463.

[39] R. Cogni, C. Kuczynski, S. Koury, E. Lavington, E.L. Behrman, K.R. O'Brien, P.S. Schmidt, W.F. Eanes, The intensity of selection acting on the couch potato gene-spatial-temporal variation in a diapause cline, Evolution (N. Y). 68 (2014) 538–548. https://doi.org/10.1111/evo.12291.

[40] C.A. Botero, F.J. Weissing, J. Wright, D.R. Rubenstein, Evolutionary tipping points in the capacity to adapt to environmental change, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 184–189. https://doi.org/10.1073/pnas.1408589111.

[41] Y. Cohet, J. Vouidibio, J.R. David, Thermal tolerance and geographic distribution: A comparison of cosmopolitan and tropical endemic Drosophila species, J. Therm. Biol. 5 (1980) 69–74. https://doi.org/10.1016/0306-4565(80)90002-9.

[42] M.H. Hoffmann, Biogeography of Arabidopsis thaliana (L.) Heynh. (Brassicaceae), J. Biogeogr. (2002). https://doi.org/10.1046/j.1365-2699.2002.00647.x.

[43] M.A. Hannah, D. Wiese, S. Freund, O. Fiehn, A.G. Heyer, D.K. Hincha, Natural genetic variation of freezing tolerance in arabidopsis, Plant Physiol. (2006). https://doi.org/10.1104/pp.106.081141.

[44] P. Lee, S. Smith, J. Linderman, A.B. Courville, R.J. Brychta, W. Dieckmann, C.D. Werner, K.Y. Chen, F.S. Celi, Temperature-acclimated brown adipose tissue modulates insulin sensitivity in humans, Diabetes. 63 (2014) 3686–3698. https://doi.org/10.2337/db14-0513.

[45] T.F. Sharbel, B. Haubold, T. Mitchell-Olds, Genetic isolation by distance in Arabidopsis thaliana: Biogeography and postglacial colonization of Europe, Mol. Ecol. (2000). https://doi.org/10.1046/j.1365-294X.2000.01122.x.

[46] J.R. Lasky, D.L. Des Marais, J.K. McKay, J.H. Richards, T.E. Juenger, T.H. Keitt, Characterizing genomic variation of Arabidopsis thaliana: The roles of geography and climate, Mol. Ecol. 21 (2012) 5512–5529. https://doi.org/10.1111/j.1365-294X.2012.05709.x.

[47] S. SISODIA, B.N. SINGH, Resistance to environmental stress in Drosophila ananassae:

latitudinal variation and adaptation among populations, J. Evol. Biol. 23 (2010) 1979–1988. https://doi.org/10.1111/j.1420-9101.2010.02061.x.

[48] F. Rodríguez-Trelles, R. Tarrío, M. Santos, Genome-wide evolutionary response to a heat wave in Drosophila, Biol. Lett. 9 (2013). https://doi.org/10.1098/rsbl.2013.0228.

[49] P.S. Schmidt, D.R. Conde, ENVIRONMENTAL HETEROGENEITY AND THE MAINTENANCE OF GENETIC VARIATION FOR REPRODUCTIVE DIAPAUSE IN DROSOPHILA MELANOGASTER, Evolution (N. Y). (2006). https://doi.org/10.1554/05-430.1.

[50] E.L. Behrman, S.S. Watson, K.R. O'Brien, M.S. Heschel, P.S. Schmidt, Seasonal variation in life history traits in two Drosophila species, J. Evol. Biol. 28 (2015) 1691–1704. https://doi.org/10.1111/jeb.12690.

[51] B.S. Subaşı, Drosophila melanogaster'de Coğrafi ve Mevsimsel Diyapoz Değişimi ve İnsülin-Benzeri Reseptör (lnR) Polimorfizmi ile İlişkisinin Araştırılması, Hacettepe Üniversitesi, 2017.

[52] C.F. Aksoy, Enlemsel Klinlerin Drosophila melanogaster'de Dca Genotipi ve Kanat Büyüklüğü Bakımından Araştırılması, Hacettepe University, 2018.

[53] E.L. Behrman, V.M. Howick, M. Kapun, F. Staubach, A.O. Bergland, D.A. Petrov, B.P. Lazzaro, P.S. Schmidt, Rapid seasonal evolution in innate immunity of wild Drosophila melanogaster, Proc. R. Soc. B Biol. Sci. 285 (2018). https://doi.org/10.1098/rspb.2017.2599.

