Doctoral Thesis

博士論文

Morphological and genetic characterization

of mulberry (genus *Morus*)

（クワ（Morus 属）の形態学的および遺伝的特性評価）

Muhonja Linah

ムホンジャ　リナ

Contents

List of Tables

List of Figures

Abstract

Morphological and genetic characterization of mulberry (genus *Morus*)

Laboratory of Applied Bioresources

Student ID: 47-177311

Muhonja Linah

Introduction

Mulberry belongs to the genus *Morus* and the family of Moraceae with over 68 species recognized (Datta, 2002). After the first classification of mulberry that unveiled seven species (Linnaeus, 1753), several studies on mulberry classifications have proposed different number of mulberry species such as 24 species and one subspecies (Koidzumi, 1917), 35 species (Hotta, 1958), 16 species (Zhou and Gilbert, 2003) and the need for correct reclassification still continues for breeding and utilizing mulberry all over the world. This diversity and classification have elucidated a long debate over time due to relatedness in morphological features as a result of the heterogeneous nature and influence of environment (Machii et al., 2001). Contribution through enhanced morphological and genetic studies have been done in previous studies (Sharma et al., 2000;Tikader et al., 2001;Bhattacharya and Ranade, 2001; Awasthi et al., 2004;Kafkas et al., 2008; Vijayan et al., 2008;Chikkaswamy et al., 2012), however, it remains a challenge as they focused on either morphological or genetic markers with minimum number of species and varieties as well as different technologies. This further prompted me to contribute towards knowledge base of mulberry since understanding the relationship in mulberry varieties among the *Morus* species is crucial to enhance the benefits of breeding, conservation and utilization.

My research objective is to contribute to addressing the above mentioned issues by

revealing phenotypic and phylogenetic relationship existing among 56 diploid mulberry varieties belonging to seven different *Morus* species (*M. alba, M. indica, M. bombycis, M. acidosa, M. latifolia, M. kagayamae*, and *M. rotundiloba*) and one unspecified *Morus* species ('Enbu') by combining both morphological and genetic markers, as well as unravelling the whole genome sequence of 'Enbu'.

Results and discussions

1. Morphological characterization of 56 mulberry varieties

Morphological characterization has been indispensable in plant taxonomy over time; however, the nature of plant phenotypes and diversity make it challenging to distinguish between closely related varieties among species of the same family. I used the 56 diploid varieties selected at NARO Genebank (Tsukuba) with aim of evaluating their morphological features which may be used for classifying them into species category. The varieties consist of *Morus alba* (8 varieties), *M. acidosa* (6), *M. kagayamae* (4), *M. bombycis* (13), *M. indica* (4), *M. rotundiloba* (4), *M. latifolia* (13), interspecific hybrids (3) and *Morus* sp. 'Enbu' (hereafter, 'Enbu'). In order to evaluate morphological characteristics of 56 varieties, I measured the quantitative and qualitative morphological features of individual varieties using three replicates; the quantitative ones are blade length, blade width, apex length and petiole length of mulberry leaves (**Fig. 1(a)**) and the qualitative ones are appearance of leaf blade, shoot structure and orientation.

The result revealed diversity among mulberry at variety level and within the same species exist in the morphological features (**Fig. 1(b) - (i)**). In the blade length and blade width, clear differences were observed between species with higher values observed in 'Enbu', *M. latifolia*, *M. kagayamae* and interspecific hybrids ('Rohachi', 'Shinichinose' and 'Kairyou Ichinose') while smaller values in species *M. acidosa, M. alba*, and *M. indica*, respectively. However,

similar values were observed among some species, making it difficult to isolate the species clearly. The apex length differs among the species separating three groups of long apex (*M. kagayamae*, *M. acidosa*, and *M. rotundiloba*), medium apex group ('Enbu', *M. indica*, *M. bombycis* and interspecific hybrids) and finally the small group category (*M. latifolia* and *M. alba*). The petiole length also grouped the species into long group (*M. latifolia*, *M. kagayamae*, the interspecific hybrids, and *M. alba*), medium group (*M. bombycis*, *M. indica* and 'Enbu') and the short group (*M. rotundiloba* and *M. acidosa*) respectively. Overall, similar values were observed among some species in each feature, making it difficult to isolate the species clearly. However, the leaf apex maybe useful for differentiating some mulberry varieties compared with other features. Among the mulberry varieties, 'Enbu' seemed to be having big leaf blade, medium apex and short petiole compared to other species phenotypes (**Fig. 2**). The results prompted necessity for further analysis using SNP markers in order to bring out the phylogenetic relationships among them through genetic markers.

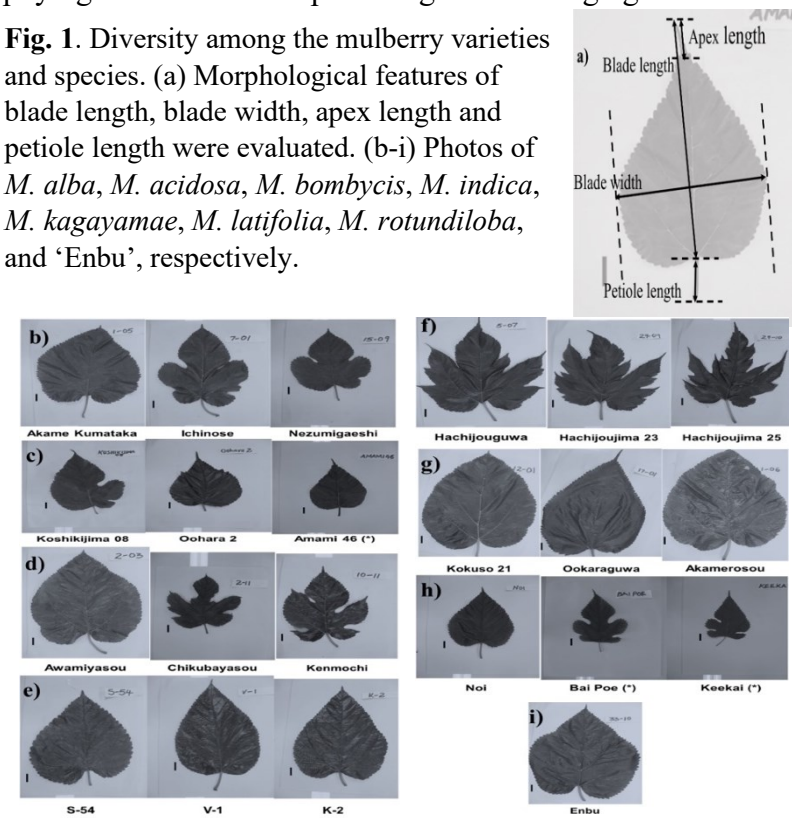

Fig. 1. Diversity among the mulberry varieties and species. (a) Morphological features of blade length, blade width, apex length and petiole length were evaluated. (b-i) Photos of *M. alba*, *M. acidosa*, *M. bombycis*, *M. indica*, *M. kagayamae*, *M. latifolia*, *M. rotundiloba*, and 'Enbu', respectively.

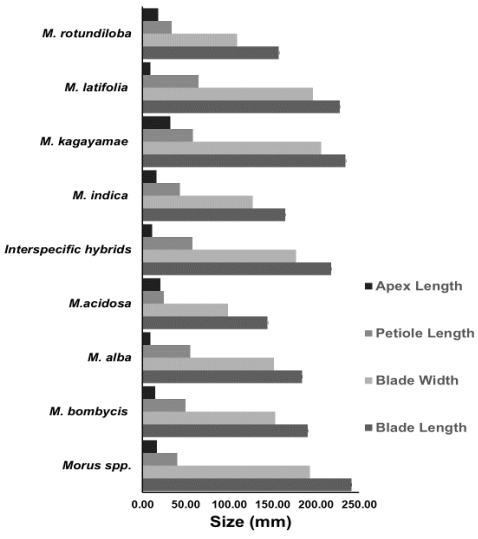

Fig. 2. Variation among the mulberry species in blade length, blade width, petiole length and apex length.

2. Genome-wide SNP marker discovery and phylogenetic analysis of mulberry varieties using double-digest restriction site-associated DNA sequencing

Due to difficulties observed in differentiating the mulberry species by morphological phenotypes, further investigations focused on genetic interrelationship among the varieties was conducted using double digest restriction site associated DNA sequencing (ddRAD-seq). Genome-wide 2,229 homozygous SNPs were identified among the 54 mulberry varieties in the eight species by ddRAD-seq. The results of the phylogenetic analysis revealed existence of three clear monophyletic clades in two Japanese native species, *M. acidosa* (**C**) and *M. kagayamae (***K***)*, distributed in different geographically isolated islands, and a Thai native species, *M. rotundiloba* (K*)*, whereas the other species were non-monophyletic (**Fig. 3)**. Varieties of *M. bombycis* (**B***)*, another Japanese native species, were roughly classified into three groups (B1, B2 and unclustered varieties). Of these, two groups were monophyletic with *M. acidosa* (**C**) and *M. kagayamae* (**K**), forming **BC** and **BK** clades respectively, while another group was not monophyletic. Furthermore, varieties of *M. indica* (**I**), an Indian native species,

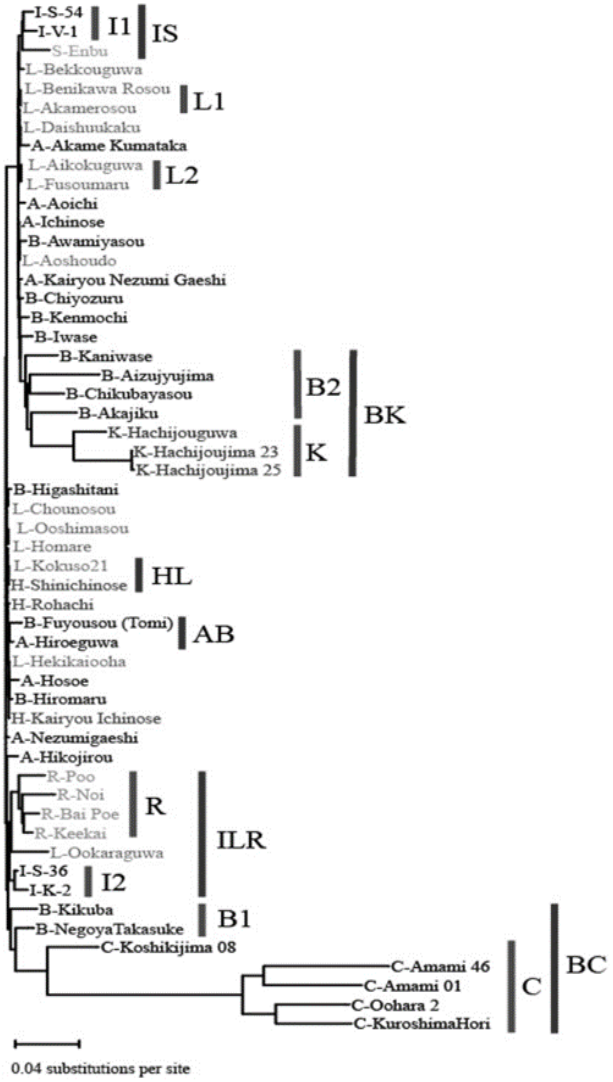

Fig. 3. Neighbour joining (NJ) phylogenetic tree of 54 mulberry varieties where **C**: *M. acidosa*; **I**: *M. indica*; **R**: *M. rotundiloba*; **K**: *M. kagayamae*; **B**: *M. bombycis*; **L**: *M. latifolia*; **A**: *M. alba*; **S**: 'Enbu'.

were classified into two different monophyletic clades. Of these, one clade was clearly monophyletic with an indigenous variety in Kenya, 'Enbu', while another clade was monophyletic with *M. rotundiloba* (**R**) varieties and one *M. latifolia* (**L**) variety. There were no clear monophyletic clades within *M. alba* (**A**)and *M. latifolia* (**L**)varieties, which could be a result of several hybridization events after their introductions from China to Japan. These results suggested that it was difficult to clearly classify the hybridized mulberry varieties even with genome-wide DNA markers. Having observed that the leaf apex could contribute towards differentiating the varieties, I evaluated relationship between the structure of the phylogenetic and morphological characteristics of mulberry leaves. The results revealed that leaf tip ratio (LTR), calculated by apex / blade length, may correlate to genetic differences among the two *M. bombycis* groups (**B1** and **B2**) in monophyletic clades and another unclustered *M. bombycis* group in non-monophyletic clades, suggesting that LTR might be used for evaluating hybridization level among *M. bombycis* varieties. Overall, these results may provide new insights into taxonomic debate of mulberry species.

3. Admixture analysis of 56 mulberry varieties

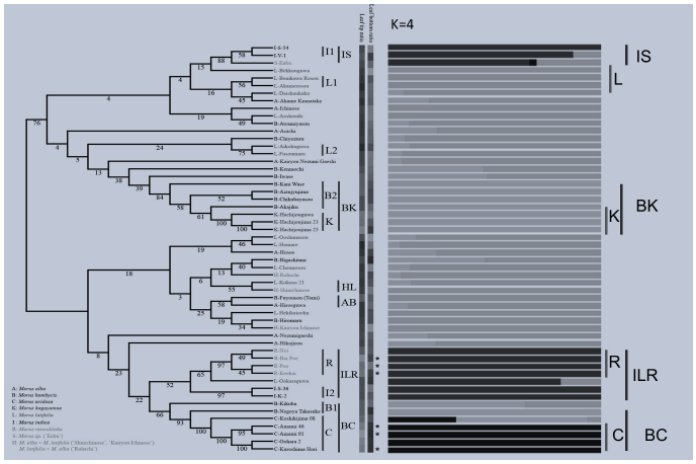

Fig. 4 Reconstructed phylogenetic tree by Neighbour joining (NJ) method combined with admixture analysis derived from 54 mulberry varieties belonging to I: *M. indica*;S: 'Enbu'; L: *M. latifolia*;K: *M. kagayamae*; B: *M. bombycis;* C: *M. acidosa*; R: *M. rotundiloba;*A: *M. alba.;*.

Interesting upon conducting the admixture analysis on varieties from the documented 7 known species plus one unknown species, there is indication of varietal interactions among the species resulting to creation of three clear clades of native species of *M. kagayamae* (**K**), *M. acidosa* (**C**), and introduced *M. rotundiloba* (**R**) as had been observed in phylogenetic analysis. However, the results further reveal presence of

admixtures that may have occurred among the varieties in their growing niches resulting to similarities among the phenotypes after introductions of *M. latifolia* (**L**), *M. indica* (**I**)and *M. alba* (**A**). *M. latifolia* (**L**)seems to be most affected by the hybridization among the species.

4. Genome sequencing of Enbu variety of mulberry with comparison to Chinese mulberry Chuansang (*Morus notabilis*)

After the admixture analysis result revealing interactions among the varieties, I isolated whole genome sequence of 'Enbu' an indigenous variety linked to Kenya in order to give insight into the genome of 'Enbu' in comparison to existing genome sequences of *Morus notabilis* in MorusDB with target of finding out the similarities and differences among the two species. Highly improved genome assembly of 'Enbu' compared with the existing genome assembly of *M. notabilis* was constructed by de novo assembly using PacBio long reads with scaffolding done by Dovetail's Chicago method. The results revealed that the size of 'Enbu' genome assembly is 349.92 Mb, (N50 is 1,17 Mbp). In the genome assembly, 35,483 predicted gene sequences were identified in 32,989 loci. Functional annotation for the predicted genes using several public databases such as NCBI-nr, InterPro, Gene ontology, and KEGG were performed.

Conclusion

Morphological features alone despite being indispensable and accessible cannot differentiate the varieties among the species clearly. My result of combining both morphological features and genetic markers may offer a promising future in differencing varieties, however, in highly hybridized varieties there is need to go deeper into single copy nucleotides markers and finger typing for proper placement of varieties among the species. The use of SNP markers and some morphological features such as LTR could help solve the dilemma in classification of mulberry species. Furthermore, exploration of the genome

sequence and comparison to public databases offers an opportunity to proper utilization of the mulberry bioresource by other researchers and farmers in breeding varieties that meet their needs.

Chapter1: Introduction and literature review

1.1 Mulberry and silk

Mulberry and silk are intertwined for productivity and utilization of the other yet the great irony lies on how silk is hyped as the queen of all textiles (Khurana and Checker, 2011; Nazim et al., 2017; Sori and Gebreselassie, 2016) without considering the whole value chain which in fact has created a pending gap slowing down the industry. The whole silk value chain for development encompasses both moriculture (Kole, 2011) that is the production of mulberry for silkworm (*Bombyx mori* L.) feeding(Khurana and Checker, 2011) and sericulture which is the magnificent art of rearing silk worm and production of silk (Rahmathulla, 2012; Rouge, 2010; Vijayan et al., 2012) yet much focus has targeted silk production and silkworm rearing over time while neglecting the source of the feeds nourishing the system. Interestingly, research has shown that the success of the value chain solely depends on production and sustainability of mulberry where the mulberry quality and quantity produced affects the silk productivity through the process of proper silkworm rearing (Khurana and Checker, 2011; Kumar et al., 2014; Mathithumilan et al., 2016; Sarkar et al., 2017).

Over time, silk production has been a preserve of a few countries, especially of the Asian continent origin such as India, China, Thailand, Korea and Japan with a slow uptake trend by other countries that tend to currently tap into the available diverse genetic resource of the Moraceae family (Venkatesh and Chauhan, 2008; Vijayan, 2010). The growth and diversity of mulberry in Asian continent could have been influenced by the fact that the mulberry root origin is thought to be at the Himalayan foothills before spreading to other areas of Asia, Africa, America and Europe (Rao et al., 2013; Sanchez, 2002). The mulberry distribution that followed throughout the world hence forth finds its route aligned to trading along the great silk road trail coupled with the ease of high adaptation of varieties to diverse microclimatic conditions ranging from tropics, subtropics to temperate (Machii et al., 2000; Sanchez, 2002).

Despite the availability of mulberry as a bioresource, the growth of the moriculture and /or sericulture industry as a whole has been dwindling with indicators showing inadequate sustainable mechanisms in production caused by reduced land area for production of mulberry, high costs of labour and industrialization coupled by competition from cheap fabrics demotivating production (Datta and Nanavaty, 2005; Mathithumilan et al., 2016). Furthermore, less investment in physiological adaptation studies conducted on various species to ascertain the changes that may have occurred over time affecting the genetic makeup has been inadequate (Liu et al., 2019).

1.2 Production and distribution mulberry

Mulberry (*Morus* sp.), belonging to Rosales the order (APG IV, 2016), in the family of Moraceae is a perennial plant. According to the Rosale classification, it occurs as a monophyletic descend (Zhang et al., 2011). Documentation of over 150 known species exist all over the world although only 68 species of the mulberry are recognised for their role in sericulture (Datta, 2000) with their centre of origin believed to be from the foothills of Himalayas and traceable towards China and Japan. The silk industry that spans over several thousand years ago as a result of domesticated silkworm (*Bombyx mori*) rearing, could have been the drive for the mulberry domestication (He et al., 2013) and further production. Moreover, the regional classification of genus *Morus* based on the morphological characterization of style length and nature of stigma with over 100 cultivars known of deciduous small trees or shrubs growing wildly in many regions of the world offered a great opportunity such as 24 species and one subspecies (Koidzumi 1917). These plants were found to be versatile in nature where they could grow to be large and with them being deciduous, they could adapt Asia, Africa, and America where the native zones ranging from tropical, subtropical and temperate regions (Sanchez, 2002). Sharma et al., (2000) further showed that

several species of significant economic importance are widely distributed around the world. Mapping the species revealed that the Asian continent has the highest diversity of moraceae species with Africa and South America having each one species respectively (Sharma et al., 2000). It is of no wonder that the mulberry centre of origin is placed to be the Japan-China, where great diversity and technology has been embedded into sericulture running over fifty centuries. However, further exploration in Africa needs to be done to reconfirm the varieties that are currently available and whether they are native or introduction. Key species among the documented include the black mulberry (*M. nigra*), white mulberry (*Morus alba)* and American mulberry (*M. rubra*) among others that can be found in Asian continent, middle East, Africa and possible spread through introductions to many regions (**Fig. 1**).

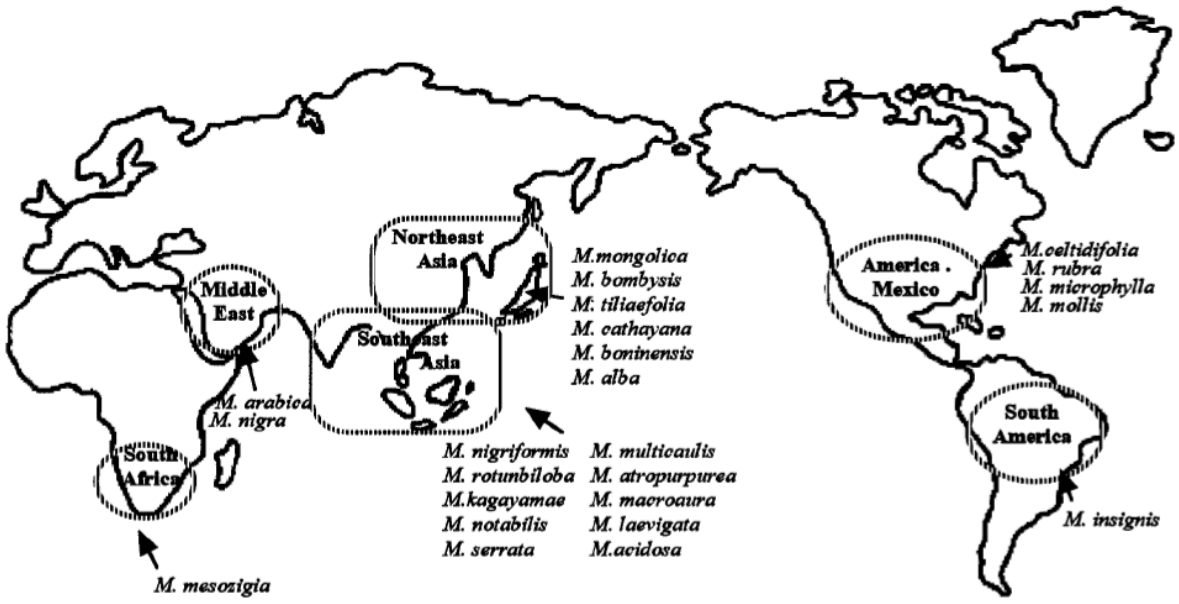

Figure 1. Distribution of various species of mulberry around the world.

(cited from Sharma et al., 2000)

Adaptation of mulberry to a wide range of climatic conditions ranging from tropical, subtropical and temperate (Machii et al.,2000; Tikader and Dandin,2005) is advantageous, although the success of production requires maximization of all factors of production such as

light, water, nutrients, good soils and proper management just like any other crop. The quality of the leaves targeted for feeding silkworm deserve key attention regarding the duration exposed to sunlight for proper photosynthesis, amount of nutrients provided, amount of rainfall / irrigation provided as well as optimum temperatures for sustainable and optimum yield production. To achieve the required quality of leaves for rearing healthy silkworm to maturity, the leaves should be tender, rich in moisture, proteins, carbohydrates and other essential minerals.

In Kenya, these conditions are available in different climatic zones existing with temperatures ranging between 24-28^{0}C, adequate rainfall of 800-2000 mm, water supply with good aeration and adequate sunlight for proper leaf development (Tuigong et al., 2015). The regions producing the mulberry have well drained soils with pH range of between 6.2-6.8 and usually well amended with farmyard manure during the planting time.

Despite the above being possible, growth of mulberry trees /shrubs is currently being affected by different abiotic and biotic constraints thus reducing its productivity. The pests and diseases, various aspects of climate change and management witnessed, poor soils and fertility management, varied altitude, erratic rainfall and extreme temperature have impacted on growth patterns as well as the silkworm rearing process resulting to low quality of raw silk and hence low prices which further do not motivate the farmers to grow the mulberry anymore affecting the whole value chain. Sustainable production and regeneration of mulberry is therefore attainable by use of both sexual and asexual methods of propagation that have been adapted world-wide and available (Machii et al., 2000; Yamanouchi et al., 2017). Besides, these valued mulberry plant has been found to have a complex breeding system of monoecious, androecious, gynodioecious and dioecious enhancing ease for cross pollination enabling ease of multiplication and manipulation in their respective niches (Datwyler and Weiblen, 2004; Tikader et al., 1995).

Mulberry culturing as a whole depends on the purpose of production; either for rearing silkworm, fruits production or landscape maintenance that dictate the management practices required for maintaining the plant. They can be maintained either as a short shrub through repeated pruning or as a tree throughout the growing period. More so with proper plant spacing, planned frequency of pruning, proper maintenance of canopy density and other management practices such as irrigation, weeding and fertilization per unit area, the average leaf production of mulberry can be boosted throughout the growth cycle taking about 2-3 years from seeding to flowering when undisturbed (Fukui, 2009; Nithya et al., 2011; Rao et al., 2013).

Genetically, mulberry crop varieties exists as diploids having been derived from the wild although sometimes polyploids occur as a result of either crossing and further human manipulation by use of colchicine (Das et al., 1970;Mathithumilan and Dandin, 2009; Ramesh et al.,2011; Dai et al.,2015) with target of increasing leaf yield, improving feed efficiency as well as the mitigating the glucose menace in lifestyle diseases related research among others requirements (Lown et al., 2017).

1.3 Importance of mulberry and economic impact to the society

Mulberry crop has received recognition world over due to its economic importance, medicinal value and ecological function unlocking the need to understand the crop more (Zeng et al., 2015). For over 50 century years, the crop has played an isolated economic role of providing feed to the silkworm for production of silk required to drive the sericulture industry (Wang et al., 2017).

The sericulture industry involving silk production has progressed through the silk value chain, economically empowering different segments from production to consumer with mulberry standing out as silkworm (*Bombyx mori* L)feeds influencing both the amount and quality of the cocoons produced (Khurana and Checker, 2011; Sarkar et al., 2017; Sori and

Gebreselassie, 2016). Mulberry dietary preferred feed by silkworm is attributed to the leaves having potential to provide required nutrients for development and the optimal amounts of sterols needed as a structural component for the cellular membrane and precursor for ecdysteroid; a moulting hormone for silkworm (Nair et al., 2005; Niwa et al., 2010; Murthy et al., 2013) hence sustaining the whole system.

