Pattern of genetic variation, fine-scale genetic structure, and footprints of natural selection in populations of *Juglans regia* L. in the southern Kyrgyz Republic

Dissertation

Submitted in partial fulfillment of the requirements for the degree of
Doctor rerum naturalium (Dr.rer.nat)

at the Department of Forest Genetics and Forest Tree Breeding
Faculty of Forest Sciences and Forest Ecology
Georg-August-Universität Göttingen

By

Nurlan Torokeldiev
Born in Moscow, Russian Federation

Göttingen, 2016

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Torokeldiev, Nurlan:
Pattern of genetic variation, fine-scale genetic structure, and footprints of natural selection in populations of Juglans regia L. in the southern Kyrgyz Republic
ISBN 978-3-86376-180-6

Referee:	Prof. Dr. Reiner Finkeldey
Co-referee:	Prof. Dr. Ursula Kües

Date of oral examination: 26.02.16

Printed and published with the support of the German Academic Exchange Service.

1. Edition 2016, Göttingen

Text- & Coverlayout: Dipl.-Kfm. Alexander Mostafa
URL: www.optimus-verlag.de
Printed in Germany
Paper is FSC certified (wood-free, chlorine free and acid-free,
and resistant to aging ANSI 3948 and ISO 9706)

Dedicated to my beloved parents Muratali and Kimsanay,

my dear sisters Gulshair, Jumazia and Begimay,

my beloved wife Azada and daughter Mariam

Acknowledgments

This dissertation could not have come to end without the enormous support from many people during my PhD. I am using this opportunity to express my gratitude and my most genuine thanks to the following people.

Foremost, I would like to thank Prof. Dr. Reiner Finkeldey for accepting me as a PhD student to the Department of Forest Genetics and Forest Tree Breeding, for his continuous support, patience and guidance, despite his many other academic and professional commitments. His wisdom, knowledge and commitment to the highest standards inspired and motivated me.

Although it is not easy for me to acknowledge someone who has contributed so much to my academic studies in a few sentences, I would like to express my deepest gratitude to my former supervisor from Kyrgyz Republic, Prof. Dr. Almazbek Orozumbekov who passed away in 2014 at a young age. I would not have experienced such a great academic atmosphere without his encouragement and endless support. Dr.Almazbek was not only my advisor, but he was also an amazing mentor and an indispensable friend. His death is a great loss for our scientific community. May your soul rest in peace my dear friend!

It has been my privilege to work closely with Prof. Dr. Martin Ziehe during the last course of my doctoral studies. I am greatly indebted to his supervision and constant encouragement, which was vital in making this dissertation a reality. I am very grateful for his translation the summary into German. He spent hours to proofread my dissertation and to give me outstanding suggestions. Moreover, he showed me the approach handling population genetics data. I will never forget the hours we spent for brainstorming, since I was full of motivation and enthusiasm whenever I left his office. Thank you for all that you have done.

I must acknowledge Prof. Dr. Ursula Kües for her continuous guidance and willingness to be my co-referee and examiner. Special thanks also go to the rest of my Thesis Committee: Prof. Dr. Hermann Behling, Prof. Dr. Holger Kreft, and Prof. Dr. Christian Ammer, for their encouragement and special interest for taking part in my examination committee.

I would like to convey my heartfelt thanks to Dr. Elizabeth Gillet, who first introduced me to the statistical methods in the analysis of population genetics data course. She was very generous in sharing her scientific experiences and helping me with different software programs. I would like to extend my deepest thanks to Dr. Barbara Vornam for her help with sequence alignment analysis as well as Prof. Dr. Konstantin Krutovsky for his valuable encouragement during the initial course of my doctoral work. A special thanks to Prof. Dr. Hans-Rolf Gregorius. His extraordinary approach to studies in population genetics extremely motivated and inspired me.

I am also grateful to Dr. Yazar Minn, Dr. Markus Müller and Dr. Devrim Semizer-Cuming for their endless help and friendship through my studies. I would like to thank Alexandra Dolynska for her guidance in the lab. My gratitude is also extended to the faculty and staff of our institute: Prof. Dr. Hans H. Hattemer, Dr. Ludger Leinemann, Regina Berkeley, Christine Radler, and Marco Winkler.

I sincerely acknowledge my indebtedness to my class teacher Gülbara Muratova, and mathematics teacher and mentor Nishangül Karaeva at On-Eki-Bel Middle School, whose pedagogical talent positively affected me to love Mathematics and Science. I am also grateful to Abiybilla Abdullaev, who was the director at this school. To my high school teachers at Kyzyl-Kiya Sebat High Lycee. I will forever be thankful to these people for always being there and looking out for me and making me what I am today.