[54] P.A. Umina, A.R. Weeks, M.R. Kearney, S.W. McKechnie, A.A. Hoffmann, Evolution: A rapid shift in a classic clinal pattern in Drosophila reflecting climate change, Science (80-.). 308 (2005) 691–693. https://doi.org/10.1126/science.1109523.

[55] J.R. Powell, Progress and Prospects in Evolutionary Biology: The Drosophila Model - Jeffrey R. Powell, Professor Director of Undergraduate Studies and Chairman of Organismal Biology, 1997. https://books.google.com.tr/books/about/Progress_and_Prospects_in_Evolutionary_B.html?id=E3Q8DwAAQBAJ&redir_esc=y (accessed June 23, 2020).

[56] T. Dobzhansky, F.J. Ayala, Temporal frequency changes of enzyme and chromosomal polymorphisms in natural populations of Drosophila., Proc. Natl. Acad. Sci. U. S. A. 70 (1973) 680–683. https://doi.org/10.1073/pnas.70.3.680.

[57] A. Futschik, C. Schlötterer, The next generation of molecular markers from massively parallel sequencing of pooled DNA samples, Genetics. 186 (2010) 207–218. https://doi.org/10.1534/genetics.110.114397.

[58] C. Schlötterer, R. Tobler, R. Kofler, V. Nolte, Sequencing pools of individuals-mining genome-wide polymorphism data without big funding, Nat. Rev. Genet. 15 (2014) 749–763. https://doi.org/10.1038/nrg3803.

[59] G.A. Watterson, On the number of segregating sites in genetical models without recombination, Theor. Popul. Biol. (1975). https://doi.org/10.1016/0040-5809(75)90020-9.

[60] M. Nei, W.H. Li, Mathematical model for studying genetic variation in terms of restriction endonucleases, Proc. Natl. Acad. Sci. U. S. A. 76 (1979) 5269–5273. https://doi.org/10.1073/pnas.76.10.5269.

[61] F. Tajima, Statistical method for testing the neutral mutation hypothesis by DNA polymorphism, Genetics. 123 (1989) 585–595.

[62] S. Wright, The genetical structure of populations, Ann. Eugen. 15 (1951) 323–354. https://doi.org/10.1111/j.1469-1809.1949.tb02451.x.

[63] M. Nei, Analysis of gene diversity in subdivided populations, Proc. Natl. Acad. Sci. U. S. A. 70 (1973) 3321. https://doi.org/10.1073/pnas.70.12.3321.

[64] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, EMBnet.Journal. 17 (2011) 10. https://doi.org/10.14806/ej.17.1.200.

[65] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics. 25 (2009) 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

[66] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The Sequence Alignment/Map format and SAMtools, Bioinformatics. 25

(2009) 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

[67] H. Li, A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data, Bioinformatics. 27 (2011) 2987–2993. https://doi.org/10.1093/bioinformatics/btr509.

[68] D.C. Koboldt, K. Chen, T. Wylie, D.E. Larson, M.D. McLellan, E.R. Mardis, G.M. Weinstock, R.K. Wilson, L. Ding, VarScan: Variant detection in massively parallel sequencing of individual and pooled samples, Bioinformatics. 25 (2009) 2283–2285. https://doi.org/10.1093/bioinformatics/btp373.

[69] P. Cingolani, A. Platts, L.L. Wang, M. Coon, T. Nguyen, L. Wang, S.J. Land, X. Lu, D.M. Ruden, A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3, Fly (Austin). 6 (2012) 80–92. https://doi.org/10.4161/fly.19695.

[70] R. Kofler, P. Orozco-terWengel, N. de Maio, R.V. Pandey, V. Nolte, A. Futschik, C. Kosiol, C. Schlötterer, Popoolation: A toolbox for population genetic analysis of next generation sequencing data from pooled individuals, PLoS One. 6 (2011). https://doi.org/10.1371/journal.pone.0015925.

[71] R. Kofler, R.V. Pandey, C. Schlötterer, PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq), Bioinformatics. 27 (2011) 3435–3436. https://doi.org/10.1093/bioinformatics/btr589.

[72] C. Lasne, B. Van Heerwaarden, C.M. Sgrò, T. Connallon, Quantifying the relative contributions of the X chromosome, autosomes, and mitochondrial genome to local adaptation*, Evolution (N. Y). 73 (2019) 262–277. https://doi.org/10.1111/evo.13647.