Although the crop has been used for feeding and silkworm rearing for a long time, very little diversification had been done until in the 90s when the value addition concept drove the diversification of the crop into many aspects affecting human livelihood and income generation through the different components. The leaves contribute towards the biggest share of the value chain through silkworm rearing where their protein content enables them to utilised as forage for livestock (Sánchez, 2000; Sanchez, 2002; Wang et al., 2012), poultry / chicken feed and aquatic fish feeds (Ustundag and Ozdogan, 2015; Al-Kirshi et al., 2010; Al-Kirshi et al., 2013; Machii and Katagiri, 1991; Ebrahim et al., 2015; Saddul et al., 2004; Sánchez, 2000; Sánchez, 2002), superfood for nutraceuticals (Krishna et al., 2018) and vegetables (Singh et al., 2013).

In addition, the plant extracts from the mulberry leaf, bark and roots have been used in dyeing fabrics (Nguyễn et al., 2017) and making beauty products (Pillaiyar et al., 2017).The pharmaceutical industry too has used products derived from mulberry to develop products that help mitigate human lifestyle and diseases. Such efforts have been experienced through recovery of 1-deoxynojirimycin (DNJ) that has been found to play part in increasing the plasma adiponectin and activating the beta oxidation system thus reducing diet based obesity known as antihyperglycemic (Venkatesh and Chauhan, 2008; Tsuduki et al., 2013; Gao et al., 2016; Thaipitakwong et al., 2018; Zhang et al., 2018; Lim and Choi, 2019). Other health benefits derived include the; antioxidative (Harauma et al., 2007; Aramwit et al., 2010; Thabti et al., 2011; Niratker et al., 2015), hypolipidemic (Chen and Li, 2007; Kobayashi et al.,2010;Yang et al., 2010; Zhang et al., 2014; Yuan and Zhao, 2017), antiatherogenic effects (Priya, 2012;

Gryn-Rynko et al., 2016; Rodrigues et al., 2019), type 2-Diabetes (Jiao et al.,2017), Parkinson disease (Kim et al.,2010), Pain reliever (De Souza et al.,2000).

Furthermore, fruits too have a major role to play for being a source for fresh fruit and different nutrients and the antioxidants (Imran et al., 2010) contributing towards the main food processing through production of marmalades and jam, soft drinks like juices and tea as well as liquors

The crop has also been tapped into with regard to protection against pests through production of biopesticides for control of termites among others (Wasano et al., 2009; Konno, 2011; Golpayegani et al., 2014; Gai et al., 2017).

Humans through innovations have utilised the hardy and light characteristics of mulberry plant parts to come up with more value-added products that include light and durable timber used to make ornaments, carvings and sports equipment's, human foods in terms of fresh leaf as vegetables, processed mulberry teas and wines (Tao et al., 2017), and poultices for wound healing (Gao et al., 2015).

The mulberry trees, wood and branches have contributed towards the paper making (Dong et al., 2017), production of wood natural dyes (Nguyễn et al., 2017) making sports equipment's such as hockey sticks and carvings as has been case of *M. kagayamae* (Dong et al., 2017) as well as contributing towards mitigating environmental degradation effects through landscaping and agroforestry due to its ability to grow on poor soils (Liu and Willison, 2013;.Jiang et al.,2017)

Overall there is the herbal medicine segment that more often never captured clearly in plant use stands out in mulberry where the Chinese have been utilising the crop with a lot of documentation to heal different ailments (Priya, 2012; Eo et al., 2014; Zhang et al., 2014; Zhang et al., 2014; Niratker et al., 2015; Chan et al., 2016; Krishna et al., 2018; Thaipitakwong

et al., 2018; Wei et al., 2018) Furthermore the profiling of different parts of *M. alba* species from different Chinese cultivars has revealed different antioxidants and their benefits (Yuan and Zhao, 2017). Noted the antioxidants have both the phenols and flavonoids known worldly to have health benefits towards antimutagenic and anticarcinogenic agents. Several studies have emphasized the importance of profiling the varieties for anthocyanins and polyphenols among others. Examples include Jin et al. (2015) that revealed the presence of 4 anthocyanins of importance as polyphenols and inhibitors of glucosidase while another study, Wang et al., (2019) emphasised on mulberry being a good source of phenols. Furthermore, studies have highlighted cyanidin, pelargonidin, delphinodin petunidin as key anthocyanin components (Kang et al., 2006; Huang et al., 2013; Jiang et al., 2013), phenolics (Kang et al.,2006; Yuan and Zhao, 2017), chlorogenic acid and protocatechuic acid as phenolic acids (Kähkönen and Heinonen, 2003; Wang et al.,2011; Tao et al.,2017; Yuan and Zhao,2017) and flavanols such as quercetin and kaempferol (Sugiyama et al., 2013; Tao et al., 2017) among others. Further studies have revealed other phenolic compounds such as maclurin, rutin, quercitrin, resveratrol, and morin among others (Chang et al.,2010; Chan et al., 2016; Thabti et al.,2014). The properties of above components have played a major role in free radical scavenging activity (Du et al., 2008), lowering blood glucose in the blood (Wang et al., 2013; Zhang et al., 2014), recovery and protection of nervous system (Kang et al., 2006; Kim et al., 2010), re-energizing the body systems (Jiang et al.,2013) reduction and treatment of high cholesterol effects (Chen et al., 2005; Jiao et al., 2017; Wang et al.,2011) among others. The abundant antioxidative components in mulberry leaves with ability to scavenge for reactive oxygen species (ROS) is a driver for more variety search and identification. In lieu of these clarity of different components extracted from mulberry equally has resulted to need for components such as glycosidase inhibitor 1-deoxynojirimycin (DNJ), that is potent on elevated blood glucose levels (Gao et al., 2016; Tsuduki et al., 2013) as well as antioxidative effect of mulberry leaf extract

such as oxidation of low density lipoprotein (Sugiyama et al., 2013), antioxidant (Bae and Suh,2007; Wang et al., 2013; Kasote et al., 2015) and products used to prevention of breakdown of melanin through ethanolic extract (Chang et al., 2010).

1.4 Mulberry production challenges and prospects

1.4.1 Climate change and its effects

Myriad of challenges have been aggravated by climatic changes affecting both the biotic and abiotic factors in nature consequently affecting plant growth and species diversity (Raza et al., 2019). Notwithstanding the above, mulberry has stood taste of time by maintaining its diversity through continued natural crossing in addition to assisted breeding (Vijayan et al., 2012) as such many varieties have been documented in different regions (Machii et al., 2000).

1.4.2 Research and development of mulberry

Mulberry research has been around since the 19th century and it's been driven by the countries majorly dealing in silk industry for silkworm rearing with major focus revolving around agronomic aspects for quality and quantity of crop. Furthermore, with increasing lifestyle diseases, research has been tailored towards utilizing mulberry to mediate the effects. Previous main research focus considered the harvesting modalities of the crop, how to preserve the forage and conduct animal trials as well as research into utilization. Notable, with the genus *Morus* having over 150 species (Awasthi et al., 2004) that can be utilised in research and development of the crop while tapping into the use of wild species existing may contribute a major crop improvement role through breeding.

Heterozygous nature of mulberry (Venkatesh and Munirajappa, 2015) has potential to allow ease outcross and result to new varieties. These generated varieties however, pose a

challenge of identification if only the morphological characteristics alone are used because the ease of crossing can modify the phenotypes of the plant hence more research and technologies are needed to determine the genotype nature of the species.

Proper identification of the species and lineages of the mulberry crop is fundamental because it can contribute towards improvement of bred varieties for high leaf yield needed for silkworm rearing as well as for food production for humans, however this requires acquisition of key skills in breeding and hybridization as well as proper selection of the varieties. Noted the improvement of both quantity and quality of leaf foliage becomes a breeder's aspiration (Tikader and Kamble, 2008).

Despite a great headway in mulberry production research progress, there is still inadequate information available on established genome sequences and complete single e nucleotide polymorphism sequence (SNPs) as well as all the chromosomes in the genus *Morus.* However, currently the first draft genome has paved way into genetic characterization of mulberry through use of wild type *M. notabilis*, (He et al., 2013) which has opened room for new discoveries of genetic information and research.

1.4.3 Sericulture industry challenges

Sericulture though a viable industry has been hindered by various challenges that require to be tackled at different levels such as technical level where proper capacity building of various specialized professions is needed along the value chain from production, processing and marketing for both research development as well as crop farm management. The growth will rely on proper infrastructure laid out for both field research on mulberry planting programs, rearing facilities, processing facilities storage as well as laboratory analysis. Furthermore, the marketing component should not be neglected it plays a key role in input output model of the

products in the value chain. Emphasis should also be places at the industrial side of the chain where quality and quantity should be adhered to for sustainability

1.5 General strategies of mulberry production and improvement

1.5.1 Identification of *Morus* species existence

Morus species are widespread although their primary origin centre is said to be China - Japan gene centre (Vavilov, 1926). The mulberry species taxonomy depended on the morphological features of the plant for decades. However, the complexity and diverse morphology existing in *Morus* species is related to the history of its naming, classification and identification. Mulberry has shown great ability to outbreed easily, with potential of high phenotypic plasticity and occurrence of interspecific hybridization coupled with mutations. These necessitated more studies to be undertaken. The varied phenotypic expressions (Gray and Gray, 1987; Awasthi et al., 2004; Burgess et al., 2005) and observed sexual traits (Datwyler and Weiblen, 2004) have been observed through mulberry studies comparison of available species and many varieties resulting from hybridization thereby creating room for a continuous identification debate theory.

Linneaus (1753) unveiled the initial classification of *Morus* resulting into seven species of *M. nigra, M. indica, M. rubra M. alba*, *M. tartarica*, *M. papyrifera* and *M. tinctoria* based on the morphological features characterization. With additional studies on observable morphological feature characteristics based on the length of the style, presence of pubescence and stigma hairiness, mulberry was further reclassified into 24 species and one subspecies across species, Koidzumi (1917). Due to many morphological features, additional methods of categorisation were sort with regard to leaf anatomical and wood characteristics which regrouped species into *M. nigra*, *M. alba*, *M. bombycis*, *M. latifolia*, *M indica*, and hybrid

crosses among others to enable narrow the arising debates on the diversity and relatedness of species (Biasiolo et al.,2004; Vijayan et al., 2004).

Necessity being mother of invention, more focus has been placed on utilization of mulberry and as such created several methods of identification of species such as assessment of the proteins and allozymes where the similarity in banding patterns of different 17 diploid cultivars helped group cultivars into *M. bombycis*, *M. alba*, and *M. latifolia* species and seven varietal groups with aim of understanding their protein quality for genetic variation and leaf protein profile in relation to silkworm production (Hisashi, 1982; Rao et al., 2011). Additionally, the assessment of enzymes like the peroxidases levels found in the leaves has been done highlighting the their roles in in enhancing plant defence mechanism against pathogens (Gai et al., 2017; Konno, 2011) as well as their use as potential to manage industrial wastes. Despite revelation, the use of peroxidases has been found to be unstable marker due to variation that may occur due to the environmental effects on them hence more research needs to be undertaken.

Currently, utilising quantitative analysis of vegetative traits of plant leaf features such as width, petiole and length as well as the qualitative ones such as the bud shape often during winter when they are in dormant state , colour of leaf, shape of leaf margins and apex in addition to the chilling requirements have formed the third identification category. Using such evaluation,(Chang et al., 2014) grouped different varieties into three clusters of "*M. laevigata*", "*M. atropurpurea*, *M. bombycis*, *M. australis* and *M. formosensis"* and third being *"M. alba* and *M. latifolia"* with a claim that even within this grouping some species though clustered had similarity features observed across them.

Furthermore, the advent of new technologies utilising molecular markers has redefined the final Classification and nomenclature of mulberry. Molecular characterization of mulberry has been prompted by the enormous differences that exist among the known species and unclear

use of morpho- biochemical classification witnessed in order to bring out the exact genetic relationships that could have emanated from aspects of mulberry nature of hybridization through processes of natural dispersion, targeted domestication, and evolution of the species. Proper characterization and utilization of mulberry resources has enhanced breeders and other interested parties to target the use of the DNA markers throughout the plant genomes to easily reproduce and analyse these differences.

Several molecular characterization initiatives have been adapted in this regard in order to differentiate and help understand the diversity within and among the species of mulberry. Key among them include the use of Polymerase Chain Reaction (PCR) technologies such as the inter-simple sequence repeat (ISSR) (Bhattacharya and Ranade, 2001; Awasthi et al.,2004; Vijayan et al., 2004; Vijayan et al., 2005; Vijayan et al., 2006; Zhao et al., 2006; Zhao et al., 2007; Ipek et al., 2012; Kalpana et al., 2012; Chikkaswamy and Chandra, 2014; Saha et al., 2016), directed amplification of minisatellite DNA (DAMD) (Bhattacharya and Ranade, 2001) and random amplified polymorphic DNA (RAPD) (Vijayan et al., 2004; Ozrenk et al., 2010; Kalpana et al., 2012; Wani et al., 2013; Vijayan et al., 2014; Naik et al., 2015; Kala et al., 2016) and markers targeting to finding out the possibility the identity of species genotypes.

ISSR markers have been used to determine the genetic diversity and protein content of some varieties (Vijayan et al., 2006; Kar et al., 2008) as well as further reclassification of the varieties based on production regions and the protein and sugar content resulting to reconstruction of dendrograms for associations. In addition, the use of nuclear DNA (nrDNA)-internal transcribed spacer (ITS) and the chloroplast trnL-trnF intergenic spacer (Nepal and Ferguson, 2012) showed some relationship complication among the phylogenies formed using the sequenced data of not being similar revealing thence deducing that *Morus* could be restricted to a non-monophyletic clade formation within Moreae.

1.5.2 Breeding of new varieties

Moraceae family classification has had a significant debate over the decades due to the species diversity as well as the different mechanism coupled with regional classification. Integrated strategies of classification have been adapted to encompass the morphological, genetic/molecular and biochemical with aim of unravelling the mystery behind the species diversity which has further deepened with the breeding of more varieties.

Studies have shown that there could be a three-tier classification system existing in classification adding to the confusion in the taxonomy that originates from the initial global level classification (Linnaeus, 1753), regional level classifications (Koizdumi, 1917; Hotta, 1958) as well as the local level classification based on the specific gene banks and scientist's role in the research of different varieties.

While basing on the population structure, Vijayan et al., (2004) noted that there is close relationships among some species (*M. latifolia, M. alba* and *M. bombycis*) that may cause confusion into the *Morus* classification of species. This emphasized that the similarities among the species still remains a major challenge to be tackled for success of formulating appropriate strategies to conserve these precious bred materials for future breeding programs with aim targeting varieties of desirable traits such as high sprouting and rooting (Yahiro, 1974), quantity of leaf yield per unit area, desired moisture content, capacity of moisture retention by the leaves, increased total shoot length as well as biotic and abiotic stresses tolerances. Majority of the mulberry species existing are diploid in nature but with a wide chromosome numbers ranging from basic 2n =14, to 2n =308 and varied ploidy levels (Dandin and Rajan, 1989; Awasthi et al.,2004; Machii et al., 2000; He et al., 2013; Yamanouchi et al., 2017).

1.5.3 Post-harvest handling of the products of mulberry

The handling of mulberry fruits and leaves has had innovations being brought onboard for industrial exploitation (Singhal et al., 2010). Key among them is the production branch

powders (Qiu and Zhang, 2019), artificial feeds for silkworm, preparation of mulberry teas and other food products such as mulberry juices and Jam (Buhroo et al., 2018; Yadav et al., 2014) as well wood as utilised in carvings and making sports.

1.6 Success of mulberry production, prospects and enhanced research

Achievements of the sericulture industry for decades has relied on the mulberry as a crop, however with the need for sustainability (Sarkar et al., 2017) under the changing climatic conditions, utilization due to advanced technologies and social living conditions, there is need for evaluating basic inputs of mulberry production with regards to monitoring for both biotic and abiotic factors such pests, diseases, temperature, climatic condition, types of soils and varieties (Babu, 2013). These factors play a role by impacting negatively on productivity of the mulberry if not controlled properly hence there is need for consistent research to be conducted while adhering to the adverse climatic changes to develop different markers for breeding new varieties that can mitigate climate change effects over time.

With advancement of the fast-growing evolution of mulberry production and utilisation, there is need to investigate the molecular structures to explore the various genetic information that can enhance the development of the industrial use (He et al., 2013). Establishment of *Morus* database (MorusDB) has advanced the mulberry research by offering a reference platform for other researchers to find out information on molecular markers, activities and functions to which mulberry can be put to use (He et al., 2013).

1.7 Phylogeny revealed by morphology and genetic materials

Correlation of expressed genes to morphological characteristic can be achieved through phylogeny construction and relationship analysis among species. However, this has not been easy as hypothesising the relationships often rely on either single characters such as genes,

morphology or the cellular component while being biased based on method and goals that are easily accessible (Nepal, 2008; Nepal and Ferguson,2012). More often the queries targeted usually give an insight into evolution of the genetics, development as well as the morphological traits based on a biased model system of research hypothesised necessitating the involvement of different strategies to be considered.

Morphological characters played key role in phylogenies due to ease of data collection and inexpensive technology however, it was found difficult to infer correctly the phylogeny as the species could be largely different in sequences and metabolic processes even though they look similar in their morphology. The challenge has since been addressed through exploiting the molecular methods where nucleotide and amino acid sequences can unravel the complexity. Combination of techniques has resulted to a better understanding of phylogeny of plants resulting to comparisons that yielded to creation of relationships within species through a process (**Fig. 2**) (Tekle et al., 2010; Patwardhan et al., 2014).

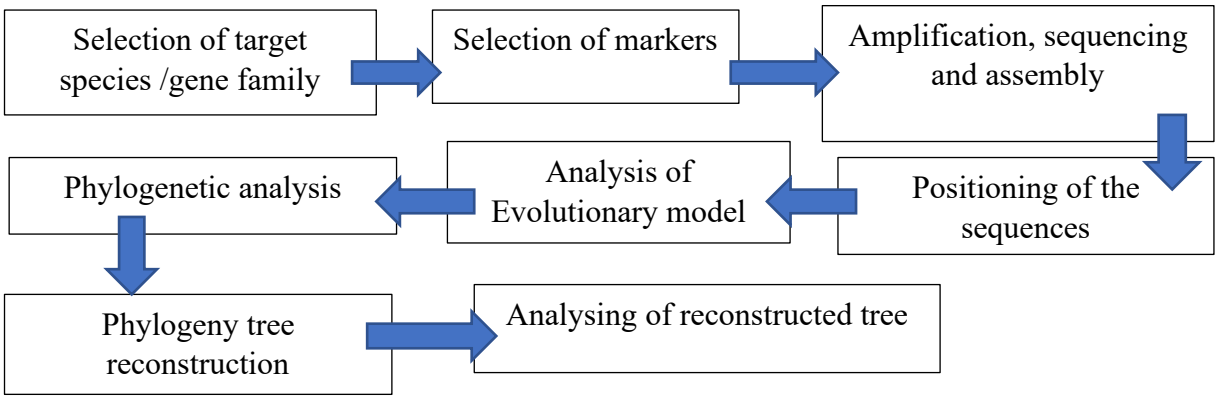

Figure 2. A common process in revealing the phylogenetic relationship of species.

1.8 Overview of mulberry phylogeny

Distribution of *Morus* species is wide and complex hence bound to be disputed. Understanding the relationships though phylogenetics that exist among them has helped try to narrow the gap through revealing the evolutionary history as well as categorising species. In

mulberry such quests have involved through use of molecular sequences such as the use of internal transcribed spacers (ITS) that has attempted to reveal the diversity in nature of *Morus* species as well as divergence time estimates through the classification of varieties into different key species namely; *M. mesozygia*, *M. insignis, M. nigra, M. celtidifolia, M. notabilis, M. rubra M. serrata and M. alba* (Zeng et al., 2015). In addition, their result indicated the formation of clear branches among the polyploids and the non-Asian species while the diploid ones did not make clear branches an indication of how closely related, they could be.

The genetic divergence time among the mulberry can be measured by phenetic appearances among the groups as well as genetic frequencies over time causing divergence. In moraceae studies conducted around the world, it has been noted that the common range of divergence falls between 5.3-30 million of years ago (MYA) according to the time tree information timescale of life (TTOL) for various pairwise taxa comparison (Kumar et al., 2017). Further attempts have revealed such comparison of *M. bombycis* against *M. acidosa; M. alba, M. notabilis* species respectively. Further reviewed data reveal that *M. bombycis vs M. latifolia* and *M. notabilis* have their divergence times ranging at 10 MYA while when compared against other species like ; *M. alba, M. cathayana*, *M. latifolia*, *M. indica* and *M. rotundiloba* their median divergence time becomes as low as 5.3 MYA. Based on such reflection we can deduce that the closeness of some varieties in their divergence may cause morphological characterization to be undependable as some of the features are almost alike compared to *M. mesozygia* whose divergence time falls at 23.4 MYA (Zarega et al., 2005; Chen et al., 2012; Zeng et al., 2015).

Apparently despite having great diversity very little has been done to establish the exact divergence of mulberry. There has been attempts that have yielded to three documented trees (Zarega et al., 2005; Chen et al., 2012; Zeng et al., 2015).

Further build-up of information on *Morus* species relationships unveiled the first genetic linkage map with the help of molecular methods that entailed use of random amplified polymorphic DNA (RAPD), Inter simple sequence repeats (ISSR), Simple sequence repeat (SSR) and two-way cross mapping strategy (Venkateswarlu et al., 2006) prompting more uncertainties and more research into breeding and relationship analysis among the *Morus*. Attempts to differentiate Japanese and Indian mulberry using the ISSR and RAPD primers showed similarities among 18 genotypes tested, and further clustering them into two groups based on geographic origin and species status location is an example (Vijayan et al., 2004).

1.9 Morphological characterization of mulberry

Taxonomists over time have used the phenotypic expression of plant morphology to classify the *Morus* genus. The characteristics mostly relied on the observable plant parts such as the plant form, female style, stigma type (hairs or protrusions), branching patterns, morphology of leaf, description of flower and fruit shapes as well as the agronomic properties (Wani et al., 2013). Such phenotypes have formed the basis for morphological characterization universally resulting to grouping of various species of mulberry into different accessions all over the growing regions. Utilising the plant descriptive characteristics of the shoots, leaves, bark, nodes, internodes, texture, phyllotaxy, flowers, fruits, leaf and fruit colour, leaf angles as well as the use of scanning electron microscopy (SEM) for epidermis observations (Biasiolo et al., 2004) among others has gained prominence.

In Kenya, not much of the characterization has been done on mulberry varieties, however different varieties of *Morus* species exist that are unclassified properly but adaptable and growing to different regions based on the rainfall requirement and the popularity in production for use in rearing the silkworm such as ‘Embu’, ‘Thika’, ‘Limuru’, and ‘Ithanga’ (**Table 1**).

Table 1. Characteristics of the local varieties grown in Kenya

Variety	Characteristics
Embu	Short internodes, reddish bark and small leaves with high drought resistance
Thika	Long internodes, whitish bark and large light green leaves with slight drought tolerance
Limuru	Short thin internodes, small finger shaped serrated leaves
Ithanga	Medium heart shaped smooth light green leaves

Although Africa is known to be origin of *M. mesozygia,* less research has been done except for utilising in local traditional medicines and wood (Toirambe and Ouattara, 2008; Machii et al., 2000), the above species has morphological characteristic features of dark green leaves with three veins arising from the leaf base yet the variety documented resembling it ('Enbu') doesn't hence the collection could be either products of native varieties or introductions done through the Overseas Technical Cooperation Agency (OTCA), Japan in collaboration with the Ministry of Agriculture in Kenya in 1972. Currently there is development and collection of mulberry species bioresources resulting to availability of *M. latifolia, M. rotundiloba*, *M .alba* and *M. indica* respectively in Kenya through introductions such as Thailand, 'Noi', 'Wasemidori' and 'Ichinose' from Japan, and India's 'Kanva 2' and 'S-36' varieties among others offering more opportunity for further research and development through multiplication and breeding.

1.10 Genetic characterization of mulberry and the importance

Although morphological and phenological characterization of mulberry has revealed a wealth of information in different regions (Balik et al., 2019; Machii et al., 1999), mulberry identification has not been easy due to ambiguous, time-consuming (Vijayan, 2010) and subjective nature while considering the highly heterozygous mulberry nature which may result

to high diversity in the gene pool through hybridization that may complicate identification as a result of inadequate proper and stable genetic markers, selection strategies, tedious nature and long breeding cycle.

Quest for easy, quick and precise methods has tapped into the genome sequencing techniques providing a range of markers through complete nucleotide sequences and physical mapping of the chromosomes. Molecular understanding the genetic material has offered an opportunity to improve them for productivity and adaptation. However, eukaryotic genomes present a major hinderance due to their heterozygous nature forming many heterozygous positions, insertions/deletions polymorphism, high copy number variation as well as small-scale re-arrangements. *Morus* genus being a eukaryote has not been spared either although currently establishment of *Morus* database having the first draft genome sequence of *Morus notabilis* available has offered opportunity for further comparison studies of in-depth understanding of the genus *Morus* (He et al., 2013; Li et al., 2014).

Genetic mechanisms in combination with the environmental influence within the different plants are known to trigger variations allowing for regulation of interspecific hybridization, polyploidization and genome change during the process of meiosis and mitosis thus bringing about diversity both at adaptation as well as at chromosomal level (Tikader and Kamble, 2008). Mulberry due to ease of hybridization, a variation in ploidy levels has been observed ranging from 28 (most of the *Morus* species) to 308 ('Kuromiguwa', a variety of *M. nigra*) (Sharma et al., 2000). Previous studies had placed the basic number of mulberry chromosomes as 14 with species having 28 chromosomes dominating the genus *Morus* (Venkatesh and Munirajappa, 2015), however, current research suggests that the basic number of chromosome stands at 7 (He et al., 2013) prompting the query of how important is the duplication in whole genome sequencing in the strive of mulberry production and identification and breeding for polyploids.