It is a pleasure for me to thank current and former PhD students: Dr. Oleksandra Kuchma, Dr. Yazar Minn, Dr. Markus Müller, Dr. Essy Harnelly, Dr. Sarah Seifert, Dr. Randy Villarin, Dr. Devrim Semizer-Cuming, Fitri Y. Amandita, Laura Cuervo, Sinan Hagenah, Johanna Ropertz and Natalie Breidenbach.

I take this as an opportunity to thank the German Academic Exchange Service (DAAD) for the financial support during my doctoral studies. I am deeply grateful to Ms. Alexandra Sandic-Magione, consultant-in-charge, for all her help.

Of course, no acknowledgements would be complete without giving thanks to my dearest parents, Muratali and Kimsanay. Dad, thanks for always believing in me and encouraging me to do the best despite hard times we had in the past. Mom, I am indebted to your personality for building up my character. The words are not enough to express my feelings how thankful I am for each of you and grateful for my smart genetic inheritance.

Azada, my dearest wife, I would like to express my heartfelt thanks for being the best true partner in my life. Thank you for your kind love and support. I am grateful to my baby girl Mariam Jeyda for showing me the beauty of life …

Table of contents

List of Tables

List of Figures

Abbreviations

AMOVA	Analysis of Molecular Variance
BLAST	Basic Local Alignment Search Tool
DNA	Deoxyribonucleic Acid
EST-SSR	Expressed Sequence Tags-derived SSR
FDR	False Discovery Rate
LD	Linkage Disequilibrium
MCMC	Markov Chain Monte Carlo
NJ	Neighbor-Joining
OLE	Oleosin
PCoA	Principal Coordinate Analysis
PCR	Polymerase Chain Reaction
SGS	Spatial Genetic Structure
SSR	Simple Sequence Repeat
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
WFFs	Walnut-Fruit Forests

1 INTRODUCTION

1.1 General introduction

Trees are the largest plant species, major elements of forest ecosystems and basic ecological models on Earth. In general, trees live a long time, preserve high levels of diversity and adapt to local conditions (Petit et al. 2006). Forest trees make up a major part of ecological systems and play important role in human society, as they provide various products, such as wood materials for building and fuel consumptions. Approximately 30% of the world`s ice-free land surface is covered by forests and woods (Council, N.R 1991). From an evolutionary viewpoint, the characteristics of forest trees in terms of maintenance of genetic variation and adaptation to environmental conditions should be an essential feature for future conservation efforts.

Understanding molecular basis and adaptation of species/populations is an important factor in population genetics (Gonzales-Martinez et al. 2006), because genetic diversity is a keystone of biological species. Diversity refers individuals, populations and species, and enables them to adapt to varying environments (Amos and Hardwood, 1998). Indeed, genetic variation within and among populations is fundamental for effective gene conservation issues (Eriksson et al. 1998). Besides, many tree species cover different geographical distribution ranges. Accordingly, local adaptations of particular populations may occur in response to variant selective forces under different environmental conditions, such as climate (Hedrick et al. 2006; Kawecki & Ebert 2004).

In general, evolutionary genetic forces, such as genetic drift, mutation, selection and gene flow shape levels of genetic diversity throughout long periods of time (Rajora et al. 2001). Perennial plant populations, for instance, generally maintain high levels of genetic variation than shorter-lived plant species, and are less likely to be altered by genetic drift (Hamrick et al. 1979). Life history and ecological factors could be a plausible explanation for the presence of high levels of genetic variation in forest tree populations (Hamrick et al. 1981).

Unfortunately, most of the forests in the world are diminishing at unusual rate. Various anthropogenic pressures and current deforestations have altered trees globally and led to extinction of more than 10% of the world´s tree species (Oldfield et al. 1998; Petit et al. 2006). As the status of forest trees is shifting globally, the genetic diversity within forest species also has gained interest and concepts in forest genetics are getting much more essential for genetic management (Ziehe et al. 1999). Because of changes in genetic structures of forest species, it is required to sustain permanent efforts (Finkeldey & Ziehe, 2004). Therefore, maintaining forest genetic resources gained much attention recently (Geburek and Turok, 2005) and is becoming one of the important task for today´s world.

A variety of DNA-based molecular markers have been developed for assessment of genetic diversity in forest trees. Molecular markers, such as Restriction Fragment Length Polymorphisms (RFLPs), Random Amplified Polymorphic DNAs (RAPDs), Amplified Fragment Length Polymorphisms (AFLPs) and microsatellites or Simple Sequence Repeats (SSRs), are often used for genetic studies of plant sciences (Mondini et al. 2009). SSRs, among these markers, are markers of choice in plant genetic studies because of their multiallelic nature, codominant inheritance, relative abundance and easy detection by using Polymerase Chain Reaction (PCR) (Thiel et al. 2003). As an alternative approach, expressed sequence tags (ESTs) have received particular attention, and such efforts may provide a different source of SSRs that combined with coding regions of the genome (EST-SSRs) (Eujayl et al. 2002). In the present study, the patterns of genetic diversity, differentiation and analysis of spatial genetic structure within and among natural walnut populations from walnut-fruit forests in the Kyrgyz Republic is assessed by using gSSR and EST-SSR markers.