[73] P.W. Hedrick, J.D. Parker, Evolutionary genetics and genetic variation of haplodiploids and X-linked genes, Annu. Rev. Ecol. Syst. 28 (1997) 55–83. https://doi.org/10.1146/annurev.ecolsys.28.1.55.

[74] P. Pamilo, Genic variation at sex-linked loci: Quantification of regular selection models, Hereditas. 91 (1979) 129–133. https://doi.org/10.1111/j.1601-5223.1979.tb01652.x.

[75] P.W. Hedrick, Sex: Differences in mutation, recombination, selection, gene flow, and

genetic drift, Evolution (N. Y). 61 (2007) 2750–2771. https://doi.org/10.1111/j.1558-5646.2007.00250.x.

[76] L.M. Berg, M. Lascoux, P. Pamilo, The infinite island model with sex-differentiated gene flow, Heredity (Edinb). 81 (1998) 63–68. https://doi.org/10.1046/j.1365-2540.1998.00358.x.

[77] R.A. Hoskins, C.D. Smith, J.W. Carlson, A.B. Carvalho, A. Halpern, J.S. Kaminker, C. Kennedy, C.J. Mungall, B.A. Sullivan, G.G. Sutton, J.C. Yasuhara, B.T. Wakimoto, E.W. Myers, S.E. Celniker, G.M. Rubin, G.H. Karpen, Heterochromatic sequences in a Drosophila whole-genome shotgun assembly., Genome Biol. (2002). https://doi.org/10.1186/gb-2002-3-12-research0085.

[78] M.C. Schatz, A.L. Delcher, S.L. Salzberg, Assembly of large genomes using second-generation sequencing, Genome Res. (2010). https://doi.org/10.1101/gr.101360.109.

[79] C.H. Chang, A.M. Larracuente, Heterochromatin-enriched assemblies reveal the sequence and organization of the drosophila melanogaster Y Chromosome, Genetics. (2019). https://doi.org/10.1534/genetics.118.301765.

[80] S.E. Fick, R.J. Hijmans, WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas, Int. J. Climatol. (2017). https://doi.org/10.1002/joc.5086.

[81] S. D. Turner, qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots, J. Open Source Softw. 3 (2018) 731. https://doi.org/10.1101/005165.

[82] H. Wickham, C. Winston, Create Elegant Data Visualisations Using the Grammar of Graphics, Packag. “Ggplot2.” (2019). https://doi.org/10.1093/bioinformatics/btr406.

[83] D.K. Fabian, J.B. Lack, V. Mathur, C. Schlötterer, P.S. Schmidt, J.E. Pool, T. Flatt, Spatially varying selection shapes life history clines among populations of Drosophila melanogaster from sub-Saharan Africa, J. Evol. Biol. 28 (2015) 826–840. https://doi.org/10.1111/jeb.12607.

[84] M. Kapun, D.K. Fabian, J. Goudet, T. Flatt, Genomic Evidence for Adaptive Inversion Clines in Drosophila melanogaster, Mol. Biol. Evol. 33 (2016) 1317–1336. https://doi.org/10.1093/molbev/msw016.

[85] E.L. Behrman, T.J. Kawecki, P. Schmidt, Rapid evolution of learning and reproduction in natural populations of Drosophila melanogaster, BioRxiv. (2020) 288696. https://doi.org/10.1101/288696.

[86] B.S. Önder, M. Kapun, C.F. Aksoy, Ş.S. Şenkal, E. Demir, S. Coşkun, The effect of chromosomal inversions on seasonal phenotypic variation in Drosophila melanogaster, 2019.

[87] T. Ohtsu, M.T. Kimura, S.H. Hori, Energy storage during reproductive diapause in the Drosophila melanogaster species group, J. Comp. Physiol. B. (1992). https://doi.org/10.1007/BF00357524.

[88] T. Dobzhansky, Genetics of natural populations. Xvi. Altitudinal and seasonal changes produced by natural selection in certain populations of Drosophila Pseudoobscura and Drosophila Persimilis, Genetics. 33 (1948) 158–15876.

[89] T. Dobzhansky, Genetics of Natural Populations IX. Temporal Changes in the Composition of Populations of Drosophila Pseudoobscura., Genetics. 28 (1943) 162–86. http://www.ncbi.nlm.nih.gov/pubmed/17247077 (accessed July 6, 2020).