Interestingly the high plasticity observed in mulberry morphology has created a high range of confusion in the classification among the species with well noted effects of polyploidy complexes existing especially in species of *M. bombycis* (Machii et al., 2000) that necessitates research to be carried out. As such utilisation of different genetic markers such as use of RAPD, ISSR, SSR had to be adapted which have yielded results for benchmarking and further research into mulberry such as the generation of the first mulberry genetic linkage map (Venkateswarlu et al., 2006) and six chloroplast genome sequences that have given an insight into the differences occurring in *Morus* genome sequence size among several species such as *M. cathayana* 159,265 bp, *M. multicaulis* 159,103 bp, *M. indica* 158,484 bp, *M. mongolica* 158,459 bp, *M. notabilis* 158,680 bp and *M. atropurpurea* 159,113 bp respectively (Kong and Yang, 2017; Li et al., 2016). The comparative studies above and inferences made to the existing database have played a major role in ascertaining the relationships among mulberry species (Kong and Yang, 2017; Li et al., 2016). Moreover, although these complimentary DNA (cDNA) sequence of Moraceae exist, not all species have been covered hence the need for continued research.

Gene duplication in mulberry has been noted as a source of variation which can be expressed in transcripts resulting to phenotypic changes during molecular evolution hence more emphasis has been tailored towards mulberry genes after polyploidization through RNA-Seq (Dai et al., 2015). Transcriptomes are being utilised to identify genes of importance being influenced by either the hormones or functions while using the established MorusDB as a reference platform for comparative studies (Li et al., 2014). The use of comparative genomics in relation to available public databases such as NCBI-nr, gene ontology (GO) and Kyoto encyclopaedia of genes and genomes (KEGG), among the many existing with curated nucleotide and protein sequences as well as pathways will enable a great insight into functionality of identified genes of *Morus* species. These public databases have been curated

with continued review on addition of information in order to play a major inference role in both the genic domains, structure and transcript analysis for different organisms as such may be useful in existing autoploid and diploid mulberry varieties there by revealing information at different levels with regard to molecular function, biological processes and cellular components thereby offering opportunity for proper utilization of the genomic data obtained from transcripts among others (Dai et al., 2015; Saeed et al., 2016).

1.11 Marker development and utilisation

Mulberry genetic enhancement and molecular breeding has been a challenge as a result of unavailable genomic resources with stable markers for adoption (Mathithumilan et al., 2013). Basic studies utilising the genomic and expressed sequence tags (ESTs) have shown high heterozygosity combined with high polymorphic information providing high potential for diversity analysis giving a hint the appropriate adoption of genome-wide markers. Through the next generation sequencing will result to ease of monitoring the heritable polymorphism can be achieved. Research has shown that identification of recombinant breakpoints important for linkage and quantitative trait loci mapping, genotyping as well as resolving the relationships existing in wild populations throughout the progenies among others may be achieved by use of molecular markers usually have a high degree of accuracy (Vijayan, 2007; Vijayan, 2010; Davey et al., 2011; Krishnan et al., 2014). By adapting the marker based technologies for proper deciphering genetic relatedness, population sizes and levels of inbreeding among different plant species (Awasthi et al., 2004; Ekblom and Wolf, 2014), enhanced mapping and reconstruction of phylogenies has been achieved. However, the above has been slow mulberry classification since the previous drive was tailored towards high leaf yield production for silkworm rearing through evaluation of the best agronomic practices, harvesting modalities, forage preservation, nutrition as well as utilization of various products and animal trials, hence

current attention needs to be directed into bringing together the varieties under a clear genealogy for both conservation and increased value-added utilization. The use of DNA markers should be adopted for utilisation in these mulberry assessments due to their stable nature of not being influenced by environment for similarity determination of the many genotypes available around the world. Efforts regarding use of DNA markers have been made through clearly documented achievements using various markers for characterization of the accessions such as ISSR, SSR, RAPD, sequence characterized amplified regions (SCAR), directed amplification of minisatellite DNA (DAMD) and amplified fragment length polymorphism (AFLP) methods based on Polymerase Chain have contributed (Ipek et al., 2012). The current smart innovations like the single molecule real time sequencing (SMART) DNA sequencing is further offering great and fast solutions for identifications of unique characteristics that encompass long reads, lack of G-C bias, pulse width and base modifications among others (Kim et al., 2014).

1.12 Current research on genes in mulberry

Genome sequencing strategy has been adapted in mulberry to facilitate the understanding of the genetic functioning of inherent components (Khurana and Checker, 2011). As such several crop improvement advances have been fronted such finding microsatellite markers for high degree of polymorphism in improvement of *M. indica* (Aggarwal et al., 2004). Furthermore, in endeavour to adjust to climatic mitigation strategies, more research is being directed towards different varieties for acclimatization to a wide range of climatic conditions through search and discovering of more tolerance genes required for adaptation against biotic and abiotic stresses (Checker and Khurana, 2013). Examples include the finding markers for heavy metal detoxification (Fan et al., 2017), surviving the drought (Gai et al., 2018), improving growth and development of the plant (Luo et al., 2016; Luo et al., 2018), ethylene

biosynthesis (Liu et al., 2015; Shang et al., 2014), combinations for heat, cold, drought and salt stress (Baranwal and Khurana, 2017) among others using different plant tissues like leaf, root, winter bud, bark and male flower.

Based on the above, my motivation arose with the drive to conduct research for the development of sericulture industry now and for future generations through conducting morphological and genetic characterization considering the state of dwindling mulberry and silk production worldwide. In Kenya based on survey carried out by the High Tech Silk Project (HTSP) 2014 report revealed that although there is high potential, the mulberry production and silkworm rearing is being conducted on a small scale and requires to be up-scaled to the industry level for both mulberry and silk production through proper research methods along the value chain. However, this scaling-up is currently being impeded by a myriad of challenges not limited to lack of properly established knowledge of mulberry, established research mechanism from selection, breeding and multiplication as well as the insufficient technical capacity building coupled with other biotic and abiotic stresses. The need to dive into commercialisation of the mulberry for better returns in Kenya therefore requires proper understanding of the system from the basic knowledge required for production, maintenance and further improvement of the crop. This knowledge base can be unlocked if there is proper identification of existing varieties and further isolating their lineage of 'Enbu' while comparing it further to other varieties existing in the NARO Genebank which will further enhance development of better production, breeding and conservation strategies with less negative influence on the environment and epistasis while offering a platform for further research.

1.13 General objectives

Overall objective was to conduct morphological and genomic characterization of various mulberry varieties while targeting the 'Enbu' genome sequencing mulberry to establish its lineage and to identify potential markers.

1.14 Specific objectives

The specific objectives entailed;

- To collect and characterize diploid mulberry varieties by both primary and secondary morphological features.
- To utilise the double digest restriction site associated DNA sequencing technique for marker discovery and phylogenetic relationship among selected mulberry varieties.
- To generate a high-quality genome sequence of 'Enbu' variety and compare it with existing variety 'Chuansang' *Morus notabilis* from the MorusDB.
- Development of conservation and preservation strategies

1.15 Research hypothesis

My hypothesis was grounded on the fact that mulberry being a perennial crop that outbreeds easily may exhibit high degree of heterozygosity causing variations in the varieties. Currently in Kenya Embu variety is grown widely for silkworm rearing however, the true species identity of this variety is not known. Furthermore, the need to understand the lineage of this indigenous mulberry variety 'Enbu' has been key noting that it had been registered as an African mulberry *M. mesozygia* a NARO Genebank of Japan, yet the morphological features appear different. My research therefore targets to upgrade and develop knowledge for further utilization through understanding the origin, habitat locations, its production, sampling mechanisms and

documentation of the DNA content. In addition, research on characterization of various mulberry varieties including 'Enbu' targets at identification of different markers at both morphological and genetic level for mulberry enhancement and utilization.

1.16 References

Aggarwal R. K., Udaykumar D., Hendre P. S., Sarkar A., and Singh,L. I. (2004). Isolation and characterization of six novel microsatellite markers for mulberry (*Morus indica*). *Molecular Ecology Notes*, 4(3), 477–479.

Al-KirshiR. A., Alimon A., Zulkifli I., Atefeh S., Zahari M. W., and Ivan M. (2013). Nutrient digestibility of mulberry leaves (Morus alba). *Italian Journal of Animal Science*, *12*(2), 219–221.

Al-Kirshi R., Alimon A., Zulkifli I., Sazili A., Zahari M. W., and Ivan M. (2010). Utilization of mulberry leaf meal (Morus alba) as protein supplement in diets for laying hens. *Italian Journal of Animal Science*, *9*(3), 265–267.

Aramwit P., Bang N., and Srichana T., (2010). The properties and stability of anthocyanins in mulberry fruits. *Food Research International*, 43(4), 1093–1097.

Awasthi A. K., Nagaraja G. M., Naik G. V., Kanginakudru S., Thangavelu K., and Nagaraju J. (2004). Genetic diversity and relationships in mulberry (genus *Morus*) as revealed by RAPD and ISSR marker assays. *BioMed Central Genetics*, 5(1), 1–9.

Babu K. M. (2013). Silk: Processing, Properties and Applications; In *Woodhead Publishing Series in Textiles* (1st Edition). Woodhead Publishing Limited. https://doi.org/10.1533/9781782421580.

Bae S. H., and Suh H. J. (2007). Antioxidant activities of five different mulberry cultivars in

Korea. *LWT - Food Science and Technology*, 40(6), 955–962.

Balik A., Geçer M. K., and Aslantaş R. (2019). Diversity of biochemical content in fruits of some indigenous mulberry genotypes. *Turkish Journal of Agriculture and Forestry*, 43(1), 28–35.

Baranwal V. K., and Khurana P. (2017). Major intrinsic proteins repertoire of Morus notabilis and their expression profiles in different species. *Plant Physiology and Biochemistry*, 111, 304–317.

Barchi L., Lanteri S., Portis E., Acquadro A., Valè G., Toppino L., and Rotino G. L. (2011). Identification of SNP and SSR markers in eggplant using RAD tag sequencing. *BioMed Central Genomics*. 12, 304-313.

Bhattacharya E., and Ranade S. A., (2001). Molecular distinction amongst varieties of Mulberry using RAPD and DAMD profiles. *BioMed Central Plant Biology*, 1(1), 3.

Biasiolo M., Teresa M., Canal D. A., and Tornadore N., (2004). Micromorphological Characterization of Ten Mulberry Cultivar*s (Morus* Spp.*)*, Economic Botany. *5*8(4), 639–646.

Buhroo Z. I., Bhat M. A., Kamili A. S., Ganai N. A., Bal, G. K., Khan I. L., and Aziz A. (2018). Trends in Development and Utilization of Sericulture Resources for Diversification and Value Addition. *Journal of Entomology and Zoology Studies*, 6(4), 601–615 .

Burgess K. S., Morgan M., Deverno L., and Husband B. C. (2005). Asymmetrical introgression between two *Morus* species (*M. alba, M. rubra*) that differ in abundance. *Molecular Ecology*, 14(11), 3471–3483.

Byng J. W., Chase M W. C, Christenhusz, M. J. M., Fay M. F.,Judd W. S.,Mabberley D. J.,

Sennikov A. N., Soltis D. E., Soltis P. S., and Stevens, P.F.(2016).An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society*. 181, 1-20.

Chan E.W., Lye P., Wong S.,(2016). Phytochemistry, pharmacology, and clinical trials of *Morus alba*. *Chinese Journal of Natural Medicines*. 14(1), 0017-0030

Chang L., Juang L., Wang B., Wang M., Tai H., Hung W., Chen Y., and Huang M.,. (2010). Antioxidant and antityrosinase activity of mulberry (Morus alba L.) twigs and root bark. *Food and Chemical Toxicology*, 49(4), 785–790.

Chang L., Li, K., Yang, W., Chang, J., and Chang, M., (2014). Phenotypic classification of mulberry (*Morus)* species in Taiwan using numerical taxonomic analysis through the characterization of vegetative traits and chilling requirements. *Scientia Horticulturae*, 176, 208–217.

Checker V. G., and Khurana P. (2013). Molecular and functional characterization of mulberry EST encoding remorin (MiREM) involved in abiotic stress. *Plant Cell Reports*, 32(11), 1729–1741.

Chen C., Liu L.., Hsu J., Huang H., Yang M. , and Wang C.., (2005). Mulberry extract inhibits the development of atherosclerosis in cholesterol-fed rabbits. *Food Chemistry*, 91(4), 601–607.

Chen J., and Li X. (2007). Hypolipidemic effect of flavonoids from mulberry leaves in triton WR-1339 induced hyperlipidemic mice. *Asia Pacific Journal of Clinical Nutrition*, *16*(SUPPL.1), 290–294.

Chen R., Zhang Z., Sun W., Fu Y., Zhang Q., Wei L., Liang Y., and Dong R., (2012). System Position and Divergent Time of Based on ITS Sequence Humulus scandens. *American Journal of Plant Sciences*, 3(10), 1470–1475.

Chikkaswamy B. K., and Chandra R. (2014). An investigation of RAPD and ISSR molecular marker studies in Mulberry varieties. *IntenationalJournal of .Current Microbiology and Applied Sciences*, 3(6), 986–995.

Dai F., Wang Z., Luo G., and Tang C., (2015). Phenotypic and transcriptomic analyses of autotetraploid and diploid mulberry (*Morus alba* L.). *International Journal of Molecular Sciences*, 16(9), 22938–22956.

Dandin S. B., and Rajan M. V., (1989). Microsporogenesis in Hexaploid *Morus serrata* Roxb. *Cytologia*, 54, 747–751.

Das B.C, Prasad D.N.,and Sikdar A.K.,(1970).Cochicine induced tetraploids of mulberry.*Caryologia*, 23(3), 283-293.

Datta, R K. (2000). Mulberry Cultivation and Utilization in India. *FAO Electronic Conference on Mulberry For Anımal Productıon (Morus-L)*, 1–24.

Datta, Rajat K, and Nanavaty, M. (2005). *Global Silk Industry: A Complete Source book*. Universal Publisher USA. http://www.bookpump.com/bwp/pdf-b/1124937b.pdf.

Datwyler, S. L., and Weiblen, G. D. (2004). On the origin of the fig: Phylogenetic relationships of Moraceae from NDHF sequences. *American Journal of Botany*, 91(5), 767–777.

Davey J., Hohenlohe P., Etter P., Boone J., Catchen J. M., and Blaxter M. L., (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, 12(7), 499–510.

De Souza M. M., Bittar M., Cechinel-Filho V., Yunes R. A., Messana I., Monache F. D., and Ferrari F., (2000). Antinociceptive properties of morusin, a prenylflavonoid isolated from Morus nigra root bark. *Zeitschrift Fur Naturforschung Section C Journal of*

Biosciences, 55(3–4), 256–260.

De Storme N., and Mason A., (2014). Plant speciation through chromosome instability and ploidy change: Cellular mechanisms, molecular factors and evolutionary relevance. *Current Plant Biology*, 1, 10–33.

Dong Z., Ding Z., Zhang S., Zhang Y., Fan H., and Yang Y., (2017). Natural fibers from the bark of mulberry branches for textile application. *Fibres and Textiles in Eastern Europe*, 25(3), 20–25.

Du Q., Zheng J., and Xu Y.,(2008). Composition of anthocyanins in mulberry and their antioxidant activity. *Journal of Food Composition and Analysis*, 21(5), 390–395.

Ebrahim A., Animut G., Urge M., and Hailemichael A., (2015). Mulberry (*Morus alba*) as emerging potential opportunity for livestock feed development in Northern Ethiopia. *Journal of Agricultural Science and Food Technology*, 1(5), 66–72.

Ekblom R., and Wolf J. B. W., (2014). A field guide to whole-genome sequencing, assembly and annotation. *Evolutionary Applications*, 7(9), 1026–1042.

Ellegren H., (2014). Genome sequencing and population genomics in non-model organisms. In *Trends in Ecology and Evolution,* 29(1), 51–63.

Eo H., Park J., Park, G., Lee, M., Lee J., Koo J., and Jeong J., (2014). Anti-inflammatory and anti-cancer activity of mulberry (*Morus alba* L.) root bark. *BioMed CentralComplementary and Alternative Medicine*, 14(1), 200–209.

Fan W., Guo Q., Liu C., Liu X., Zhang M., Long D., Xiang Z., and Zhao A., (2017). Two mulberry phytochelatin synthase genes confer zinc/cadmium tolerance and accumulation in transgenic Arabidopsis and tobacco. *Gene*, 645, 95–104.

Fukui, K. (2009). Effects of temperature on growth and dry matter accumulation in mulberry

saplings. *Plant Production Science*, *3*(4), 404–409.

Gai Y., Yuan S., Zhao, Y, Zhao, H., Zhang H., and Ji X.,(2018). A Novel LncRNA, MuLnc1, associated with environmental stress in mulberry (*Morus multicaulis*). *Frontiers in Plant Science*, 9, 669-682.

Gai Y., Zhao, Y., Zhao H., Yuan C., Yuan S., Li S., Zhu B., and Ji X., (2017). The latex protein MLX56 from mulberry (*Morus multic*aulis) protects plants against insect pests and pathogens. *Frontiers in Plant Science*, 8, 1475–1486.

Gao K., Zheng C., Wang T., Zhao H., Wang J., Wang Z., Zhai X., Jia Z., Chen J., Zhou Y., and Wang W., (2016). 1-Deoxynojirimycin: Occurrence, extraction, chemistry, oral pharmacokinetics, biological activities and in silico target fishing. *Molecules*, 21(11), 1600–1615.

Gao M., Nguyen T. T., Suckow, M. A. Wolter W. R., Gooyit M., Mobashery S., and Chang M., (2015). Acceleration of diabetic wound healing using a novel protease–anti-protease combination therapy. *Proceedings of the National Academy of Sciences*, 112(49), 15226–15231.

Golpayegani A. S., Thévenon M. F., Gril J., Masson E., and Pourtahmasi K., (2014). Toxicity potential in the extraneous compounds of white mulberry wood (Morus alba). *Maderas: Ciencia y Tecnologia*, 16(2), 227–238.

Gray E., and Gray R. E., (1987). Leaf Lobation Patterns in Mulberry. *Castanea*, 52(3), 216–224.

Gryn-Rynko A., Bazylak G., and Olszewska-Slonina D., (2016). New potential phytotherapeutics obtained from white mulberry (*Morus alba* L.) leaves. *Biomedicine and Pharmacotherapy*, 84, 628–636.

Harauma A., Murayama T., Ikeyama K., Sano H., Arai H., Takano R., Kita T., Hara S., Kamei K., and Yokode M., (2007). Mulberry leaf powder prevents atherosclerosis in apolipoprotein E-deficient mice. *Biochemical and Biophysical Research Communications*, 358(3), 751–756.

He N., Zhang C., Qi X., Zhao S., Tao Y., Yang G., Lee T., Wang X., Cai Q., Li D., Lu M., Liao S., Luo G., He R., Tan X., Xu Y., Li T., Zhao A., Jia L., … Xiang Z. (2013). Draft genome sequence of the mulberry tree *Morus notabilis. Nature Communications*, 4, 2445–2454.

Hisashi H., (1982). Varietal Differences of leaf protein profiles in mulberry. *Phytochemistry*, 21 (7), 1513–1518.

Hotta T., (1958). Taxonomical studies on the *Morus* plants and their distributions in Japan and its vicinities. Japanese Society for Promotion of Science, Tokyo.

Huang H., Ou T., and Wang C., (2013). Mulberry (Sang Shèn Zǐ) and its Bioactive Compounds, the Chemoprevention Effects and Molecular Mechanisms In Vitro and In Vivo. *Journal of Traditional and Complementary Medicine*, 3(1), 7–15.

Imran M., Khan H., Shah M., Khan R., and Khan F., (2010). Chemical composition and antioxidant activity of certain *Morus* species. *Journal of Zhejiang University- Science B (Biomedicine &Biotechnology)*, 11(12), 973–980.

Ipek M., Pirlak L., and Kafkas S., (2012). Molecular characterization of mulberry (*Morus* spp.) genotypes via RAPD and ISSR. *Journal of the Science of Food and Agriculture*, 92(8), 1633–1637.

Jiang D., Guo Y., Xu D., Huang Y., Yuan K., and Lv Z. , (2013). Antioxidant and anti-fatigue effects of anthocyanins of mulberry juice purification (MJP) and mulberry marc purification (MMP) from different varieties mulberry fruit in China. *Food and Chemical*

Toxicology, 59, 1–7.

Jiang Y., Huang R., Yan X., Jia C., Jiang S., and Long T., (2017). Mulberry for environmental protection. *Pakistan Journal of Botany*, 49(2), 781–788.

Jiao Y., Wang X., Jiang X., Kong F., Wang S., and Yan C., (2017). Antidiabetic effects of *Morus alba* fruit polysaccharides on high-fat diet- and streptozotocin-induced type 2 diabetes in rats. *Journal of Ethnopharmacology*, 199, 119–127.

Jin Q., Yang J., Ma L., Cai J., and L, J., (2015). Comparison of polyphenol profile and inhibitory activities against oxidation and α-Glucosidase in mulberry (genus *Morus)* cultivars from China. *Journal of Food Science*, 80(11), C2440–C2451.

Kähkönen M. P., and Heinonen M., (2003). Antioxidant activity of anthocyanins and their aglycons. *Journal of Agricultural and Food Chemistry*, 51(3), 628–633.

Kala P., Zargar S. M., Bali R. K., Gupta N., Salgotra R. K., and Koul A., (2016). Assessment of genetic diversity in mulberry using morphological and molecular markers. *Electronic Journal of Plant Breeding*, 7(1), 94–103.

Kalpana D., Choi S., Choi T., Senthil K., and Lee Y., (2012). Assessment of genetic diversity among varieties of mulberry using RAPD and ISSR fingerprinting. S*cientia Horticulturae,* 134, 79–87.

Kang T. H., Hur J. Y., Kim H. B., Ryu J. H., and Kim S. Y., (2006). Neuroprotective effects of the cyanidin-3-O-β-D-glucopyranoside isolated from mulberry fruit against cerebral ischemia. *Neuroscience Letters*, 391(3), 168–172.

Kar P. K., Srivastava P.P., Awasthi A. K., Urs S. R., (2008). Genetic variability and association of ISSR markers with some biochemical traits in mulberry (*Morus* spp.) genetic resources available in India. *Tree Genetics and Genomes*, 4(1), 75-83.

Kasote D. M., Katyare S. S., Hegde M. V., and Bae H., (2015). Significance of antioxidant potential of plants and its relevance to therapeutic applications. *International Journal of Biological Sciences*, 11(8), 982–991.

Kobayashi Y., Miyazawa M., Kamei A., Abe K., and Kojima A., Ameliorative effects of mulberry (Morus alba L.) leaves on hyperlipidemia in Rats fed a high-fat diet: Induction of fatty acid oxidation, inhibition of lipogenesis, and suppression of oxidative stress. *Bioscience, Biotechnology, and Biochemistry*, 74(12), 2385-2395.

Koidzumi C., (1917). Taxonomical discussion on *Morus* plants. .Bull. Imp. sericult.Exp.Stat. Japan, 3, 1-62.

Kong W., and Yang J., (2017). The complete chloroplast genome sequence of *Morus cathayana* and *Morus multicauli*s a comparative analysis within genus *Morus* L. *Peer Journal*, 62(1), 165–172.

Khurana P., and Checker V. G., (2011). The advent of genomics in mulberry and perspectives for productivity enhancement. *Plant Cell Report,* 30(5), 825–838.

Kim H. G., Ju M. S., Shim, J. S., Kim, M. C., Lee, S. H., Huh, Y., Kim, S. Y., and Oh, M. S. (2010). Mulberry fruit protects dopaminergic neurons in toxin-induced Parkinson's disease models. *British Journal of Nutrition*, 104(1), 8–16.

Kim K. E., Peluso P., Babayan P., Yeadon P. J., Yu C., Fisher W. W., Chin C. S., Rapicavoli N. A., Rank D. R., Li J., Catcheside D. E. A., Celniker S. E., Phillippy A. M., Bergman C. M., and Landolin J. M., (2014). Long-read, whole-genome shotgun sequence data for five model organisms. *Scientific Data*, *1*, 140045–140055.

Kole C. (Ed.)., (2011). *Wild Crop Relatives:Genomic and breeding resources: Tropical and subtropical fruits.* Springer Heidelberg Dordrecht London New York.

Kong W., and Yang J. (2017). The complete chloroplast genome sequence of Morus cathayana and Morus multicaulis,and a comparative analysis within genus Morus L. *Peer Journal*, 62(1), 165–172.

Konno K., (2011). Plant latex and other exudates as plant defense systems: Roles of various defense chemicals and proteins contained therein. *Phytochemistry*, 72(13), 1510–1530.

Koren S., Walenz B. P., Berlin K., Miller J. R., Bergman N H., and Phillipy. A M., (2017). CANU: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, 27, 722–736.

Krishna H., Singh D., Singh R. S., Kumar L., Sharma B. D., and Saroj P. L., (2018). Morphological and antioxidant characteristics of mulberry *(Morus* spp.) genotypes. *Journal of the Saudi Society of Agricultural Sciences*, doi.org/10.1016/j.jssas.2018.08.002.

Krishnan R. R., Naik V. G., Ramesh S. R., and Qadri S. M. H., (2014). Microsatellite marker analysis reveals the events of the introduction and spread of cultivated mulberry in the Indian subcontinent. *Plant Genetic Resources*, 12(1), 129–139.

Kumar R., V., Kumar D., and Pher R. (2014). Varietal influence of mulberry on silkworm, *Bombyx mori* L. growth and development. *International Journal of Advanced Research*, 2(3), 921–927.

Kumar S., Stecher G., Suleski M., and Hedges S. B., (2017). TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Molecular Biology and Evolution*, 34(7), 1812–1819.

Linneaus C.(1753). *Morus* in Species plantarum. Stockholm, Sweden ,Vol.2 p.986

Li Q., Guo, J, Yan N., and Li C., (2016). Complete chloroplast genome sequence of

cultivated *Morus* L . species. *Genetics and Molecular Research*, 15(4), 60–73.

Li T., Qi X., Zeng Q., Xiang Z., and He N., (2014). MorusDB: A resource for mulberry genomics and genome biology. *Database(Oxford)*, 2014, bau054.

Lim S., and Choi C., (2019). Pharmacological properties of *Morus nigra* L. (Black Mulberry) as a promising nutraceutical resource. *Nutrients*, 11(2), 437–455.

Liu C., Zhao A., Zhu P., Li J., Han L., Wang X., Fan W., Lü R., Wang C., Li Z., Lu C., and Yu M., (2015). Characterization and expression of genes involved in the ethylene biosynthesis and signal transduction during ripening of mulberry fruit. *PLoS ONE*, 10(3), 1–17.

Liu Y, Ji D., Turgeon R., Chen J., Lin T., Huang J., Luo J., Zhu Y., Zhang C., and Lv Z., (2019). Physiological and proteomic responses of mulberry trees (*Morus alba*. L.) to combined salt and drought stress. *International Journal of Molecular Sciences*, 20(10), 2486–2506.