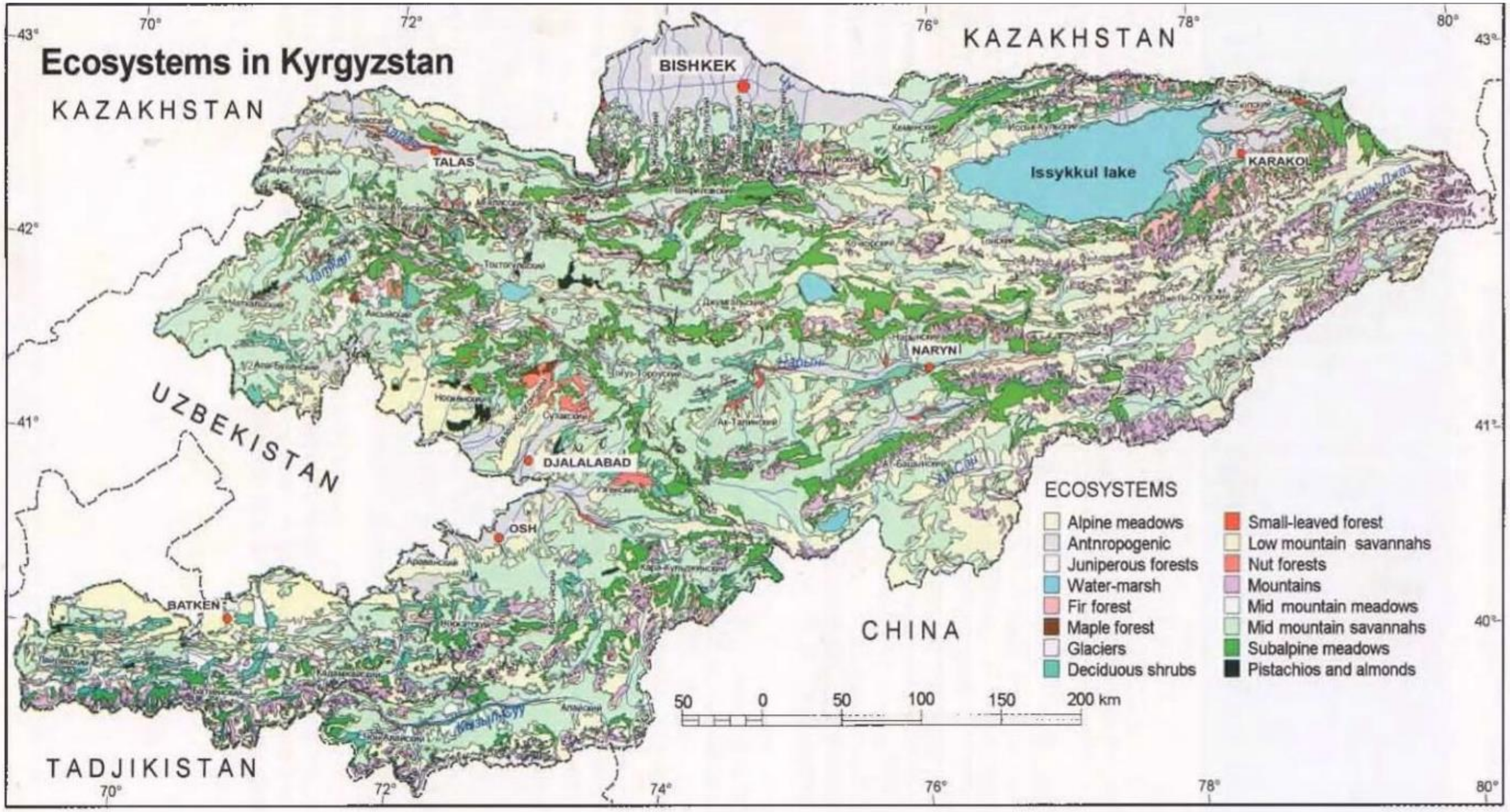
1.2 Biodiversity in Kyrgyz Republic

Biodiversity is the variety of species, populations, ecosystems and genetic diversity within and among populations (Frankham et al. 2004). Biodiversity has a major role in ecosystem balance. Actually, the biological diversity is quickly damaging by human activities. Consequently, maintaining biodiversity for the survival of living organisms on the planet is urgently needed.