[90] A.G. Clark, M.B. Eisen, D.R. Smith, C.M. Bergman, B. Oliver, T.A. Markow, T.C. Kaufman, M. Kellis, W. Gelbart, V.N. Iyer, D.A. Pollard, T.B. Sackton, A.M. Larracuente, N.D. Singh, J.P. Abad, D.N. Abt, B. Adryan, M. Aguade, H. Akashi, W.W. Anderson, C.F. Aquadro, D.H. Ardell, R. Arguello, C.G. Artieri, D.A. Barbash, D. Barker, P. Barsanti, P. Batterham, S. Batzoglou, D. Begun, A. Bhutkar, E. Blanco, S.A. Bosak, R.K. Bradley, A.D. Brand, M.R. Brent, A.N. Brooks, R.H. Brown, R.K. Butlin, C. Caggese, B.R. Calvi, A. Bernardo De Carvalho, A. Caspi, S. Castrezana, S.E. Celniker, J.L. Chang, C. Chapple, S. Chatterji, A. Chinwalla, A. Civetta, S.W. Clifton, J.M. Comeron, J.C. Costello, J.A. Coyne, J. Daub, R.G. David, A.L. Delcher, K. Delehaunty, C.B. Do, H. Ebling, K. Edwards, T. Eickbush, J.D. Evans, A. Filipski, S. Findeiß, E. Freyhult, L. Fulton, R. Fulton, A.C.L. Garcia, A. Gardiner, D.A. Garfield, B.E. Garvin, G. Gibson, D. Gilbert, S. Gnerre, J. Godfrey, R. Good, V. Gotea, B. Gravely, A.J. Greenberg, S. Griffiths-Jones, S. Gross, R. Guigo, E.A. Gustafson, W. Haerty, M.W. Hahn, D.L. Halligan, A.L. Halpern, G.M. Halter, M. V. Han, A. Heger, L.D. Hillier, A.S. Hinrichs, I. Holmes, R.A. Hoskins, M.J. Hubisz, D. Hultmark, M.A. Huntley, D.B. Jaffe, S. Jagadeeshan, W.R. Jeck, J. Johnson, C.D. Jones, W.C. Jordan,

G.H. Karpen, E. Kataoka, P.D. Keightley, P. Kheradpour, E.F. Kirkness, L.B. Koerich, K. Kristiansen, D. Kudrna, R.J. Kulathinal, S. Kumar, R. Kwok, E. Lander, C.H. Langley, R. Lapoint, B.P. Lazzaro, S.J. Lee, L. Levesque, R. Li, C.F. Lin, M.F. Lin, K. Lindblad-Toh, A. Llopart, M. Long, L. Low, E. Lozovsky, J. Lu, M. Luo, C.A. Machado, W. Makalowski, M. Marzo, M. Matsuda, L. Matzkin, B. McAllister, C.S. McBride, B. McKernan, K. McKernan, M. Mendez-Lago, P. Minx, M.U. Mollenhauer, K. Montooth, S.M. Mount, X. Mu, E. Myers, B. Negre, S. Newfeld, R. Nielsen, M.A.F. Noor, P. O'Grady, L. Pachter, M. Papaceit, M.J. Parisi, M. Parisi, L. Parts, J.S. Pedersen, G. Pesole, A.M. Phillippy, C.P. Ponting, M. Pop, D. Porcelli, J.R. Powell, S. Prohaska, K. Pruitt, M. Puig, H. Quesneville, K. Ravi Ram, D. Rand, M.D. Rasmussen, L.K. Reed, R. Reenan, A. Reily, K.A. Remington, T.T. Rieger, M.G. Ritchie, C. Robin, Y.H. Rogers, C. Rohde, J. Rozas, M.J. Rubenfield, A. Ruiz, S. Russo, S.L. Salzberg, A. Sanchez-Gracia, D.J. Saranga, H. Sato, S.W. Schaeffer, M.C. Schatz, T. Schlenke, R. Schwartz, C. Segarra, R.S. Singh, L. Sirot, M. Sirota, N.B. Sisneros, C.D. Smith, T.F. Smith, J. Spieth, D.E. Stage, A. Stark, W. Stephan, R.L. Strausberg, S. Strempel, D. Sturgill, G. Sutton, G.G. Sutton, W. Tao, S. Teichmann, Y.N. Tobari, Y. Tomimura, J.M. Tsolas, V.L.S. Valente, E. Venter, J. Craig Venter, S. Vicario, F.G. Vieira, A.J. Vilella, A. Villasante, B. Walenz, J. Wang, M. Wasserman, T. Watts, D. Wilson, R.K. Wilson, R.A. Wing, M.F. Wolfner, A. Wong, G. Ka-Shu Wong, C.I. Wu, G. Wu, D. Yamamoto, H.P. Yang, S.P. Yang, J.A. Yorke, K. Yoshida, E. Zdobnov, P. Zhang, Y. Zhang, A. V. Zimin, J. Baldwin, A. Abdouelleil, J. Abdulkadir, A. Abebe, B. Abera, J. Abreu, S. Christophe Acer, L. Aftuck, A. Alexander, P. An, E. Anderson, S. Anderson, H. Arachi, M. Azer, P. Bachantsang, A. Barry, T. Bayul, A. Berlin, D. Bessette, T. Bloom, J. Blye, L. Boguslavskiy, C. Bonnet, B. Boukhgalter, I. Bourzgui, A. Brown, P. Cahill, S. Channer, Y. Cheshatsang, L. Chuda, M. Citroen, A. Collymore, P. Cooke, M. Costello, K. D'Aco, R. Daza, G. De Haan, S. Degray, C. Demaso, N. Dhargay, K. Dooley, E. Dooley, M. Doricent, P. Dorje, K. Dorjee, A. Dupes, R. Elong, J. Falk, A. Farina, S. Faro, D. Ferguson, S. Fisher, C.D. Foley, A. Franke, D. Friedrich, L. Gadbois, G. Gearin, C.R. Gearin, G. Giannoukos, T. Goode, J. Graham, E. Grandbois, S. Grewal, K. Gyaltsen, N. Hafez, B. Hagos, J. Hall, C. Henson, A. Hollinger, T. Honan, M.D. Huard, L. Hughes, B. Hurhula, M. Erii Husby, A. Kamat, B. Kanga, S. Kashin, D. Khazanovich, P. Kisner, K. Lance, M. Lara, W. Lee, N. Lennon, F. Letendre, R. Levine, A. Lipovsky, X. Liu, J. Liu, S. Liu, T. Lokyitsang, Y. Lokyitsang, R. Lubonja, A. Lui, P. MacDonald, V. Magnisalis, K. Maru, C. Matthews, W. McCusker, S. McDonough,