Liu Y., and Willison J. H. M., (2013). Prospects for cultivating white mulberry (*Morus alba*) in the drawdown zone of the three gorges reservoir, China. *Environmental Science and Pollution Research*, 20(10), 7142–7151.

Lown M., Fuller R., Lightowler H., Fraser A., Gallagher A., Stuart B., Byrne C., and Lewith G., (2017). Mulberry-extract improves glucose tolerance and decreases insulin concentrations in normoglycaemic adults: Results of a randomised double-blind placebo-controlled study. *PLoS ONE*, 12(2), 1–14.

Luo Y., Ma B., Zeng Q., Xiang Z., and He N., (2016). Identification and characterization of lateral organ boundaries domain genes in mulberry, *Morus notabilis*. *Meta Gene*, 8, 44–50.

Luo Y., Li H., Xiang Z., and He N.,(2018). Identification of *Morus notabilis* MADS-box genes and elucidation of the roles of MnMADS33 during endodormancy. *Scientific Reports* 8(1), 5860-5876.

Machii H., and Katagiri K., (1991). Varietal differences in nutritive values of mulberry leaves for rearing silkworms. *Japan Agricultural Research Quaterly*, 25, 202–208.

Machii H., Koyama A., and Yamanouchi H., (2000). Mulberry breeding, cultivation and utilization in Japan. *FAO Electronic Conference on Mulberry for AnimalProduction* 8, 63-72.

Machii H., Koyama A., and Yamanouchi H., (1999). Fruit traits of genetic mulberry resources. *Journal of Sericultural Science of Japan*, 68(2), 145–155.

Mammadov J., Aggarwal R., Buyyarapu R., and Kumpatla S., (2012). SNP markers and their impact on plant breeding. *International Journal of Plant Genomics*, *2*012, 387–413.

Mathithumilan B.S., and Dandin B.,(2009).Genetic Analysis of Diploid and Colchi-Tetraploid Mulberry (*Morus indica* and *Morus alba*) by molecular and morphological Markers. *International Journal of plant breeding,* 3(1), 58-64.

Mathithumilan B., Sajeevan R. S., Biradar J., Madhuri T., Nataraja K. N., and Sreeman, S. M., (2016). Development and characterization of genic SSR markers from Indian mulberry transcriptome and their transferability to related species of Moraceae. *PLoS ONE*, 11(9), 1–18.

Mathithumilan B.,Kadam N. N.,Biradar J.,Reddy S. H.,Ankaiah M.,Narayanan M. J., Makarla U., Khurana P.,and Sreeman S. M., (2013). Development and characterization of microsatellite markers for *Morus* spp. and assessment of their transferability to other closely related species. *BioMed Central Plant Biology*, 13(1), 194–215.

Naik V. G., Dandin S. B., Tikader A., and Pinto M. V.,(2015). Molecular diversity of wild

mulberry (*Morus* spp .) of Indian subcontinent. *Indian Journal of Biotechnology*, 14, 334–343.

Nair S. K., Yun-Gen M., and Kumar N. S.,(2005).Differential response of silkworm, *Bombyx mori* L. to phytoecdysteroid depending on the time of administration. *Journal of Applied Sciences and Environmental Management*, 9(3), 81-86.

Nazim N., Buhroo Z., Mushtaq N., Javid K., Rasool S., and Mir G., (2017). Medicinal values of products and by products of sericulture. *Journal of Pharmacognosy and Phytochemistry*, 6(5), 1388–1392.

Nepal M., and Ferguson C. J., (2012). Phylogenetics of *Morus* (Moraceae) inferred from ITS and trnL - trnF sequence data. *Systematic Botany*, 37(2), 442–450.

Nepal M. P. (2008). Systematics and reproductive biology of the genus *Morus* L. (Moraceae). Doctorate thesis in *Kansas state university*, 1-189.

Niwa R., Namiki T., Ito K.,Shimada-Niwa Y., Kiuchi M., Kawaoka S., Kayukawa T., Banno Y., Fujimoto Y., Shigenobu S., Kobayashi S., Shimada T.,Katsuma S., and Shinoda T.,(2010). Non-molting glossy/shroud encodes a short-chain dehydrogenase/reductase that functions in the 'Black Box' of the ecdysteroid biosynthesis pathway *Development,*137(12), 1991-1999.

Nguyễn T. K. T., Kuroda A., Urakawa H., and Yasunaga H., (2017). Dyeing fabrics by using extracts from mulberry branch/trunk 1. Dyeability and fluorescence property. *American Journal of Plant Sciences*, 8(08), 1888–1903.

Niratker C. R., Preeti and Malti., (2015). Antimicrobial activity of leaf extract of Morus indica (Mulberry) from Chhattisgarh. *Asian Journal of Plant Science and Research*, 5(1), 28–31.

Nithya D., Poornima S. M., Pazhani R. M., Gopikrishnan V., Radhakrishnan M., Bhivi D.,

and Balagurunathan R.,(2011). Influence of biofertilizers and irrigation systems for the growth and yield of mulberry plants. *International Journal of Plant, Animal and Environmental Sciences*, 1(3), 93–99.

Ozrenk K., Gazioglu S. R.., Erdinc C., Guleryuz M., and Aykanat A., (2010). Molecular characterization of mulberry germplasm from Eastern Anatolia. *African Journal of Biotechnology*, 9(1), 1–6.

Patwardhan A. Samit R., Amit R., (2014). Molecular Markers in Phylogenetic Studies-A Review. *Journal of Phylogenetics & Evolutionary Biology,* 2(2), 131-140.

Peterson B. K., Weber J. N., Kay E. H., Fisher H. S., and Hoekstra H. E., (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, *7*(5), e37135.

Pillaiyar T., Manickam M., and Namasivayam V., (2017). Skin whitening agents: Medicinal chemistry perspective of tyrosinase inhibitors. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 32(1), 403–425.

Priya S. (2012). Medicinal Values of Mulberry-An Overview. *Journal of Pharmacy Research*, 5(7), 3588–3596.

Qiu F., and Zhang Y., (2019). Metabolic effects of mulberry branch bark powder on diabetic mice based on GC-MS metabolomics approach. *Nutrition and Metabolism*, 16(1), 1–16.

Rahmathulla V. K., (2012). Management of climatic factors for successful silkworm (*Bombyx mori* L.) crop and higher silk production: A review. *Psyche:A Journal of Entomology*, *2*012, 1–12.

Ramesh H. L., Murthy V. N. Y., and Munirajappa (2011).Colchicine induced morphological variation in mulberry variety M 5. *The Bioscan*, 6(1), 115-118.

Rao A. A., Vijayan K., Krubakaran M., Borpujari M. M., and Kamble C. K., (2011). Genetic diversity in mulberry (*Morus* spp.) revealed by isozyme markers. *Journal of Horticultural Science and Biotechnology*, 86(3), 291–297.

Rao M. R. R. D., Jhansilakshmi K., Saraswathi P., Rao, A. A., Ramesh S. R., Borpuzari M. M., and Manjula A., (2013). Scope of pre-breeding in mulberry crop improvement - A review. *Weekly Science Research Journal*, 1(6), 1–18.

Raza A., Razzaq A., Mehmood S. S., Zou X., Zhang X., Lv Y., and Xu J., (2019). Impact of climate change on crops adaptation and strategies to tackle its outcome: A review. *Plants*, 8(2), 34–63.

Rodrigues E. L., Marcelino G., Silva G. T., Figueiredo P. S., Garcez W. S., Corsino J., Guimarães R. de C. A., and Freitas K. de C., (2019). Nutraceutical and medicinal potential of the *Morus* species in metabolic dysfunctions. *International Journal of Molecular Sciences*, 20(2), 301–317.

Rouge K., (2010). Supporting Sericulture Rehabilitation.*FAO Report, February*, 3–4.

Saddul D., Jelan Z. A., Liang J. B., and HalimR. A., (2004). The potential of mulberry (Morus alba) as a fodder crop: The effect of plant maturity on yield, persistence and nutrient composition of plant fractions. *Asian-Australasian Journal of Animal Sciences*, 17(12), 1657–1662.

Saeed B., Baranwal V. K., and Khurana P., (2016). Comparative transcriptomics and comprehensive marker resource development in mulberry. *BMC Genomics*, 17(1), 98–112.

Saha S., Adhikari S., Dey T., and Ghosh P., (2016). RAPD and ISSR based evaluation of genetic stability of micropropagated plantlets of *Morus alba* L. variety S-1. *Meta Gene*,

7, 7–15.

Sanchez M. D., (2002). World distribution and utilization of mulberry, potential for animal feeding. *Proceedings of FAO Electronic Conference on Mulberry for Animal Production.*, 1–10.

Sánchez M. D., (2000). Mulberry: an exceptional forage available almost worldwide. *World Animal Review*, 93(1), 1–21.

Sarkar T., Mogili T., and Sivaprasad V., (2017). Improvement of abiotic stress adaptive traits in mulberry (*Morus* spp.): an update on biotechnological interventions. *3 Biotech*, 7(3), 214–228.

Shang J., Song P., Ma B., Qi X., Zeng Q., Xiang Z., and He N., (2014). Identification of the mulberry genes involved in ethylene biosynthesis and signaling pathways and the expression of MaERF-B2-1 and MaERF-B2-2 in the response to flooding stress. *Functional and Integrative Genomics*, 14(4), 767–777.

Sharma A., Sharma R., and Machii H., (2000). Assessment of genetic diversity in a *Morus* germplasm collection using fluorescence-based AFLP markers. *Theoretical and Applied Genetics*, 101, 1049–1055.

Singh R., Bagachi A., Semwal A., Satinder K., and Bharadwaj A., (2013). Traditional uses, phytochemistry and pharmacology of *Morus alba* Linn.: A review. *Journal of Medicinal Plants Research*, 7(9), 461–469.

Singhal B. K., Khan M. A., Dhar A., Baqual,F. M., and Bindroo B. B., (2010). Approaches to industrial exploitation of Mulberry (Mulberry sp.) fruits. *Journal of Fruit Ornamental Plant Research*, 18(1), 83–99.

Sori W., and Gebreselassie W., (2016). Evaluation of mulberry (Morus spp.) genotypes for

growth, leaf yield and quality traits under Southwest Ethiopian condition. *Journal of Agronomy*, 15(4), 173–178.

Sugiyama M., Katsube T., Koyama A., and Itamura H. (2013). Varietal differences in the flavonol content of mulberry (Morus spp.) leaves and genetic analysis of quercetin 3-(6-malonylglucoside) for component breeding. *Journal of Agricultural and Food Chemistry*, 61(38), 9140–9147.

Tao Y., Wang Y., Yang J., Wang Q., Jiang N., Chu D., Han Y., and Zhou J., (2017). Chemical composition and sensory profiles of mulberry wines as fermented with different Saccharomyces cerevisiae strains. *International Journal of Food Properties*, 20(sup2), 2006–2021.

Tekle Y. I., Grant J. R., Kovner A. M., Townsend, J.P., and Katz L. A.,(2010).Identification of new molecular markers for assembling the eukaryotic tree of life. *Molecular Phylogenetics and Evolution,* 55(3), 1177-1182.

Thabti I., Marzougui N., Elfalleh W., and Ferchichi A., (2011). Antioxidant composition and antioxidant activity of white (*Morus alba* L.), black (*Morus nigra* L.) and red (*Morus rubra* L.) mulberry leaves. *Acta Botanica Gallica*, 158(2), 205–214.

Thabti I., Elfalleh W.,Tlili N.,Ziadi M.,Campos M. G.,Ferchichi A., (2014). Phenols, flavonoids and antioxidant and antibacterial activity of leaves and stem bark of Morus species. *International Journal of Food Properties*.17(4), 842-854

Thaipitakwong T., Numhom S., and Aramwit P., (2018). Mulberry leaves and their potential effects against cardiometabolic risks: A review of chemical compositions, biological properties and clinical efficacy. *Pharmaceutical Biology*, *56*(1), 109–118.

Tikader A., and Dandin S. B., (2005). Biodiversity, Geographical Distribution, Utilization and Conservation of Wild Mulberry. *Caspian Journal of Environmental Sciences*, 3(2),

179–186.

Tikader A., and Kamble C. K., (2008). Mulberry wild species in India and their use in crop improvement – A review. *Australian. Journal of Crop Science*, 2(2), 64–72.

Tikader,A., Vijayan K., Raghunath M. K., Chakroborti S. P., Roy B. N., and Pavankumar, T., (1995). Studies on sexual variation in mulberry (*Morus* spp.). *Euphytica*, *84*(2), 115–120.

Toirambe B., and B Ouattara., (2008). *Morus mesozygia* stapf.IN: Louppe D., Oteng-Amoako A.A and Brink M. (Editors). PROTA (Plant resources of tropical africa// Ressources végétales de l'Afrique tropicale), Wageningen, Netherlands, 1–7.

Tsuduki T., Kikuchi I., Kimura, T., Nakagawa, K., and Miyazawa, T. (2013). Intake of mulberry 1-deoxynojirimycin prevents diet-induced obesity through increases in adiponectin in mice. *Food Chemistry*, 139(1–4), 16–23.

Tuigong D. R., Kipkurgat T., and Madara D., (2015). Mulberry and Silk Production in Kenya. *Journal of Textile Science & Engineering*, 5(6), 201–209.

Ustundag A. O., and Ozdogan M., (2015). Usage possibilities of mulberry leaves in poultry. *Scientific Papers Series D. Animal Science*, LVIII, 170–178.

Venkatesh K. H., and Munirajappa., (2015). Comparative micro-morphological and karyomorphological studies in three mulberry varieties (*Morus* spp.). *Cytologia*, *80*(1), 31–35.

Vavilov N. I. (1926). Studies on the origin of cultivated plants. *Bulletin of applied Botany*. Vol XVI, No. 2

Venkatesh K. R., and Chauhan S., (2008). Mulberry : Life enhancer. *Journal of Medicinal Plants Research*, 2(10), 271–278.

Venkateswarlu M., Urs S. R., Nath B. S., Shashidhar H. E., Maheswaran M., Veeraiah T. M., and Sabitha M. G., (2006). A first genetic linkage map of mulberry (*Morus* spp.) using RAPD, ISSR, and SSR markers and pseudotestcross mapping strategy. *Tree Genetics and Genomes*, 3(1), 15–24.

Vijayan K., Awasthi A. K., Srivastava P. P., and Saratchandra B., (2004). Genetic analysis of Indian mulberry varieties through molecular markers. *Hereditas*, 141(1), 8–14.

Vijayan K., Nair C. V., and Chatterjee S. N., (2005). Molecular characterization of mulberry genetic resources indigenous to India. *Genetic Resources and Crop Evolution*, 52(1), 77–86.

Vijayan K., Raju P. J., Tikader A., and Saratchnadra B., (2014). Biotechnology of mulberry (Morus L.) - A review. *Emirates Journal of Food and Agriculture*, 26(6), 472–496.

Vijayan K., Srivastava P. P., and Awasthi A. K., (2004a). Analysis of phylogenetic relationship among five mulberry (*Morus*) species using molecular markers. *Genome*, *47*(3), 439–448.

Vijayan K., Tikader A., Kar P. K., Srivastava P. P., Awasthi A. K., Thangavelu K., and Saratchandra B.,(2006). Assessment of genetic relationships between wild and cultivated mulberry (*Morus*) species using PCR based markers. *Genetic Resources and Crop Evolution*, 53(5), 873–882.

Vijayan K., (2007). Molecular Markers and their Application in mulberry breeding. *International Journal of Industrial Entomology*, *15*(2).

Vijayan K., (2010). The emerging role of genomic tools in mulberry (*Morus)* genetic improvement. *Tree Genetics and Genomes*, 6(4), 613–625.

Vijayan K.,Srivastava P. P., Raju P. J., and Saratchandra B., (2012). Breeding for higher

productivity in mulberry. *Czech Journal of Genetics and Plant Breeding*, 48(4), 147–156.

Wang C., Yang Y., Peng C., Huang, C, Yang, M., and Chan, K., (2011). Mulberry leaf polyphenols possess antiatherogenesis effect via inhibiting LDL oxidation and foam cell formation. *Journal of Agricultural and Food Chemistry*, 59(5), 1985–1995.

Wang D., Dong Z., Zhang Y., Guo K., Guo P., Zhao P., and Xia Q., (2017). Proteomics provides insight into the interaction between mulberry and silkworm. *Journal of Proteome Research*, 16(7), 2472–2480.

Wang W.,Yang H., Bo Y., Ding S., and Cao B., (2012). Nutrient composition, polyphenolic contents, and in situ protein degradation kinetics of leaves from three mulberry species. *Livestock Science*, *1*46(2–3), 203–206.

Wang Y., Xiang L., Wang C., Tang C., and He X., (2013). Antidiabetic and antioxidant effects and phytochemicals of mulberry fruit (*Morus alba* L.) polyphenol enhanced extract. *PLoS ONE*, 8(7), e71144.

Wang Z., Lin Y., Li T., Dai F., Luo G., Xiao G., and Tang C., (2019). Phenolic profiles and antioxidant capacities of mulberry (Morus atropurpurea Roxb.) juices from different cultivars. *International Journal of Food Properties*, *22*(1), 1340–1352.

Wani S A, Bhat M. A., Malik G. N., Zaki F. A., Mir,M. R., Wani N., and Bhat K. M. (2013). Genetic diversity and relationship assessment among mulberry (Morus spp) genotypes by simple sequence repeat (SSR) marker profile. *African Journal of Biotechnology*, 12(21), 3181–3187.

Wani S.A, Bhat M. A., Zafar B., Ganai M., and Majid N., (2013). Role of molecular markers in silkworm improvement. *International Journal of Recent Scientific Research*, 4(5),

515–523.

Wei H., Liu S., Liao Y., Ma C., Wang D., Tong J., Feng J., Yi T., and Zhu L., (2018). A Systematic review of the medicinal potential of mulberry in treating Diabetes mellitus. *The American Journal of Chinese Medicine*, 46(8), 1743–1770.

YadavP., Garg N., and Kumar S., (2014). Improved shelf stability of mulberry juice by combination of preservatives. *Indian Journal of Natural Products and Resources*, 5(1), 62–66.

Yahiro M. ,(1974). Dormancy and Sprouting of winter buds of a mulberry grown in a greenhouse. *Tropical Farming*, 18(4), 189–193.

Yamanouchi H., Koyama, A., and Machii, H. (2017). Nuclear DNA amounts of mulberries (*Morus* spp.) and related species. *Japan Agricultural Research Quarterly*, 51(4), 299–307.

Yang X., Yang L., and Zheng H., (2010). Hypolipidemic and antioxidant effects of mulberry (*Morus alba* L.) fruit in hyperlipidaemia rats. *Food and Chemical Toxicology*, 48(8–9), 2374–2379.

Yuan Q., and ZhaoL., (2017). The Mulberry (*Morus alba* L.) Fruit - A Review of Characteristic Components and Health Benefits. *Journal of Agricultural and Food Chemistry*, 65(48), 10383–10394.

Zarega N.J.C, Clement W. L., Datwyler S. L., and Weiblen G. D. (2005). Biogeography and divergence times in mulberry family Moraceae. *Molecular phylogenetics and Evolution*, 37(2), 402-416.

Zeng Q., Chen H., Zhang C., Han M., Li T., Qi X., Xiang Z., He N., and Xu C., (2015). Definition of eight mulberry species in the genus morus by internal transcribed spacer-

based phylogeny. *PLoS ONE*, 10(8), 1–13.

Zhang H., Ma Z., Luo X., and Li X., (2018). Effects of Mulberry Fruit (*Morus alba* L.) Consumption on Health Outcomes: A Mini-Review. *Antioxidants*, 7(5), 69.

Zhang S. , Soltis D.,Yang Y., Li D., and Yi T., (2011). Multi-gene analysis provides a well-supported phylogeny of Rosales. *Molecular Phylogenetics and Evolution*, 60(1), 21–28.

Zhang Y., Ren C., Lu G., Cui W., Mu Z., Gao H., and Wang Y., (2014). Purification, characterization and anti-diabetic activity of a polysaccharide from mulberry leaf. *Regulatory Toxicology and Pharmacology*, 70(3), 687–695.

Zhang Y., Ren, C., Lu, G., Mu, Z., Cui, W., Gao, H., and Wang Y., (2014). Anti-diabetic effect of mulberry leaf polysaccharide by inhibiting pancreatic islet cell apoptosis and ameliorating insulin secretory capacity in diabetic rats. *International Immunopharmacology*, 22(1), 248-257.

Zhao W. , Miao X., Zang B., Zhang L., Pan Y., and Huang, Y., (2006). Construction of Fingerprinting and Genetic Diversity of Mulberry Cultivars in China by ISSR Markers. *Acta Genetica Sinica*, 33(9), 851–860.

Zhao W., Zhou Z., Miao X., Zhang Y., Wang S., Huang J., Xiang H., Pan Y., and Huang Y., (2007). A comparison of genetic variation among wild and cultivated Morus Species (Moraceae: *Morus*) as revealed by ISSR and SSR markers. *Biodiversity and Conservation*, 16(2), 275–290.

Chapter 2: Morphological diversity and characterisation of 55 mulberry varieties of Asian Origin

2.1 Introduction

Morphological classification in most plants relies on the study of the form and structures that may include either the shoot system or root system or both. In mulberry, this strategy has been the main form of classification of the species and varieties for a long time revealing the diversity in the Moraceae family of which mulberry belongs creating a subject of continuous great debate and reviews. Currently, 68 species well-recognised for utilisation under the Moraceae family (Datta, 2000).

Emergence of new varieties in mulberry has been enhanced by the predominantly dioecious, heterozygous and outbreeding characteristic of the plant (Awasthi et al., 2004; Venkatesh and Munirajappa, 2015). Furthermore, majority of the varieties grown are diploid in nature however polyploids exist as a result of either natural means or human intervention through breeding (Dai et al., 2015; Yamanouchi et al., 2010).

Development of superior heterozygous breeds that has reduced the chances of inbreeding depression has been enhanced by the co-existence of different varieties in addition to introductions from different regions in same ecological niches. This is due to the inherent characteristics of mulberry for survival arising from ease of adaptation to different environmental conditions ranging from cool to warm conditions (Hotta, 1958; Machii et al., 2000) as such understanding their morphological characteristics and diversity becomes the key driving factor for utility and diversification of products.

Mulberry plant phenotypic characteristics occur as a result of combined effects of their inherent genetic makeup, environmental influence and the interactions effects that may cause the plasticity in diversity (Gray and Gray, 1987) through the intraspecific and interspecific variations while taking it position in adapting to different ecological niches such as Japan.

Japan is considered among the China -Japan gene centre (Vavilov, 1926), currently has over 1500 varieties across several species being conserved and utilised for various purposes (Machii et al., 2001). However, among these species available, *M. latifolia*, *M. alba* and *M. bombycis* are predominantly standing out for sericulture purpose with some varieties of *M. kagayamae* used for either rearing silkworm or for the wood. Morphological characterization of the varieties across the seven species aimed at elucidating the differences among varieties under the genus *Morus* using morphological features.

2.2 Materials and method

2.2.1 Plant material

Evaluated plant materials for 55 mulberry varieties were cultivated and maintained NARO genebank of Japan where their selected quantitative and qualitative measurements on morphoplogical features of *M. bombycis, M. acidosa , M. kagayamae, M. latifolia M. alba*,*M. indica* and *M. rotundiloba* selected varieties (**Table 2**). The above ground mulberry plant features used included the plant shoot/stem height, shoot internode distance, orientation of the plant, 10 cm base diameter, leaf angle, leaves size through; blade length and blade width, base depth, petiole length, length of leaf scar, width of leaf scar, texture of the leaf and buds through ;bud length, bud width, bud thickness, bud orientation. The longest shoot/stem and leaf measurements were done on three repetitive samples of each variety. The leaf data was based on destructive sampling method (where the leaves were sampled from the selected varieties). While the bud's data was collected on five bud samples of each variety during winter when in dormancy and with full features observable. Quantitative measurements were done using a metre rule and data recorded with the leaf scores, mean averages and standard deviations calculated. Measurements on the leaf were based on determined method shown (**Figure 3**).