The Kyrgyz Republic has valuable sources of biological diversity that occur mainly in mountainous areas. The territory of the country is about 19.99 million ha (4.3% are forests, 4.4% - water surface, 54% - agricultural lands and 37.3% - other lands), almost 95% of the land is mountainous at more than 1500 m a.s.l and is located between two mountain systems (Tien Shan and Pamir) (Orozumbekov et al. 2009). The mountainous region of the Kyrgyz Republic covers varieties of ecosystems; however, only about 4 % covered with forested area. The ecosystems include deciduous and evergreen forests, shrublands, grasslands, desserts, various wetlands (Akimaliev et al. 2013).

The mountainous area, such as western Tien-Shan and southern Fergana region, is an original place for many cultivated plant species (Dzunusova et al. 2008). There are four types of forests: spruce forests (*Picea schrenkiana*, as a main species), walnut-fruit forests (*Juglans regia, Malus* and *Prunus* sp.), and juniper (Archa) forests (*Juniperus* sp.). In the north of the country, forests mostly consist of spruce and poplar; however, forests in the southern part, with drier climate, occur in mixed types with walnut, maple, apple, cherry, plum, *Crataegus* and almond (Orozumbekov et al. 2009) (Fig. 1).

The wild fruit and nut forests, particularly walnut-fruit forests in southern part of the country, are stated to be the largest forest area in the present time, and gained special interest globally in terms of genetic resources (Fisher et al. 2004). According to the Vavilov (1931, 1951), center and origin of the Persian walnut (*Juglans regia*) is likely to be in the mountains of Central Asia, but exact history of the walnut forests has been subject of debate. Unfortunately, all levels of ecosystem are under anthropogenic pressures, including natural walnut forests in walnut-fruit forests. For instance, during periods between 1897 and 1995, the total area of nut-fruit forests (Persian walnut, pistachio, etc) decreased by 50% from 1,200,000 to 630,000 ha, including approximately 20,000 ha of walnut forest trees (Ashimov, 1995 & 1998; Molnar et al. 2011).



Note: Borders shown on the map are not definitive.
Source: Emil Shukurov, 2005.

Figure 1 Ecosystems in Kyrgyz Republic (*Source*: from ECODIT, 2013; based on Shukurov, 2005 and United Nations Development [UNDP])

1.3 The Persian Walnut (*Juglans regia*)

1.3.1 Taxonomy and botanical aspects

The family *Juglandaceae*, known as walnut family, consists of seven genera and about 60 deciduous tree species located mostly in the Northern Hemisphere of both Old and New World (Manos et al. 2001). Walnut trees belong to the genus *Juglans L.*, usually the seeds of these trees are known as walnuts. Besides genus *Juglans L.*, *Juglandaceae* family consists of six more genera that are commercially essential nut-producing trees, namely *Carya Nutt.*, *Oremunnea Oerst.*, *Platycarya Sieb* and *Zucc, Pterocarya* Kunth., *Engelhardia* Lesche. *Ex* Blume and *Alfaroa* Standl. (Manning et al. 1978). About 21 species are present in the genus *Juglans* (Miller et al 1976; Manning et al. 1978), which are divided into four sections mostly based on morphology and wood anatomy; *Cardiocaryon* Dode, *Trachycaryon* Dode *ex* Mann., *Rhysocaryon* Dode and *Juglans*. The Persian walnut (*Juglans regia,* literally "royal walnut") is well-known in this genus. In general, *Juglans regia* owns different names, such as Carpathian, Himalayan and English walnut. In Kyrgyz Republic, for instance, *J*.regia is called "Грецкий орех (Greece nuts)".

One of the most extensive walnut forests are in Kyrgyz Republic, which are important genetic resource for *in situ* conservation efforts (Hemery et al. 1998). *J. regia* is a large and deciduous tree growing up to 35 m in the wild (Fig. 2) with tomentose juvenile shoots. It is a light-demanding species, sensitive to winter and late spring frosts (Mohni et al. 2009). The root systems of walnut trees are well-developed and deep that provide significant stress and drought tolerance (Molnar et al. 2011). The leaves are alternate, 25-40 cm in long, odd-pinnate with 5-9 leaflets (Fig. 2). The male flowers are in catkins of 5-10 cm in length, and the female flowers are terminal (Fig. 3). The fruit is drupe-like and spherical, with a green pericarp, which releases the nut when mature. All *Juglans* species produce edible nuts, although size and extrability differ considerably. In fact, *Juglans regia* has impressive diversity in nut types (shape, size, thickness, kernel, color, flavor, etc), and nuts are mainly sold in markets, known as bazaars (Fig. 4; Fig. 5).