T. Mehta, J. Meldrim, L. Meneus, O. Mihai, A. Mihalev, T. Mihova, R. Mittelman, V. Mlenga, A. Montmayeur, L. Mulrain, A. Navidi, J. Naylor, T. Negash, T. Nguyen, N. Nguyen, R. Nicol, C. Norbu, N. Norbu, N. Novod, B. O'Neill, S. Osman, E. Markiewicz, O.L. Oyono, C. Patti, P. Phunkhang, F. Pierre, M. Priest, S. Raghuraman, F. Rege, R. Reyes, C. Rise, P. Rogov, K. Ross, E. Ryan, S. Settipalli, T. Shea, N. Sherpa, L. Shi, D. Shih, T. Sparrow, J. Spaulding, J. Stalker, N. Stange-Thomann, S. Stavropoulos, C. Stone, C. Strader, S. Tesfaye, T. Thomson, Y. Thoulutsang, D. Thoulutsang, K. Topham, I. Topping, T. Tsamla, H. Vassiliev, A. Vo, T. Wangchuk, T. Wangdi, M. Weiand, J. Wilkinson, A. Wilson, S. Yadav, G. Young, Q. Yu, L. Zembek, D. Zhong, A. Zimmer, Z. Zwirko, P. Alvarez, W. Brockman, J. Butler, C. Chin, M. Grabherr, M. Kleber, E. Mauceli, I. MacCallum, Evolution of genes and genomes on the Drosophila phylogeny, Nature. 450 (2007) 203–218. https://doi.org/10.1038/nature06341.

[91] B.Ş. Önder, S.N. Coşkun, C.F. Aksoy, S.Ş. Yağlı, E. Demir, M. Kapun, Rapid adaptation of stress related traits in Drosophila melanogaster to seasonal changing environment, ESEB, Turku, Finland, 2019.

[92] R. Cogni, C. Kuczynski, S. Koury, E. Lavington, E.L. Behrman, K.R. O'Brien, P.S. Schmidt, W.F. Eanes, The intensity of selection acting on the couch potato gene-spatial-temporal variation in a diapause cline, Evolution (N. Y). 68 (2014) 538–548. https://doi.org/10.1111/evo.12291.

[93] M.F. Rodrigues, M.D. Vibranovski, R. Cogni, Natural selection and parallel clinal and seasonal changes in Drosophila melanogaster, BioRxiv. (2020) 2020.03.19.999011. https://doi.org/10.1101/2020.03.19.999011.