Table 2. Listing for varieties across seven and one unknown species used in morphological characterization

Description	**Identification**	
***Morus acidosa* Griff**.		
Amami 01	JP NO.	166772
Amami46	JP NO.	186972
Koshikijima 08	JP NO.	166779
Kuroshima Hori	JP NO.	239623
Oohara 2	JP NO.	204005
***Morus alba* L.**		
Akame Kumataka	JP NO.	165686
Aoichi	JP NO.	165737
Hikojirou	JP NO.	165729
Hiroeguwa	JP NO.	165893
Hosoe	JP NO.	165904
Ichinose	JP NO.	165752
Kairyou Nezumi Gaeshi	JP NO.	165775
Nezumigaeshi	JP NO.	165725
***Morus bombycis* Koidz.**		
Aizujyujima	JP NO.	165736
Akajiku	JP NO.	165741
Awamiyasou	JP NO.	165747
Chikubayasou	JP NO.	165869
Chiyozuru	JP NO.	165872
Fuyousou (Tomi)	JP NO.	165901

Higashitani	JP NO.	165728
Hiromaru	JP NO.	165895
Iwase	JP NO.	165688
Kani Wase	JP NO.	166329
Kenmochi	JP NO.	165807
Kikuba	JP NO.	166335
Negoya Takasuke	JP NO.	165890
***Morus indica* L.**		
K-2	JP NO.	166929
S-36	JP NO.	166933
S-54	JP NO.	166934
V-1	JP NO.	166928
***Morus kagayamae* Koidz.**		
Hachijoujima 23	JP NO	239641
Hachijoujima 25	JP NO	239642
Hachijoujima 35	JP NO	239643
Hachijouguwa	JP NO	165726
***Morus latifolia* Poir**		
Aikokuguwa	JP NO	165735
Akamerosou	JP NO	165742
Aoshoudo	JP NO.	165685
Bekkouguwa	JP NO.	165731
Benikawa Rosou	JP NO.	165732
Chounosou	JP NO.	165722
Daishuukaku	JP NO.	165860

Fusoumaru	JP NO.	165900
Homare	JP NO.	165905
Kokusou 21	JP NO.	165813
Ookaraguwa	JP NO.	165761
Ooshimasou	JP NO.	165762
***Morus rotundiloba* Koidz.**		
Bai Poe	JP NO.	166000
Keekai	JP NO.	166009
Noi	JP NO.	166012
Poo	JP NO.	166015
***Morus* sp**		
Enbu	JP NO.	165945
Interspecific hybrids		
Kairyou Ichinose	JP NO.	165773
Shinichinose	JP NO.	165836
Rohachi	JP NO.	165734

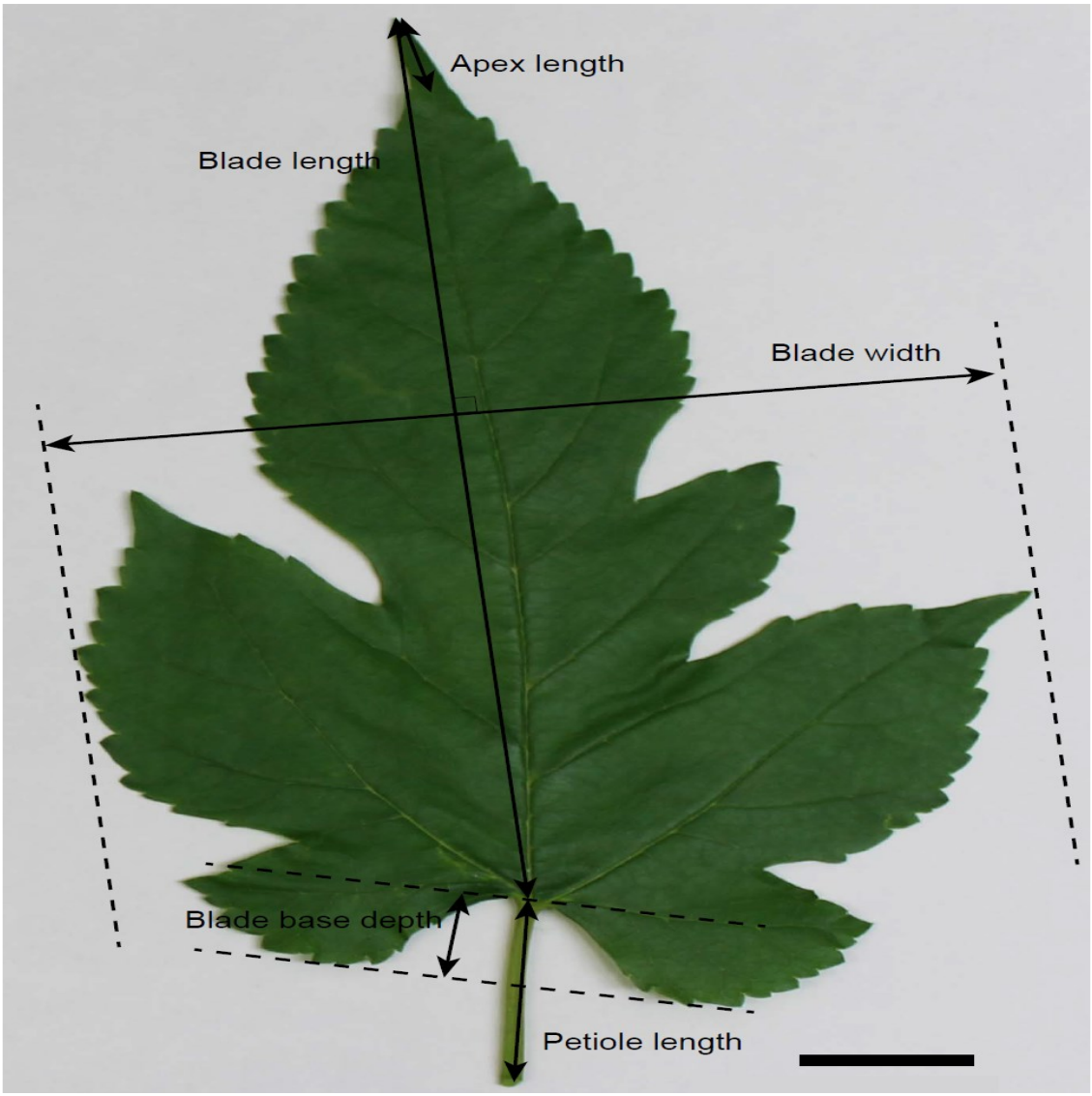

Figure 3. Dimension of the leaf measurements bearing on the mulberry leaf blade

The apex length was measured from the tip of the leaf to slightly where the leaf starts to broaden, the Blade width at the widest point of the leaf perpendicular to the length while the Blade length was measured from the tip of the leaf down to the intersection point with the petiole. The petiole measurements were done from the point of detachment from the branch till the interaction point with the start of the blade length. The blade base depth was measured by the parallel distance existing between the lower edge of the base of the leaf and the intersection point of length and petiole (**Fig. 3**).

2.3 Results

The result revealed diversity among mulberry at variety level and within the same species existing based on the morphological features (**Figure 4-11**). In the blade length and blade width, clear differences were observed between species with higher values observed in 'Enbu', *M. latifolia*, *M. kagayamae* and interspecific hybrids ('Rohachi', 'Shinichinose' and 'Kairyou Ichinose') while smaller values in varieties belonging to *M. acidosa, M. indica* and *M. alba*, species, respectively. However, similar values were observed among some species, making it difficult to isolate the species clearly.

The apex length differed among the species separating them into three groups of long apex (*M. kagayamae*, *M. acidosa*, and *M. rotundiloba*), medium apex group ('Enbu', *M. indica*, *M. bombycis* and interspecific hybrids) and finally the small group category (*M. latifolia* and *M. alba*).

The petiole length also grouped into three categories among the species of long petiole group (*M. latifolia*, *M. kagayamae*, the interspecific hybrids, and *M. alba*), medium petiole group (*M. bombycis*, *M. indica* and 'Enbu') and the short petiole group (*M. rotundiloba* and *M. acidosa*) respectively. Overall, similar values were observed among some species in each feature, making it difficult to isolate the species clearly.

However, the leaf apex stood out as a feature that maybe useful for differentiating some mulberry varieties compared with other features. Among the mulberry varieties, 'Enbu' seemed to be having big leaf blade, medium apex and short petiole compared to other species phenotypes (**Figure 12**). The leaf pictures of the mulberry plants either maintained as potted plants denoted with (asterisk) and those from the field below **(Figure 4-11)** of each variety across the species were taken using a stabilized digital camera(Cannon IXY 600 F model) and the size of scale bar of 2 cm embedded in each jpeg picture file of all the leaves using Image J (Schneider et al., 2012).

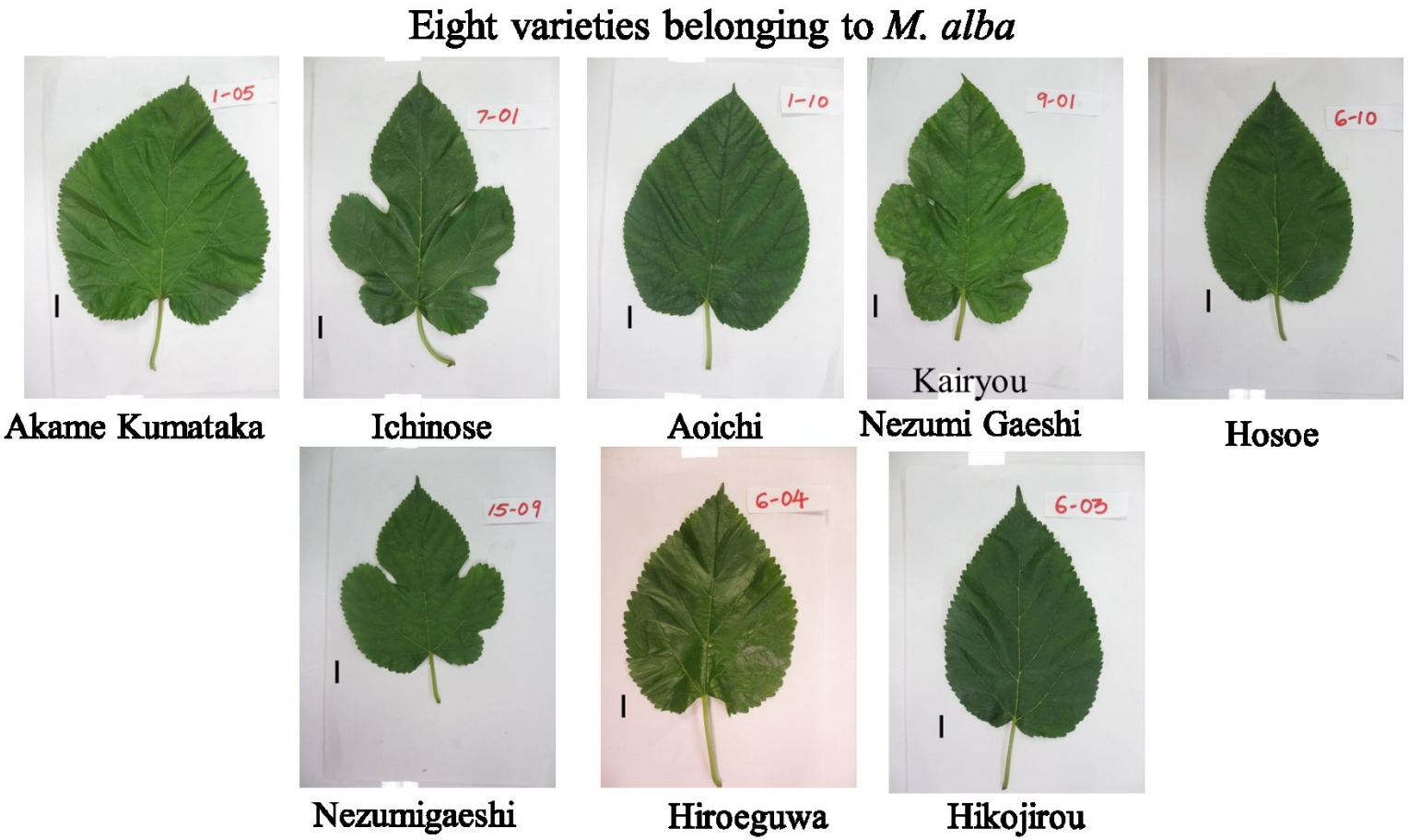

Figure 4. Varieties of *M. alba* evaluated for morphological studies.

The leaves were collected through destructive sampling by detaching them from the shoot and stored in labelled sample bags for measuring and further analysis. Variation occurred in leaf blade sizes, length of apex, number of lobbation as well as the size of petiole within the species.

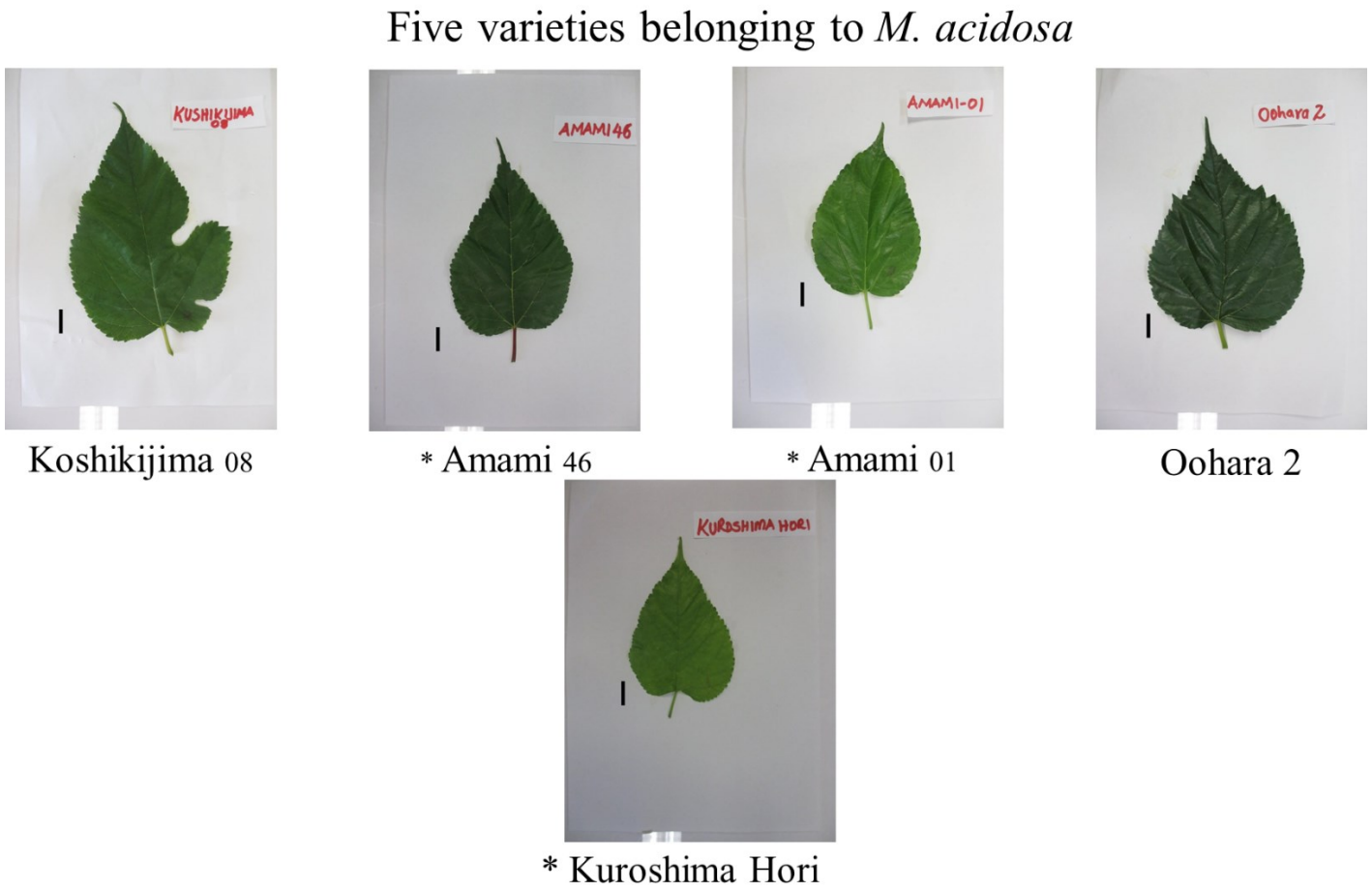

Figure 5. Varieties of *M. acidosa* evaluated for morphological studies.

The leaves of *M. acidosa* have a bigger shallow base and long apex with shorter petiole however, three (*) of the five samples were potted plants grown in greenhouse environment different from the others established in the field.

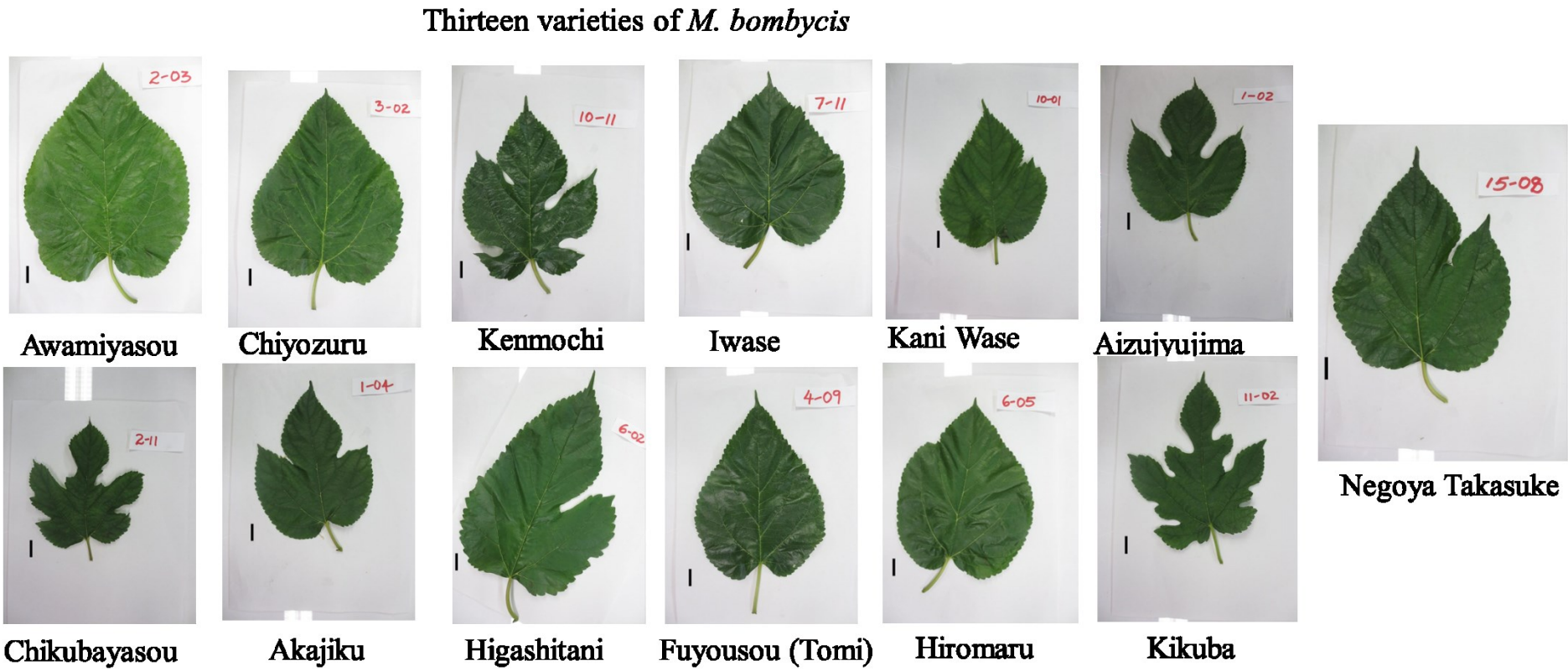

Figure 6. Varieties of *M. bombycis* varieties used for the morphological studies

The varieties have mixed characteristics of different leaf apex, unlobed to multilobate, different pigmentation as well as short petioles.

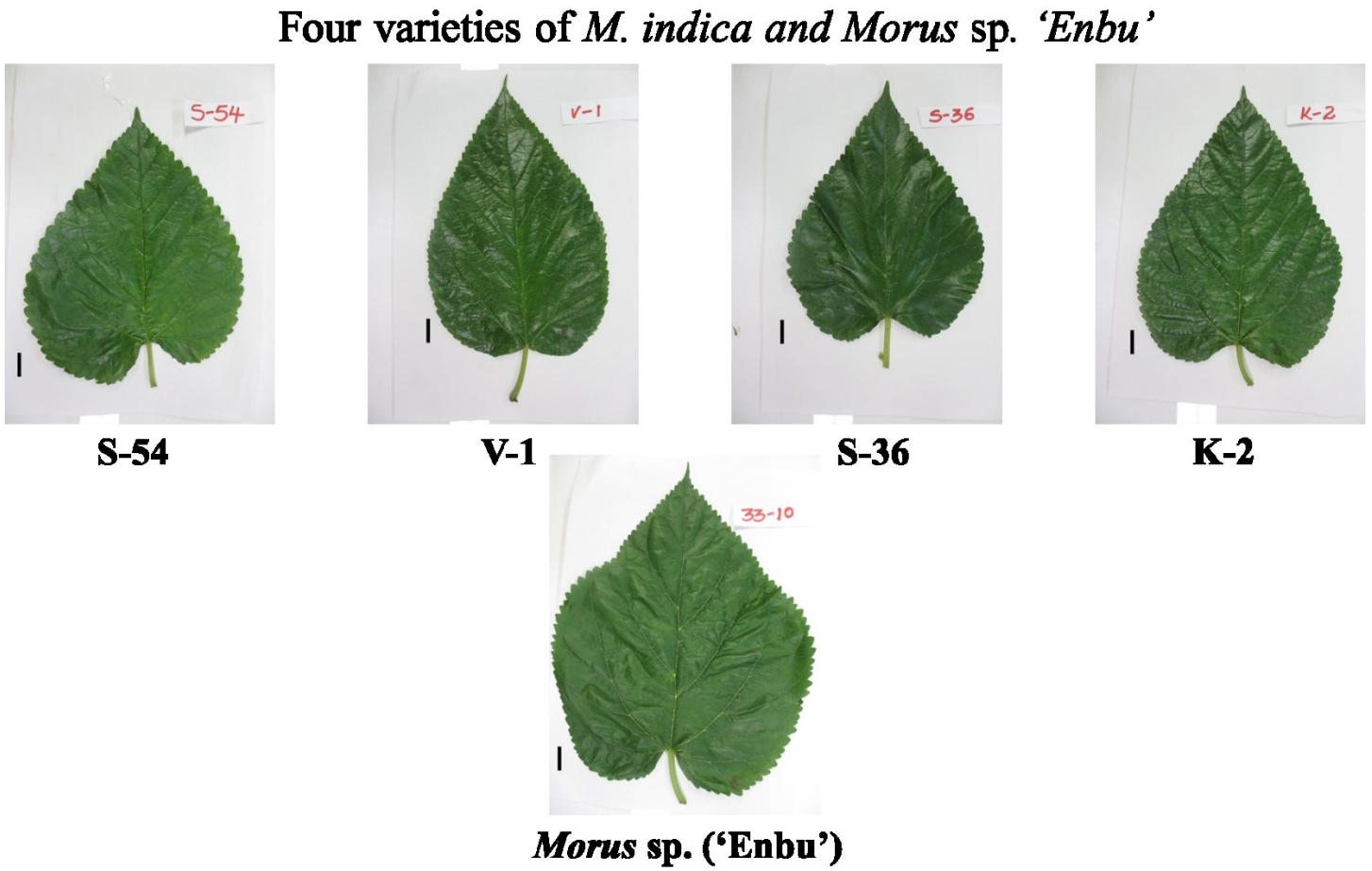

Figure 7. Varieties of *M. indica* and one unspecified *Morus* sp. 'Enbu' used for the morphological studies.

The varieties in *M. indica* are single lobed, have an acute leaf tip with short apex that unifies them and a broad base like the *Morus* sp. 'Enbu'

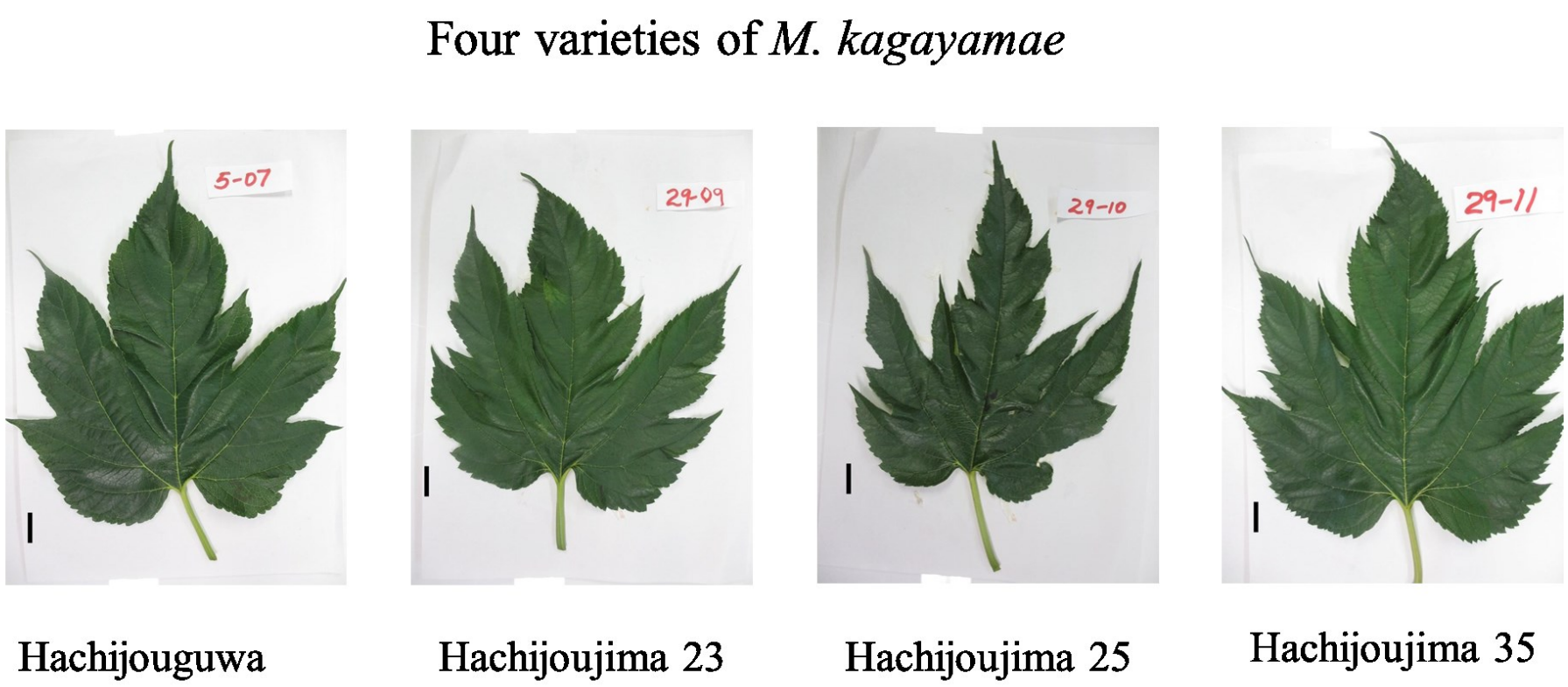

Figure 8. Varieties of *M. kagayamae* used for the morphological studies.

The varieties are darker in pigmentation, multilobed with long apex and short thick petioles compared to other varieties

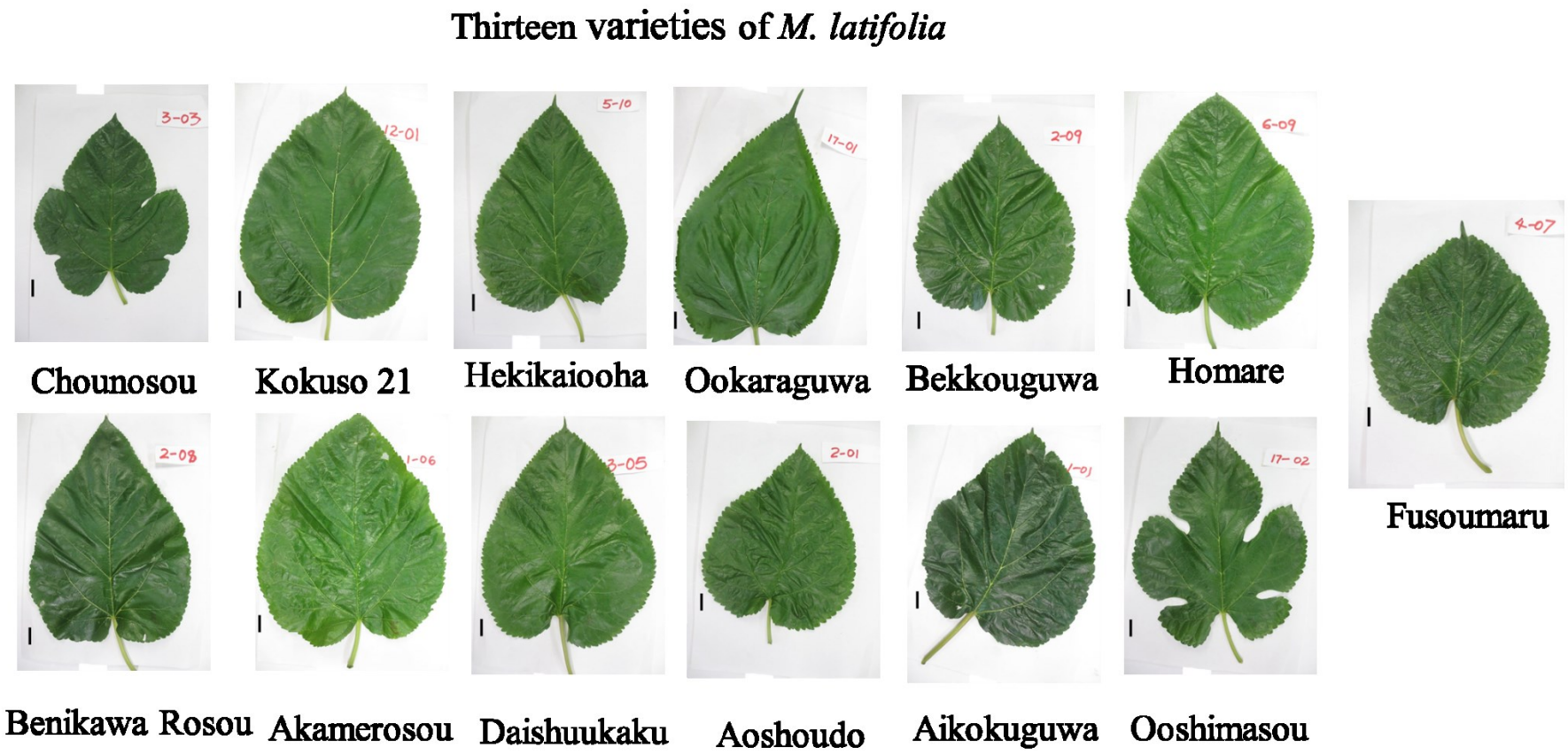

Figure 9. Varieties of *M. latifolia* used for the morphological studies.

The pictures were captured at a scale of 2cm.The varieties have big leaf blades, longer petioles, mixed lobbation and short to rounded apex formation.

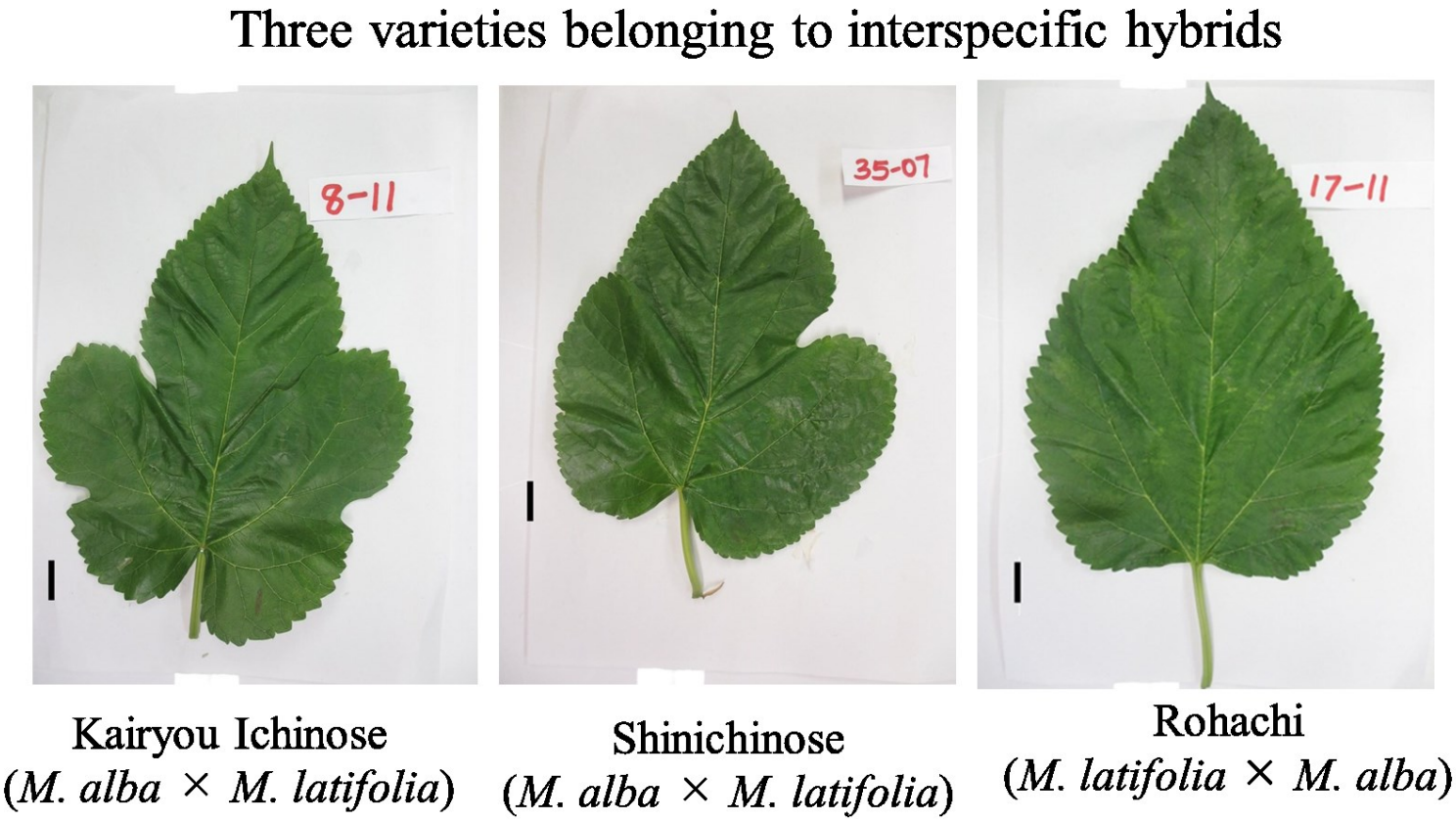

Figure 10. Varieties of interspecific hybrids used for the morphological studies

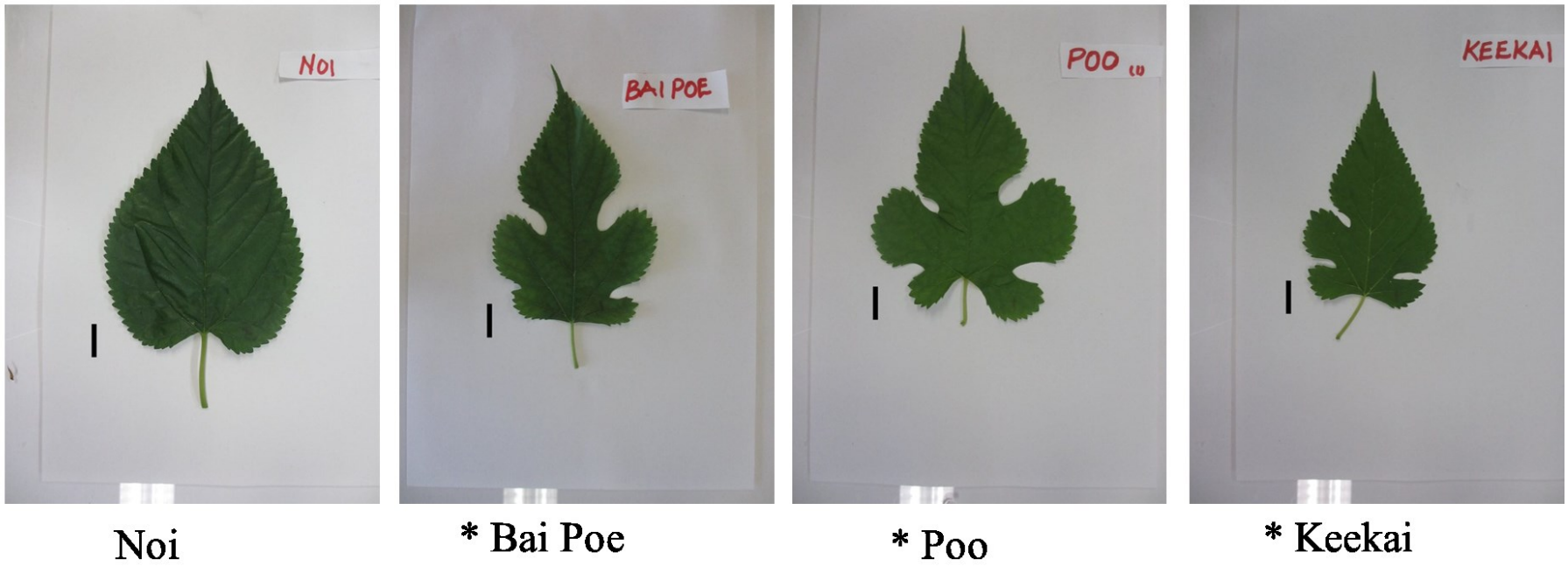

Figure 11. Varieties of *M. rotundiloba* used for the morphological studies.

The varieties have long apex mostly multilobate and smaller in blade size

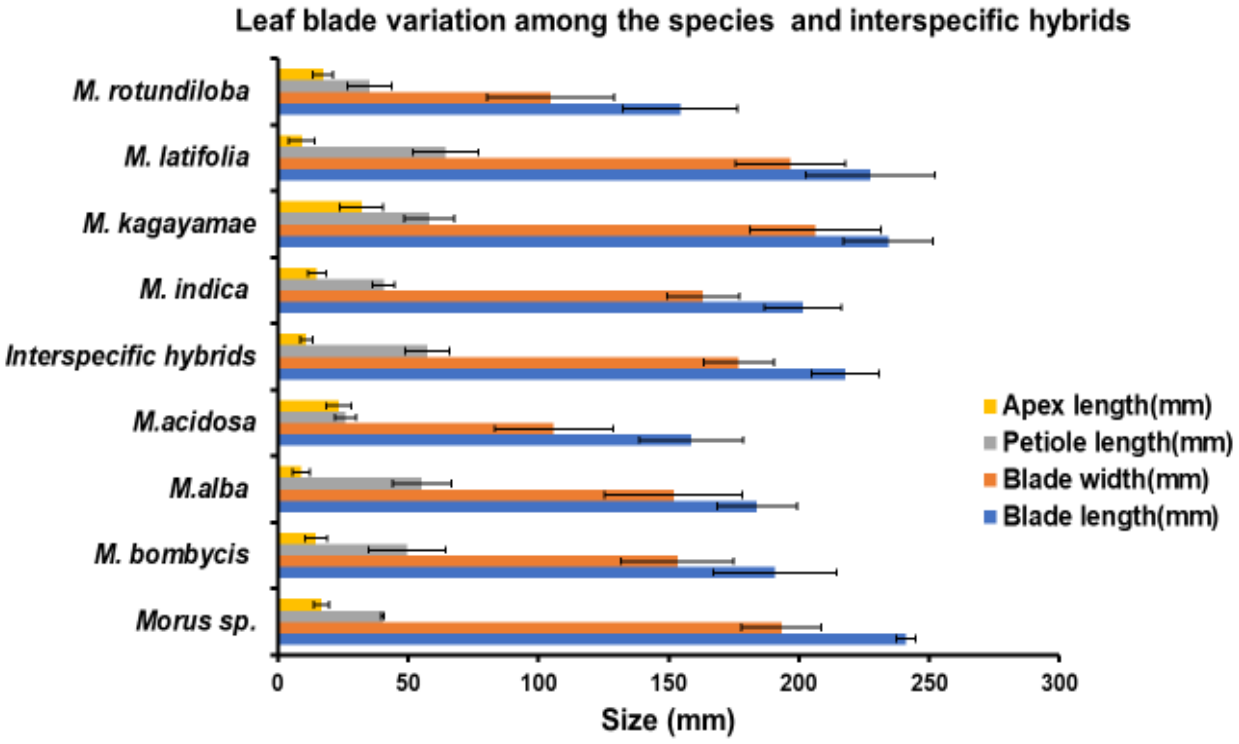

Figure 12. Variation among the mulberry species in blade length, blade width, petiole length and apex length

There is variation in the apex among the species with *M. kagayamae* having the longest petiole while *M. acidosa* the least. The species further can be categorized into three groups based on petiole size of long group *(M. latifolia, M. kagayamae*, the interspecific hybrids, and *M. alba*), medium group (*M. bombycis, M. indica* and *Morus* sp. 'Enbu') and the short group (*M. rotundiloba* and *M. acidosa*) respectively. *Morus* sp., interspecific hybrids *M. kagayamae* and *M. latifolia* have bigger leaf parameters than the indigenous *M. bombycis, M. acidosa*, and introduced *M. rotundiloba*.

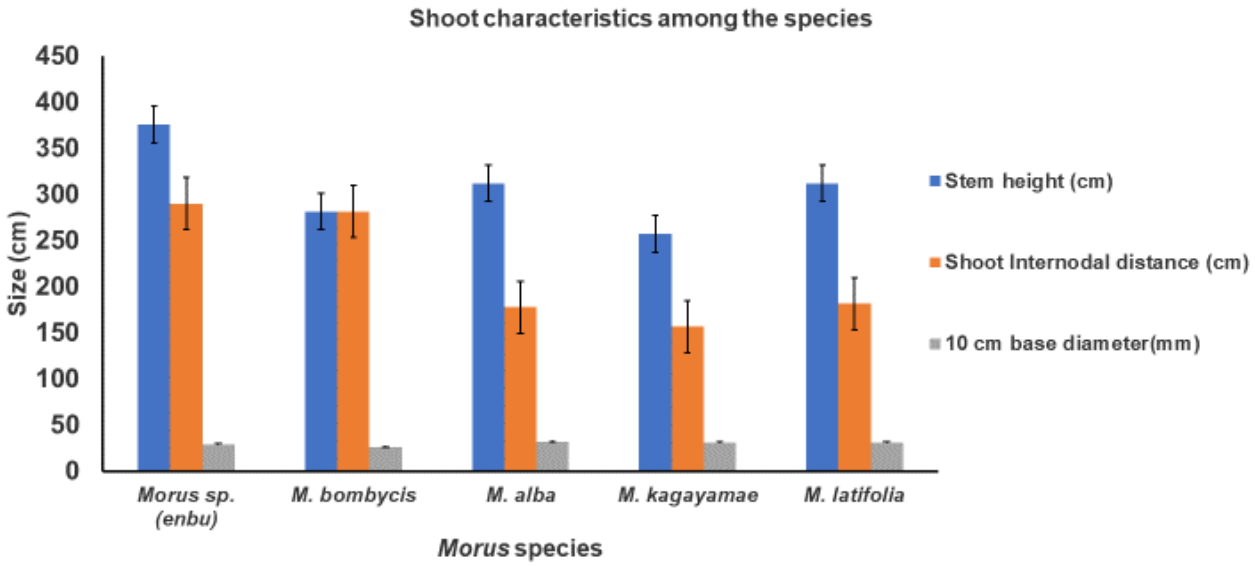

Figure 13. Variation among the stem height, internodal distance and the 10 cm base diameter measurement

The stem height was measured on the tallest shoots on the stool, the internodal distance was measured between five consecutive nodes 50cm below the tip of the shoot while the 10 cm base diameter was measured using a digital vernier calliper at a height 10 cm from the ground in millimeters.

There was observed differences among the height of the shoot of the species however the internodal distances varies from varieties adapted to warmer regions where (*M. alba, M. kagayamae* and *M. latifolia)* having shorter than those in colder areas (*M. bombycis*).

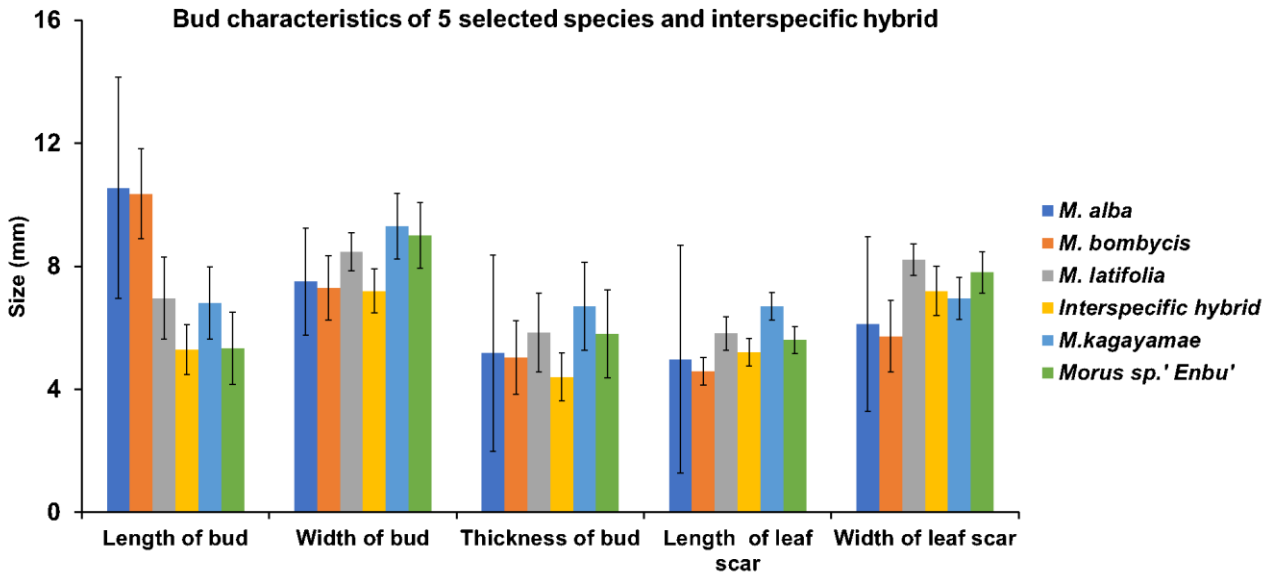

Figure 14. Relationship of buds among the species and interspecific hybrids the widely grown varieties for rearing silkworm (*Bombyx mori*).

The varieties for cold areas have long narrow bud that is thick with small leaf scar to be able to adopt easily to the conditions while those that adopt to the warm temperatures such as *M. latifolia* and *M. alba* have medium to small and wider bud with a big leaf scar.

Table 3.Commonly grown mulberry species of origins in Japan, and their characters

Characters for species identification	**Features**	***M. bombycis***	***M. alba***	***M. lhou* (= *M. latifolia*)**
Shoot characters	Shoot length	Long	Slightly long	Very long
	Shoot thickness	Thin or medium	Thin or medium	Very thick
	Shoot number	Many	Slightly many	Small number
	Secondary shoot No. (Elongated shoots from sprouted axillary buds)	Many	Many	A few
	Shoot straightness	Zigzag	Straight	
	Shoot uprightness	Slightly lying	Erect	Lying
	Colour of lignified shoot	Brown, dark Brown (dark reddish-brown)	Brownish grey, white grey	Brownish grey, grey
	Smoothness of lignified shoot	Rough	Slightly smooth	Smooth
Bud characters	Shape of winter bud	Egg shape, rounded head	Triangle	Triangle
	Size of winter bud	Long and big	Small and short	Small and short
	Colour of winter bud	Reddish brown	Brownish grey, brown	Brownish grey, pale brown
Leaf characters	Size of leaf blade	Large	Medium	Very large
	Shape of leaf blade	Lobed (2, 3,5~many)	Lobed (5), not lobed	Not lobed
	Shape of leaf apex	Caudate, acuminate	Slightly caudate, acute	Roundate, Obtuse
	Shape of leaf blade base	Shallow cuspidate, similar to ear lobe	Deep cuspidate	Deep cuspidate
	Characters of leaf blade surface	Rough, Many hairs on abaxial side	Smooth	Smooth, wrinkled
	Colour of leaf blade	Dark green	Bright green	Bright green
	Glossiness of adaxial surface of leaf blade	No glossiness	Medium	Strong
	Serrate	Sharp serrate (sometimes biserrate)	Crenate	Crenate
	Phyllotaxy	½	2/5	2/5
Flower characters	Style length	Long	Short or less	Short or less

Table 4.Characterization of 40 mulberry varieties using selected winter morphological features of internode, % leaf cover and bud presentation

Species Name	Variety Name	% Leaf cover	Longest stem height (cm)	10cm base diameter (mm)	Internode	Orientation	Bud presentation
M. alba	AKAME KUMATAKA	0	291	38	Slightly zigzag	Slight wavy	Slight protrusion
	AOICHI	0	343	29	Straight	Medium wavy	Slight protrusion
	HIKOJIROU	60	307	26	Zigzag	Straight	No protrusion
	HIROEGUWA	20	350	29	Zigzag	Straight	Slight protrusion
	HOSOE	0	303	30	Straight	Straight	No protrusion
	NEZUMIGAESHI	0	268	21	Straight	Straight	No protrusion
	ICHINOSE	40	307	26	Slightly zigzag	Slight wavy	Slight protrusion
	KAIRYOU NEZUMI GAESHI	80	329	27	Slightly zigzag	Slight wavy	Slight protrusion
M. bombycis	AIZU JYUJIMA	0	192	22	Straight	Strongly wavy	Medium protrusion
	AKAJIKU	0	299	21	Slightly zigzag	Straight	Flat/no protrusion
	CHIKUBAYASOU	0	279	21	Slightly zigzag	Straight	No protrusion
	AWAMIYASOU	40	340	28	Slightly zigzag	Slight wavy	Slight protrusion
	CHIYOZURU	80	332	39	Slightly zigzag	Slight wavy to straight	No protrusion
	FUYOUSOU (TOMI)	60	368	26	Slightly zigzag	Straight	No protrusion
	HIGASHITANI	40	410	37	Medium zigzag	Medium wavy	Slight protrusion
	HIROMARU	0	352	29	Slightly zigzag	Slight wavy	Slight protrusion
	IWASE	40	343	33	Zigzag	Wavy	No protrusion
	KENMOCHI	20	309	28	Slightly zigzag	Very wavy	No protrusion
	NEGOYA TAKASUKE	0	188	19	Slightly zigzag	Wavy	No protrusion
	KANI WASE	0	195	25	Strongly zigzag	straight	No protrusion
	KIKUBA	0	259	24	Zigzag	Slight wavy	No protrusion
M. kagayamae	HACHIJOUGUWA	0	257	31	Slightly zigzag	Straight	No protrusion
	HACHIJOUJIMA 23	20	220	22	Zigzag	Straight	No protrusion

M. latifolia	HACHIJOUJIMA 25	20	216	20	Zigzag	Straight	No protrusion
	HACHIJOUJIMA 35	0	197	19	Zigzag	Straight	No protrusion
	ALKOKUGUWA	20	309	28	Slightly zigzag	Slight wavy	Slight protrusion
	AKAMEROSOU	40	317	36	Slightly zigzag	Medium wavy	Slight protrusion
	BEKKOUGUWA	0	352	29	Slightly zigzag	Slight wavy	Medium protrusion
	BENIKAWA ROSOU	60	329	36	Slightly zigzag	Straight	Medium protrusion
	AOSHOUDO	60	312	32	Slightly zigzag	Medium wavy	No protrusion
	CHOUNOSOU	60	329	30	Slightly zigzag	Straight	No protrusion
	DAISHUUKAKU	0	320	32	Slightly zigzag	Medium wavy	Slight protrusion
	FUSOUMARU	0	361	38	Slightly zigzag	Strongly wavy	Medium protrusion
	HEKIKAIOOHA	20	293	29	Slightly zigzag	Slight wavy	Medium protrusion
	HOMARE	0	293	28	Straight	Straight	Slight protrusion
	KOKUSOU 21	60	346	38	Slightly zigzag	Wavy	Slight protrusion
	OOKARAGUWA	0	275	29	Slightly zigzag	Slight wavy	No protrusion
	OOSHIMASOU	20	384	35	Zigzag	Medium wavy	Slight protrusion
M. latifolia × *M. alba*	ROHACHI	0	254	27	Slightly zigzag	Slight wavy	No protrusion
Morus sp.	ENBU	0	376	29	Zigzag	Slight wavy	Slight protrusion

- % leaf cover is the estimated amount of leaves remaining on the plant during winter
- 10 cm base diameter is the measurement taken at the height of 10 cm form base of shoot/stem

Among the species, some varieties exhibit hardiness during the cold stress as shown by the remaining percentage leaf cover during the cold acclimatization period. Notably, *M. bombycis* (‘Awamiyasou‘, ‘Chiyozuru‘, ‘Fuyousou (Tomi) ‘, ‘Higashitani‘, ‘Iwase‘ and ‘Kenmochi‘) has many of their varieties tolerant to the cold stress followed by *M. latifolia* (‘Bennikawa Rosou‘, ‘Aushoudo‘, ‘Chounosou‘, ‘Hekikaiooiha‘, ‘Kokuso21‘ and ‘Ooshimasou‘) with a few of *M. alba* varieties (‘Hikojirou‘, ‘Hiroeguwa‘, ‘Ichinose‘ and ‘Kairyou Nezumi Gaeshi‘)

2.4 Discussion

Morphological features are heritable but can also be influenced by the interaction of genetic and environments over time as the plant develops and establishes in a given habitat. Most of the morphological characteristics have been known to be unstable and subjective hence not suitable to be used as markers (Vijayan, 2007). Following such, the 55 mulberry varieties studied belonging to species of *M. bombycis, M. acidosa, M. kagayamae, M. latifolia, M. alba*, *M. indica, M. rotundiloba* and unknown *Morus* sp ‘Enbu’ studied gave definite characteristics of resemblance based on the species characterization at the genebank. However, there was intermediate aspects observed among the varieties with regard to similarities which I may attribute to either influence of environment or hybridization that may have occurred among the varieties in their environment since the mulberry plant outcrosses easily. This being the case observed in some of the phenotypes among different species with some having similarities towards *M. latifolia*, we can attribute this as a result of the introduction of *M. alba* and *M. latifolia* species to Japan could have caused effects of hybridization in some of the native species of *M. bombycis* where I noted resemblance of varieties like ‘Awamiyasou‘ and ‘Chiyozuru‘ having some resemblance to some of the *M. latifolia* varieties. With regard to the environmental influence, I further noticed some tendencies of cold tolerances among the species **(Table 4**), where some hardiness ranging between 20%-80% in leaf cover retainance

being shared between the varieties in the species *M. bombycis*, *M. alba,* and *M. latifolia* which could be contributing to the survival of these varieties in cold to snowy areas. The presentation of variation of some of the characteristics such as wide big leaf that initially was a preserve for *M. latifolia* and *M. alba* can be seen cutting across some varieties such as 'Awamiyasou', 'Chiyozuru' and 'Iwase'. These results concur with those of Sharma et al., (2000) where he stated that it was equally difficult to differentiate the varieties easily since most had been introduced and naturalized thus adapting to their regions of introductions.

2.5 References

Awasthi A. K., Nagaraja G. M., Naik G. V., Kanginakudru S., Thangavelu K. and Nagaraju J., (2004). Genetic diversity and relationships in mulberry (genus *Morus*) as revealed by RAPD and ISSR marker assays. *BioMed Central Genetics*,5(1), 1-9.

Dai F., Wang Z., Luo G., and Tang C., (2015). Phenotypic and transcriptomic analyses of autotetraploid and diploid mulberry (*Morus alba* L.). *International Journal of Molecular Sciences,* 16 (9), 22938-22956.

Datta R.K., (2000). Mulberry Cultivation and Utilization in India, in: FAO Electronic Conference on Mulberry For Anımal Productıon (*Morus*-L), 1–24.

Gray E., and Gray R.E., (1987). Leaf Lobation Patterns in Mulberry. *Castanea* 52, 216–224.

Hotta T., (1958). Taxonomical studies on the *Morus* plants and their distributions in Japan and its vicinities. Japanese Society for Promotion of Science, Tokyo.

Machii H, Koyama A., Yamanouchi H., Matsumoto K., Kobayashi S., and Katagiri K., (2001) A list of morphological and agronomic traits of mulberry genetic resources. Miscellaneous Publications of the National Institute of Sericulture and Entomological Science, 29:1-307 ISSB 0915-2679.

Machii H., Koyama A., and Yamanouchi,H., (2000). Mulberry breeding, cultivation and utilization in Japan. *FAO Electronic Conference on Mulberry for AnimalProduction,* 8, 63-72.

Schneider C.A., Rasband W.S., and Eliceiri K.W., (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature. Methods,* 9, 671–675.

Sharma A., Sharma R., Machii H., (2000). Assessment of genetic diversity in a *Morus* germplasm collection using fluorescence-based AFLP markers. *Theoriticaland Applied. Genetics.* 101, 1049–1055.

Vijayan K.,. (2007). Molecular Markers and their Application in mulberry breeding. *International Journal of Industrial Entomology*, *15*(2), 1-12.

Vavilov N. I. (1926). Studies on the origin of cultivated plants. *Bulletin of applied Botany*. Vol XVI, No. 2

Yamanouchi H., Koyama A., Takyu T., and Muramatsu N., (2010) Nuclear DNA amounts in diploid mulberry species (*Morus* sp.). *Journal of Insect Biotechnology and Sericology*, 79(1), 1-8.

Chapter 3: Utilization of double-digest restriction associated DNA sequencing technique for marker discovery and phylogenetic relationship among mulberry varieties

3.1 Introduction

Mulberry varieties exist wildly and have been cultivated for silkworm rearing. The breeding nature and ease of crossing, which results to hybridization, make them more diverse. The diversity leads to a complicated relationship among *Morus* species. Despite there being many species of mulberry, only 68 of them are recognised and recorded for commercial use in the sericulture industry (Datta, 2002).

The heterozygous nature of *Morus* has played a fundamental role in increasing the diversity in mulberry that requires proper identification in order to distinguish them between the different species while trying to understand their evolutionary relationships. In lieu of these, comparisons among the varieties has relied on morphological, anatomical and cytological features to try separate them. Notably, some of these features are not stable to be used as markers for classification of mulberry because they are subject to environmental influence effects thus altering either their phenotypes as well as their genetic makeup (Vijayan, 2007). Furthermore, morphological classification can be tedious, subjective and requires long periods of evaluation and time to be able to have good reliable results. Since the classification of varieties of *Morus* for a long time relied on the morphological characterization, various taxonomists have had to keep on reviewing the different observable traits which may change based on the genotype or effects of interactions with environment. Therefore, it is difficult to classify the many existing varieties among the mulberry species by morphological methods. Although there are many mulberry species identified for sericulture industry for rearing silkworm (Datta, 2002), reclassification of them by other methods in their taxonomy clusters are required. In Japan, *M. alba*, *M. bombycis*, and *M. latifolia* are known as major mulberry

species. Of these, *M. alba* and *M. latifolia* are introduced species from China, while *M. bombycis* is native to Japan (Hotta, 1958). They have contributed enormously towards the growth of silk industry through silkworm rearing, and many Japanese mulberry varieties have been habitually classified into one of them by morphological characteristics of the three major species. However, it is believed that natural hybridizations have occurred among the mulberry species, which can affect correct classification of the many Japanese mulberry varieties based on morphological characteristics.

To decipher the phylogenetic relationships among the many mulberry varieties at species level, several genetic methods have been developed and used: amplified fragment length rapid polymorphism (AFLP) method (Chumchuen and Kanekatsu, 2011, Sharma et al., 2000), internal transcribed spacer (ITS) method (Nepal, 2008; Nepal and Ferguson, 2012; Zeng et al., 2015), simple sequence repeats (SSR) method (Zhao et al., 2007; Mathithumilan et al., 2013; Mathi Thumilan et al., 2016), rapid amplified polymorphism DNA (RAPD) method (Awasthi et al.,2004), and inter simple sequence repeats (ISSR) method (Awasthi et al.,2004; Ipek et al.,2012; Vijayan et al., 2004). However, it is still difficult for these methods to completely decipher the phylogenetic relationships. This may be due to lack of enough genetic information (e.g. higher number of markers) in the methods which may be potentially required for deciphering the relationships. Furthermore, Vijayan, (2010) found that some of these methods such as RAPD and ISSR may be unreliable with regard to reproducibility.

Therefore, genetic methods that can provide genome-wide scale genetic information with high reproducibility can be a key to solve the problems. One such method is the use of double digest restriction site associated DNA method (ddRAD-seq) that has potential to generate many single nucleotide polymorphism (SNPs) in genome-wide scale from a diverse population of varieties among different species (Peterson et al., 2012). The usefulness of the

method for marker development have been verified in several crops (Emerson et al., 2010; Barchi et al., 2011; Scaglione et al., 2012).

To investigate the relationships among 56 mulberry varieties, which consist of 53 of the 56 mulberry varieties evaluated in the previous chapter and three additional varieties, both the phylogenetic analysis using genome-wide SNPs by the ddRAD-seq method and selected morphological characteristics of the 56 mulberry varieties were performed.

3.2 Materials and methods

3.2.1 Plant materials

Leaves of the 56 mulberry varieties (diploid) shown in **Table 5** were sampled at the NARO Genebank in Tsukuba and used for further investigation in the laboratory. The 56 mulberry varieties belong to seven known and one unknown mulberry species which consist of *M. latifolia* (13 varieties), *M. bombycis* (13 varieties), *M. alba (*8 varieties*), M. acidosa* (6 varieties), *M. indica* (4 varieties*), M. kagayamae* (5 varieties), *M. rotundiloba* (4 varieties), interspecific hybrids (3 varieties) and *Morus* sp. ('Enbu'). Of these, place of the origin in Japan is known for 28 mulberry varieties (**Fig. 15**). The choice of diploid varieties was informed from the point of view that most varieties derived from wild individuals, especially from the native ones such as *M. acidosa* and *M. kagayamae* together with the introduced varieties, are considered to be basically diploid and therefore using diploid varieties was suitable for adjusting experimental conditions. Each variety of mulberry had been maintained through clonal vegetative propagation to maintain their characteristics similar to most woody crops. In Asian continent, Japan has the highest variety numbers of *M. bombycis* (583), *M. kagayamae* (23) and *M. rotundiloba* (24) despite having been introduced. India has the highest gene pool for *M. indica* (350) while China has the highest gene pool for *M. latifolia* (750) and *M. alba*

(762), (Vijayan et al., 2012). The mulberry species of *M. bombycis, M. acidosa*, and *M. kagayamae* are native to Japan while the two species, *M. latifolia* and *M. alba,* are known to be introduced from China. *Morus rotundiloba and M. indica* are species developed in India (Datta, 2002) and in Thailand, respectively. The mulberry varieties of *M. bombycis*, *M. alba* and *M. latifolia* used in this study are breed varieties or indigenous varieties in Japan with one exception of 'Benikawa Rosou' (*M. latifolia*) whose origin is from China. On the other hand, the *M. kagayamae* and *M. acidosa* varieties, which grow in different islands in Japan, were clonally-derived from wild individuals. 'Enbu' and the *M. rotundiloba* varieties are indigenous varieties in Kenya and Thailand, respectively. Whilst, the *M. indica* varieties are breed varieties in India. For each variety, name of species and its origin are basically described based on registered information at NARO Genebank. However, there are exceptions for some varieties described below. The species name of 'Enbu' is denoted as *Morus* sp. (unspecified mulberry species) because its morphology is apparently different from that of *M. mesozygia* ('Enbu' is registered as this mulberry species at NARO Genebank) and therefore actual species is currently unknown (**Table 5**). Furthermore, 'Rohachi' (registered as *M. latifolia*), 'Kairyou Ichinose' and 'Shinichinose' (registered as *M. alba*) were denoted as interspecific hybrids because these varieties are known to be derived from cross of two species (*M. alba* and *M. latifolia*) as shown in **Table 5**.

Table 5. The varieties of mulberry varieties analysed in this study

Variety	JP number	Species	Place of origin
Amami 01	166772	*M. acidosa*	Amami Oshima Island, Japan
Amami 46	186972	*M. acidosa*	Tokunoshima Island, Japan
Koshikijima 08	166779	*M. acidosa*	Koshikijima Islands, Japan
Kuroshima Hori	239623	*M. acidosa*	Kuroshima Island, Japan *1
Oohara 2	204005	*M. acidosa*	Taketomi Island, Japan
Akame Kumataka	165686	*M. alba*	Tokyo Pref., Japan
Aoichi	165737	*M. alba*	Shizuoka Pref., Japan
Hikojirou	165729	*M. alba*	Shiga Pref., Japan
Hiroeguwa	165893	*M. alba*	Japan (unknown Pref.)
Hosoe	165904	*M. alba*	Shiga Pref., Japan
Ichinose	165752	*M. alba*	Yamanashi Pref., Japan
Kairyou Nezumi Gaeshi	165775	*M. alba*	Kumamoto Pref., Japan
Nezumigaeshi	165725	*M. alba*	Nagano Pref., Japan
Aizujyujima	165736	*M. bombycis*	Fukushima Pref., Japan *2
Akajiku	165741	*M. bombycis*	Japan (unknown Pref.)
Awamiyasou	165747	*M. bombycis*	Japan (unknown Pref.)
Chikubayasou	165869	*M. bombycis*	Japan (unknown Pref.)

Chiyozuru	165872	*M. bombycis*	Japan (unknown Pref.)
Fuyousou (Tomi)	165901	*M. bombycis*	Japan (unknown Pref.)
Higashitani	165728	*M. bombycis*	Niigata Pref., Japan
Hiromaru	165895	*M. bombycis*	Tokyo Pref., Japan
Iwase	165688	*M. bombycis*	Shizuoka Pref., Japan
Kani Wase	166329	*M. bombycis*	Japan (unknown Pref.)
Kenmochi	165807	*M. bombycis*	Niigata Pref., Japan
Kikuba	166335	*M. bombycis*	Yamanashi Pref., Japan
Negoya Takasuke	165890	*M. bombycis*	Niigata Pref., Japan
K-2	166929	*M. indica*	India
S-36	166933	*M. indica*	India
S-54	166934	*M. indica*	India
V-1	166928	*M. indica*	India
Hachijouguwa	165726	*M. kagayamae*	Japan (unknown Pref.)
Hachijoujima 23	239641	*M. kagayamae*	Hachijo Island, Japan
Hachijoujima 25	239642	*M. kagayamae*	Hachijo Island, Japan
Hachijoujima 35	239643	*M. kagayamae*	Hachijo Island, Japan
Miyakejima 18	N/A	*M. kagayamae*	Miyake Island, Japan
Aikokuguwa	165735	*M. latifolia*	Kyoto Pref., Japan

Akamerosou	165742	*M. latifolia*	Ehime Pref., Japan
Aoshoudo	165685	*M. latifolia*	Japan (unknown Pref.)
Bekkouguwa	165731	*M. latifolia*	Hokkaido Pref., Japan
Benikawa Rosou	165732	*M. latifolia*	China
Chounosou	165722	*M. latifolia*	Kochi Pref., Japan
Daishuukaku	165860	*M. latifolia*	Japan (unknown Pref.)
Fusoumaru	165900	*M. latifolia*	Saitama Pref., Japan
Hekikaiooha	165730	*M. latifolia*	Aichi Pref., Japan
Homare	165905	*M. latifolia*	Japan (unknown Pref.)
Kokusou 21	165813	*M. latifolia*	N/A
Ookaraguwa	165761	*M. latifolia*	Japan (unknown Pref.)
Ooshimasou	165762	*M. latifolia*	Gunma Pref., Japan
Bai Poe	166000	*M. rotundiloba* *4	Thailand
Keekai	166009	*M. rotundiloba* *4	Thailand
Noi	166012	*M. rotundiloba* *4	Thailand
Poo	166015	*M. rotundiloba* *4	Thailand
Enbu	165945	*Morus* sp. *5	Africa
Kairyou Ichinose	165775	*M. alba* × *M. latifolia* *6	N/A

Shinichinose	165836	M. alba × M. latifolia *6	N/A
Rohachi	165734	*M. latifolia* × *M. alba* *7	N/A

*1 The place of origin of ‘Kuroshima Hori’ in Japan was confirmed to be Kuroshima Island based on the report of its collector.

*2 The place of origin of ‘Aizujyujima’ in Japan is denoted as Fukushima Pref based on high similarity of morphological features and the variety names between ‘Aizujyujima’ and another *M. bombycis* variety ‘Jujima’ originated from Aizu region, which is cold and sonwy region, in Fukushima Pref.

*3 Datta, 2002

*4 Incorrect species name of *M. rotunbiloba* in NARO Genebank was corrected as *M. rotundiloba*.

*5 The species name of ‘Enbu’ is denoted as *Morus* sp. (unspecified mulberry species) because its morphology is apparently different from that of *M. mesozygia* (‘Enbu’ is registered as this mulberry species at NARO Genebank) and therefore actual species is currently unknown

*6 ‘Kairyou Ichinose’ and ‘Shinichinose’ were denoted as interspecific hybrids (*M. alba* × *M. latifolia*) because these varieties are known to be derived from the cross of *M. alba* and *M. latifolia* species.

*7 ‘Rohachi’ was denoted as an interspecific hybrid (*M. latifolia* × *M. alba*) because the variety is known to be derived from the cross of *M. latifolia* and *M. alba* species.

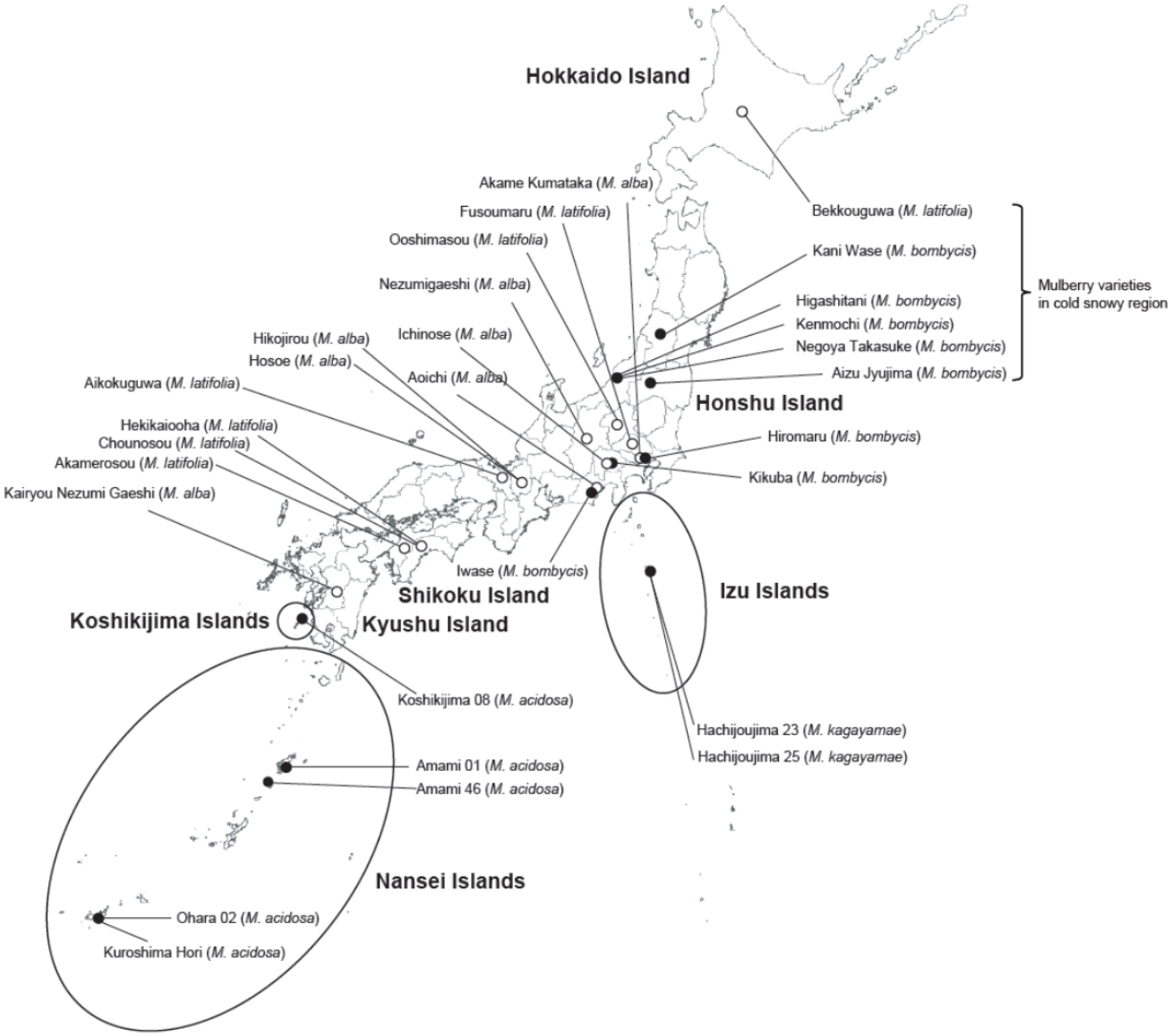

Figure 15. Mapping of 28 of the 56 mulberry varieties to known origin of collection in Japan based on descriptions in NARO Genebank.

Places of origin of 28 mulberry varieties in **Table 5** were summarized according to the description registered at NARO Genebank. The place of the origin of other Japanese varieties were unknown. The places of the origin of six mulberry varieties, which are shown in the upper

right in this figure, are located in a cold snowy region which is shown on the web site of the ministry of the environment of Japan (https://www.env.go.jp/en/nature/npr/wetland/vegetation.html). The Japan map was created using Geospatial Information Authority of Japan website (https://maps.gsi.go.jp/).

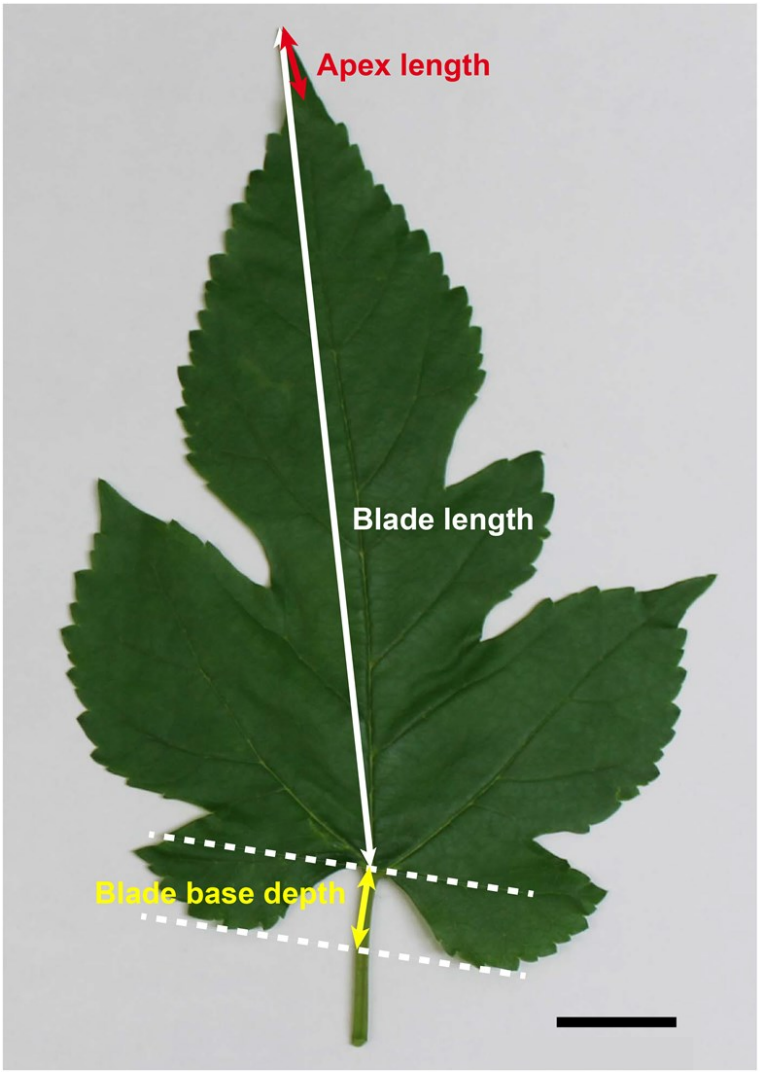

Figure 16. Measurement targeted regions for of morphological characteristics in mulberry leaves

The apex length, blade length, and blade base depth in mulberry leaves were measured using a ruler in mm for 54 varieties evaluated in the phylogenetic analysis. The values were used for calculating and evaluating the leaf tip ratio ("apex length / blade length") and leaf bottom ratio ("blade base depth / blade length").

3.2.2 DNA extraction and ddRAD-seq analysis

Using each of the 56 individual samples of immature shoot tips selected, DNA extraction was performed using Doyle and Doyle method (Doyle and Doyle, 1987) with cetyltrimethyl ammonium bromide (CTAB) modification. Construction of the ddRAD-seq library was performed using the extracted genomic DNA based on the protocol of the ddRAD-seq (Peterson et al., 2012) where double digestion of the DNA samples was performed using two restriction enzymes (EcoRI and MspI). Ligation with unique pair of inline barcodes to each of the EcoRI and MspI associated DNA samples was done for identifying each mulberry variety. Out-sourcing of sequencing of the library was then done in a single lane using Illumina Hiseq 2500 generating 101 bp paired-end reads by Macrogen Japan Corp. The sequenced raw reads were then processed using modules in STACKS version 1.44 (Catchen et al.,2013). Into the STACKS module, four different processes were conducted that involved demultiplexing and cleaning process using "process_radtags" module multiplexing the raw reads to each sample and then removing the lower quality bases, followed by building RAD-loci using the cleaned reads using "ustacks''module with options of -m = 3 and -M = 3, assembling the created RAD-loci using "cstacks" module with an option of -n = 2, and finally searching and linking individual loci using "sstacks" module. "populations" module with -phylip option was then conducted to calculate genotype and extract available homozygous and heterozygous SNPs among the 56 varieties. The nucleotide sequences of the concatenated homozygous SNPs were generated for the phylogenetic analysis.

3.2.3 Reconstruction of phylogenetic tree

The nucleotide sequences of the concatenated homozygous SNPs of the 56 varieties except for the two varieties of 'Hachijoujima 35' and 'Miyakejima 18' (removed due to lack of sufficient SNP data) were used for the phylogenetic analysis. Construction of the neighbour-

joining (NJ) phylogenetic tree (Saitou and Nei, 1987) using the sequence data was performed by MEGA X program (Kumar et al., 2018) using 1000 replicates of bootstrap support where all the unclear positions were removed for each sequence pair. For clarity and improvement of the neighbour joining phylogenetic tree, iTOL v4 (Letunic and Bork, 2019) was used and the bootstrap values were included next to the branches.

3.2.4 Measurement of morphological characteristics

In order to evaluate relationship between the structure of phylogenetic tree and morphological characteristics in the 54 mulberry varieties, some morphological characteristics of mulberry leaves were evaluated for the 54 varieties. Apex length, blade length, and base depth length in mulberry leaves (**Fig. 16**) were measured by a ruler using three leaves selected from five or more leaves, and then leaf tip ratio (LTR) and leaf bottom ratio (LBR) were calculated for each variety. The LTR was calculated by dividing the apex length by the blade length. Similarly, the LBR was calculated by dividing the blade base depth by the blade length. The reason why the LTR and LBR were selected for this analysis is because low blade base depth and high apex length have been observed as typical morphological characteristics in *M. bombycis* (Minamizawa, 1984). Five or more healthy leaves were collected from the field and/or pot for each variety where leaves at 20-50 cm from top of vigorously growing shoots were selected in the field. Three leaves used in the measurement were then narrowed down as follows: (1) damaged leaves and abnormal leaves were removed (if any), (2) a leaf of the lowest blade length and a leaf of the highest blade length were removed in turn if extra leaves remained, (3) three leaves were randomly selected if extra leaves still remained. Pictorial images of one representative leaf were extracted for each mulberry variety using a Canon IXY 600F digital camera. A scale bar (2 cm) was embedded in each picture file (jpeg format) by ImageJ (Schneider et al., 2012).

3.2.5 Evaluation of the genome-wide distribution of the SNPs

RAD loci including SNPs were mapped to the genome sequences of *M. notabilis* (He et al., 2013) in MorusDB (Li et al., 2014) by BLASTN search (threshold e-value is 1e-5). The distribution of mapped RAD loci on the 29 longest scaffolds (more than 1 Mbp) were evaluated to ascertain if they were randomly and genome-widely distributed in the whole genome sequences.

3.3 Results

3.3.1 Identification of SNPs among mulberry varieties by ddRAD-seq

In total, 463.3 million raw reads were sequenced for the 56 mulberry varieties in **Table 5.** Cleaning process of the reads through demultiplexing of each variety resulted in 425.0 million clean reads used for further analysis. The Stacks modules identified 192,414 SNP markers, located in 684,223 RAD loci, among the 56 mulberry varieties. Furthermore, SNP markers genotyped in more than 49 varieties were searched, and 47,839 SNP markers in 15,605 loci were extracted. Of these, 2,229 homozygous SNP markers located in 1,855 loci were identified by "populations" module in 54 of the 56 mulberry varieties ('Hachijoujima 35' and 'Miyakejima 18' were not evaluated due to lack of enough SNP data).

3.3.2 Reconstruction of the phylogenetic tree

A NJ phylogenetic tree of the 54 mulberry varieties was reconstructed using the nucleotide sequences of the concatenated 2,229 homozygous SNP sites. (**Fig. 17** and **Fig. 18**). In the tree, two species native to Japan, *M. acidosa* and *M. kagayamae,* formed clear monophyletic clades with 100% bootstrap support, respectively. *M. rotundiloba,* which is native to Thailand, also formed clear monophyletic clade with 97% bootstrap support. Longer

branch lengths were observed in these varieties compared with others (**Fig. 18**). In **Fig. 17** and **Fig. 18**, C, K, and R denotes each monophyletic clade, respectively.

M. bombycis is another species native to Japan. Although no clear monophyletic clade of single species were formed in the *M. bombycis* varieties, two monophyletic clades were formed with *M. kagayamae* (a clade denoted as BK with bootstrap support 84% in **Fig. 17** and **18**) and with *M. acidosa* (a clade denoted as BC with bootstrap support 66% in **Fig. 17** and **18)**, respectively. Four *M. bombycis* varieties (denoted as B2 in **Fig. 17** and **18)** belong to the clade BK while two *M. bombycis* varieties (denoted as B1 in **Fig. 17** and **18**) belong to the clade BC. Two *M. bombycis* varieties, 'Iwase' and 'Kenmochi', were classified close to the clade BK, however, they were not monophyletic due to poor bootstrap support. The remaining five *M. bombycis* varieties, which consist of 'Awamiyasou', 'Chiyozuru', 'Fuyousou (Tomi)', 'Hiromaru', and 'Higashitani', formed non-monophyletic clades including *M. alba* and/or *M. latifolia* varieties with poor bootstrap support less than 50% except for a small monophyletic clade which consists of two varieties (denoted as AB in **Fig. 17 and 18**). In summary, the *M. bombycis* varieties were roughly classified into the three groups**.**

M. indica is a species of Indian origin. Two different small clades were formed with the *M. indica* varieties. Each clade consists of two *M. indica* varieties (denoted as I1 with bootstrap support 58% and I2 with that of 97% in **Fig. 17** and **18**). Interestingly, 'Enbu', a Kenyan variety formed a monophyletic clade with the clade I1 with bootstrap support of 88% (denoted as IS in **Fig. 17** and **18**). Similarly, the clade I2 formed another monophyletic clade (denoted as ILR with 52% bootstrap support in **Fig. 17** and **18**) including the clade R of *M. rotundiloba* varieties and one *M. latifolia* variety 'Okaraguwa'.

No monophyletic clades were formed among the varieties of *M. alba* which is one of the introduced species from China to Japan (**Fig. 17** and **18**). On the other hand, two small

monophyletic clades were formed among the varieties of *M. latifolia* which is another introduced species from China. The two small monophyletic clades consist of two varieties, respectively, and are denoted as L1 (bootstrap support of 56%) and L2 (bootstrap support of 75%) in **Fig. 17** and **18**. However, no larger clades were formed among the varieties of *M. latifolia*. Furthermore, no monophyletic clades were formed among the varieties of the two introduced species from China.

Regarding the varieties of interspecific hybrids between the two introduced species from China, only 'Shinichinose' (*M. alba* × *M. latifolia*), which is one of the three varieties of interspecific hybrids, formed a monophyletic clade (bootstrap support of 55%) with 'Kokusou 21' (*M. latifolia*) which is a paternal parent of 'Shinichinose' (**Fig. 17** and **18**). However, 'Ichinose' (*M. alba*), which is a maternal parent of the two varieties of interspecific hybrids ('Shinichinose' and 'Kairyou Ichinose' (*M. alba* × *M. latifolia*)), formed no monophyletic clades among them. 'Shirome Rosou', which is a paternal parent of 'Kairyou Ichinose' was not evaluated in this study.

3.3.3 Measurement of morphological characteristics

The two rational numbers of the leaves, LTR and LBR, were calculated for the 55 mulberry varieties based on the three morphological characteristics of the leaves in **Fig. 16** (**Table 6**). In **Fig. 17**, two color gradient bars represent relative level of the mean values of the LTR and LBR among the 54 varieties in the NJ-tree, respectively, where lighter green/red corresponds to higher/lower values. A clear trend is observed in the mulberry varieties in the clear monophyletic clades (C, K, and R) showing high (light green) LTR values (**Fig. 17**). Similar trend is also observed in the two monophyletic clades of *M. bombycis* (B1 and B2), which belong to larger monophyletic clades (BK and BC) respectively, showing relatively high

(green) LTR values compared with other mulberry varieties in nonmonophyletic clades (**Fig. 17**). An opposite trend is observed in the varieties of *M. alba*, *M. latifolia*, their interspecific hybrids and some varieties of *M. bombycis*, which formed nonmonophyletic clades, showing relatively low LTR (red) (**Fig. 17**). On the other hand, similar clear trends were not observed for the LBR. Although clearly low LBR values (light red) were consistently observed among the varieties of *M. acidosa* in the clade C which is one of the clear monophyletic clades in the NJ-tree, no such clear trends were observed in the remaining clear monophyletic clades of *M. rotundiloba* (R) and *M. kagayamae* (K) (**Fig. 17**). The same is also true with the two monophyletic clades of *M. bombycis* (B1 and B2). In the mulberry varieties in the nonmonophyletic clades, the LBR was relatively high in some varieties of *M. alba*, *M. latifolia*, their interspecific hybrids, and *M. bombycis*. For example, relatively high (green) LBR values were observed in 'Ichinose' and 'Akame Kumataka' (*M. alba*), 'Bekkouguwa' and 'Ooshimasou' (*M. latifolia*), 'Awamiyasou' and 'Chiyozuru' (*M. bombycis*), and 'Kairyou Ichinose' and 'Shinichinose' (interspecific hybrids). However, on the other hand, the LBR was relatively low in some varieties of the mulberry species. For example, relatively low (red) LBR values were observed in 'Hosoe' (*M. alba*), 'Aikokuguwa' (*M. latifolia*), 'Higashitani' (*M. bombycis)*, and 'Rohashi' (interspecific hybrids). Therefore, the LBR values among the mulberry varieties in the nonmonophyletic clades were also not consistent.

3.3.4 Relationship between the deep branching and the morphological data of the varieties.

The deep branching of the phylogeny was observed in native isolated species of *M acidosa* and *M. kagayamae* followed by introduced *M. rotundiloba* from Thailand. These varieties are favoured by the both natural isolation mechanisms in their growing niche as well as the weather that could have prevented gene flow resulting to hybridization and variations occurring unlike their counter species found in open habitats where introductions of *M.*

latifolia and *M. alba* were introduced. Furthermore, the sizes of the morphology of the varieties in cold isolated are smaller compared to those growing in warmer areas where there is ease access to factors of photosynthesis as such enough light, temperature carbon dioxide unlike the limitation occurring in cold snowy regions where these varieties are found to be native. As result of the limiting factors such varieties may adapt mechanisms of survival such as having reduced guard cells that may control the opening and closing of stomata hence controlling the gaseous exchange and transpiration. Interestingly *M. kagayamae* grows in areas with poor soil characteristics hence they have to adjust.

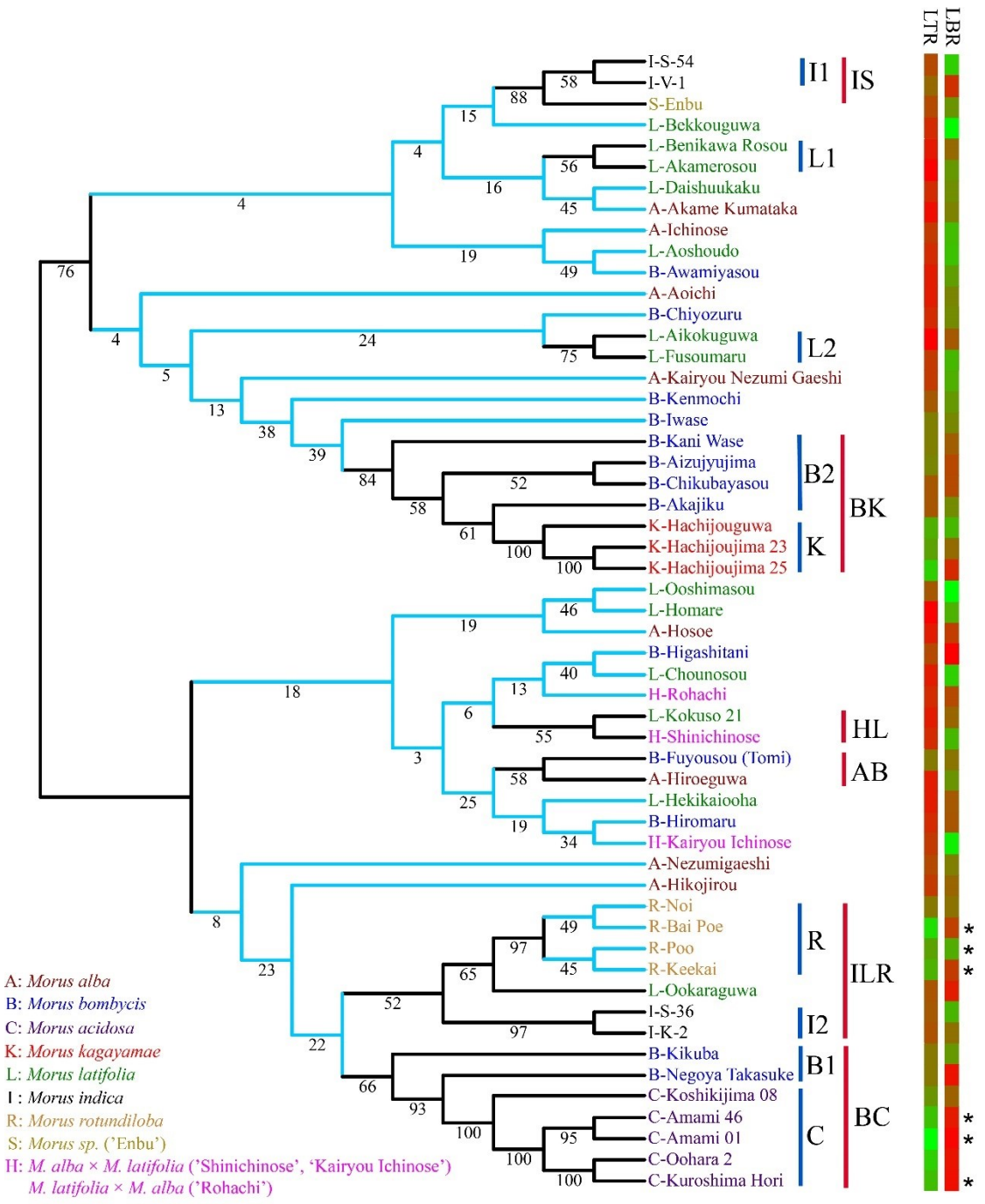

Figure 17. Neighbour joining (NJ) phylogenetic tree of concatenated 2,229 homozygous SNPs in the 54 mulberry varieties.

The values on branches are bootstrap support (%). The prefixes of the mulberry varieties denote their mulberry species: "A" (dark maroon) is *M. alba*, "B" (blue) is *M. bombycis*, "C" (purple) is *M. acidosa*, "K" (red) is *M. Kagayamae*, "L" (green) is *M. latifolia*, "I" (black) is *M. indica*, "R" (brown) is *M. rotundiloba*, "S" (gold) is *Morus* sp. ('Enbu'), and "H" (pink) is interspecific hybrids. Each light blue branch denotes low bootstrap support (<50%). Each

thick blue bar denotes a monophyletic clade of one species with enough high bootstrap support (> 50%) and each thick red bar denotes a monophyletic clade of multiple species with enough high bootstrap support (> 50%). Two color gradient bars represent relative level of the mean values of the LTR and LBR among the 54 varieties, respectively, where lighter green/red corresponds to higher/lower values. The asterisk denotes that LTR and LBR of the corresponding varieties were calculated for leaves collected from pot. The leaves collected from the field were used for remaining mulberry varieties.

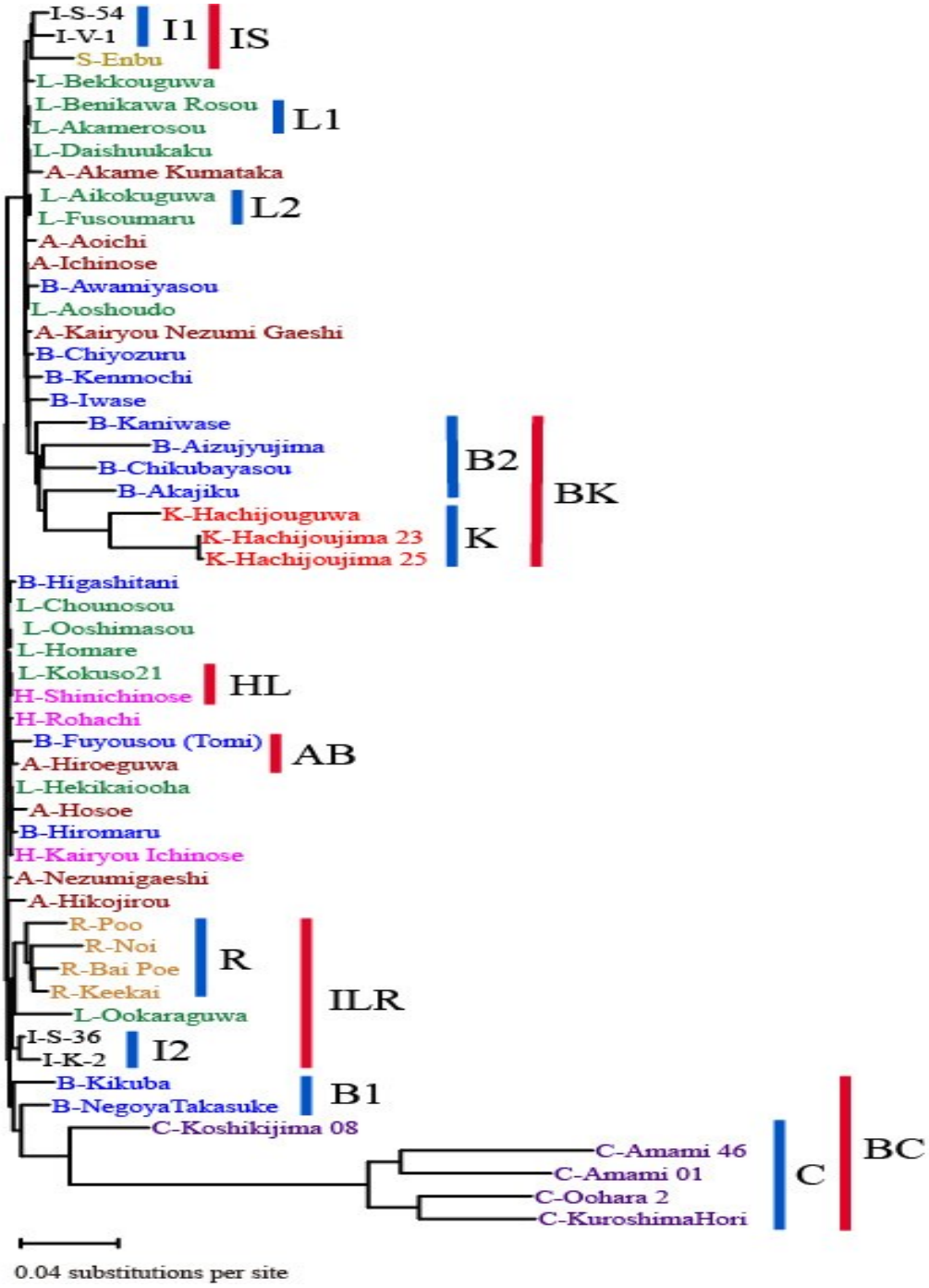

Figure 18. Neighbour joining phylogenetic tree of concatenated 2,229 homozygous SNPs among 54 mulberry varieties having bootstraps and distance.

There are 0.04 substitutions per site used at scale of 100. The color codes represent different species and interspecic hybrids denoted in **Fig. 18**.

Table 6. Morphological data of 55 mulberry varieties.

No	Species name	Variety name	Growth condition	Blade length (mm)		Apex length (mm)		Blade base depth (mm)		Leaf tip ratio (apex length/ blade length) (%)		Leaf bottom ratio (base depth / blade length) (%)	
				Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	*M. bombycis*	Aizu Jyujima	Field	164	3.86	17	3.68	10	2.36	10.20	2.34	5.88	1.28
2	*M. bombycis*	Akajiku	Field	190	9.57	13	2.94	18	2.83	6.84	1.51	9.53	1.82
3	*M. bombycis*	Awamiyasou	Field	222	1.63	9	0.47	23	2.49	4.20	0.21	10.51	1.10
4	*M. bombycis*	Chikubayasou	Field	179	8.34	12	0.47	11	1.25	6.90	0.52	5.99	0.92
5	*M. bombycis*	Chiyozuru	Field	196	14.17	10	0.47	18	2.36	4.97	0.54	9.45	1.66
6	*M. bombycis*	Iwase	Field	196	4.78	10	0.82	18	3.77	9.89	0.62	9.21	2.06
7	*M. bombycis*	Kani Wase	Field	192	2.49	19	4.32	18	2.05	9.72	2.35	7.09	1.29
8	*M. bombycis*	Kenmochi	Field	197	3.09	16	4.08	20	6.68	8.16	2.16	10.13	3.27
9	*M. bombycis*	Fuyousou (Tomi)	Field	181	13.44	17	2.16	15	2.16	9.38	0.93	8.28	0.98
10	*M. bombycis*	Higashitani	Field	245	4.08	17	2.83	5	0.00	6.96	1.27	2.04	0.03
11	*M. bombycis*	Kikuba	Field	158	3.74	16	2.36	16	1.41	10.38	1.70	10.14	0.97
12	*M. bombycis*	Negoyatakasuke	Field	187	4.71	19	3.68	6	0.47	9.96	1.77	3.04	0.30
13	*M. bombycis*	Hiromaru	Field	196	6.13	9	0.94	13	1.63	4.76	0.34	6.65	0.84
14	*M. alba*	Ichinose	Field	195	7.13	11	3.74	23	3.86	5.59	1.74	12.07	2.36

15	*M. alba*	Akame Kumataka	Field	202	0.82	6	2.94	18	2.16	2.97	1.46	8.91	1.06
16	*M. alba*	Aoichi	Field	191	4.99	8	0.82	17	3.09	4.19	0.37	9.06	1.44
17	*M. alba*	Kairyou Nezumi Gaeshi	Field	184	2.05	11	0.82	21	3.74	5.97	0.48	11.42	2.16
18	*M. alba*	Hikojirou	Field	188	10.62	12	3.86	14	1.41	6.15	1.73	7.52	1.12
19	*M. alba*	Hiroeguwa	Field	188	1.41	8	1.41	19	4.71	4.26	0.79	9.91	2.45
20	*M. alba*	Hosoe	Field	170	4.99	6	0.94	9	1.41	3.72	0.46	5.29	0.70
21	*M. alba*	Nezumigaeshi	Field	155	4.71	11	1.63	14	2.62	7.09	0.88	8.79	1.46
22	*M. alba* × *M. latifolia*	Kairyou Ichinose	Field	206	1.41	12	3.30	30	3.09	5.67	1.62	14.73	1.54
23	*M. alba* × *M. latifolia*	Shinichinose	Field	222	13.12	10	0.00	26	2.83	4.53	0.26	11.79	1.63
24	*M. latifolia*	Aikokuguwa	Field	238	10.27	5	2.87	16	1.41	1.99	1.21	6.75	0.89
25	*M. latifolia*	Akamerosou	Field	238	6.16	1	0.00	24	0.94	0.42	0.01	10.22	0.19
26	*M. latifolia*	Aoshoudo	Field	191	4.97	9	1.63	24	1.89	4.70	0.75	12.39	0.91
27	*M. latifolia*	Benikawa Rosou	Field	253	5.25	6	1.41	19	3.30	2.38	0.60	7.63	1.13
28	*M. latifolia*	Bekkouguwa	Field	224	9.42	12	2.36	35	6.34	5.18	0.82	15.44	2.56
29	*M. latifolia*	Daishuukaku	Field	240	9.18	11	1.41	23	2.16	4.61	0.74	9.61	0.99
30	*M. latifolia*	Fusoumaru	Field	216	24.51	13	3.77	25	0.00	6.14	2.55	11.75	1.36
31	*M. latifolia*	Kokuso 21	Field	258	3.30	10	0.47	19	2.05	4.00	0.18	7.49	0.87

32	*M. latifolia*	Chounosou	Field	191	4.92	7	0.00	25	4.08	3.66	0.09	13.11	2.39
33	*M. latifolia*	Hekikaiooha	Field	230	4.90	10	3.30	16	1.25	4.51	1.47	6.81	0.47
34	*M. latifolia*	Ooshimasou	Field	193	8.04	16	2.49	30	4.08	8.16	1.48	15.49	1.50
35	*M. latifolia*	Ookaraguwa	Field	247	16.36	16	6.38	9	2.94	6.65	2.87	3.61	1.08
36	*M. latifolia*	Homare	Field	237	1.89	5	0.00	27	1.89	2.11	0.02	11.52	0.85
37	*M. latifolia* × *M. alba*	Rohachi	Field	226	7.59	11	1.25	14	1.89	5.02	0.52	6.03	0.65
38	*M.acidosa*	Oohara 2	Field	166	4.97	26	0.82	5	0.47	15.68	0.65	3.21	0.18
39	*M.acidosa*	Oohara 2	Pot	157	4.71	27	2.49	8	1.89	17.00	1.24	4.86	1.08
40	*M.acidosa*	Amami 01	Pot	158	5.10	30	0.00	4	1.89	19.01	0.62	2.35	1.23
41	*M.acidosa*	Amami 46	Pot	147	5.73	22	4.78	5	1.63	15.19	3.29	3.42	1.22
42	*M.acidosa*	Kuroshima Hori	Pot	123	5.35	18	2.05	4	0.47	14.41	2.00	2.97	0.30
43	*M. acidosa*	Koshikijima 08	Field	185	6.34	21	1.89	13	2.36	11.56	0.97	7.24	1.37
44	*M. kagayamae*	Hachijoujima 35	Field	218	9.53	26	3.30	13	6.24	11.87	2.00	6.12	2.78
45	*M. kagayamae*	Hachijoujima 23	Field	229	5.79	29	6.80	18	0.47	12.87	3.15	7.73	0.31
46	*M. kagayamae*	Hachijoujima 25	Field	259	4.32	43	4.78	11	2.94	16.51	2.08	4.26	1.17
47	*M. kagayamae*	Hachijouguwa	Field	232	1.25	31	4.32	27	5.56	13.39	1.90	11.79	2.38
48	*Morus* sp.	Enbu	Field	241	2.94	17	2.36	24	2.94	6.93	1.07	9.95	1.13
49	*M. indica*	K-2	Field	197	5.72	13	1.70	16	3.30	6.42	0.73	8.01	1.92

50	*M. indica*	S-36	Field	195	2.49	15	1.63	23	1.70	7.71	0.90	11.65	0.88
51	*M. indica*	S-54	Field	190	3.40	13	2.05	25	5.44	6.68	1.06	13.06	3.13
52	*M. indica*	V-1	Field	225	4.50	20	1.25	10	0.47	8.77	0.72	4.61	0.30
53	*M. rotundiloba*	Noi	Field	171	3.30	16	0.94	14	2.16	9.17	0.39	8.18	1.14
54	*M. rotundiloba*	Noi	Pot	175	12.28	19	5.66	16	3.30	11.08	4.03	8.89	1.49
55	*M. rotundiloba*	Poo	Pot	156	1.70	21	1.41	18	2.36	13.43	0.87	11.30	1.47
56	*M. rotundiloba*	Bai Poe	Pot	134	6.94	23	1.70	7	0.94	16.96	0.91	5.50	0.79
57	*M. rotundiloba*	Keekai	Pot	140	2.45	19	0.94	7	1.70	13.82	0.88	5.23	1.19

*The mean values and standard deviation of three selected replicates were calculated for each morphological characteristic of mulberry leaves. The leaves were collected from the field and/or pot**.**

3.3.5 Evaluation of distribution of RAD-loci on the genome scaffolds of *Morus notabilis*

Overall, the 15,605 RAD-loci including 47,839 SNP markers (heterozygous and homozygous) genotyped in 50 or more mulberry varieties by the ddRAD-seq analysis were compared with the genome scaffolds of *M. notabilis* (total size is 320,378,613 bp with 16,272,868 bp gap regions) by blastn. As a result, 91.3% (14,249) of the RAD-loci were mapped onto the 1446 scaffolds covering a total size of 285,013,843 bp (88.9% of the total size) of the genome scaffolds. Of these, 2112 RAD-loci were mapped on top of the 29 longest scaffolds. The results revealed that the 2112 RAD-loci were widely and randomly distributed as shown in **Fig. 19**. Regarding 1,855 RAD-loci including 2,229 homozygous SNPs used in the phylogenetic analysis (**Fig. 17** and **18**), 1576 RAD-loci (85%) were mapped on 664 scaffolds totalling 226,849,922 bp (70.8% of the total size). Of these, 230 RAD-loci were widely and randomly mapped on the top of the 29 longest scaffolds (**Fig. 20**). The results confirmed that genome-wide SNP markers were successfully generated and used for the phylogenetic analysis.

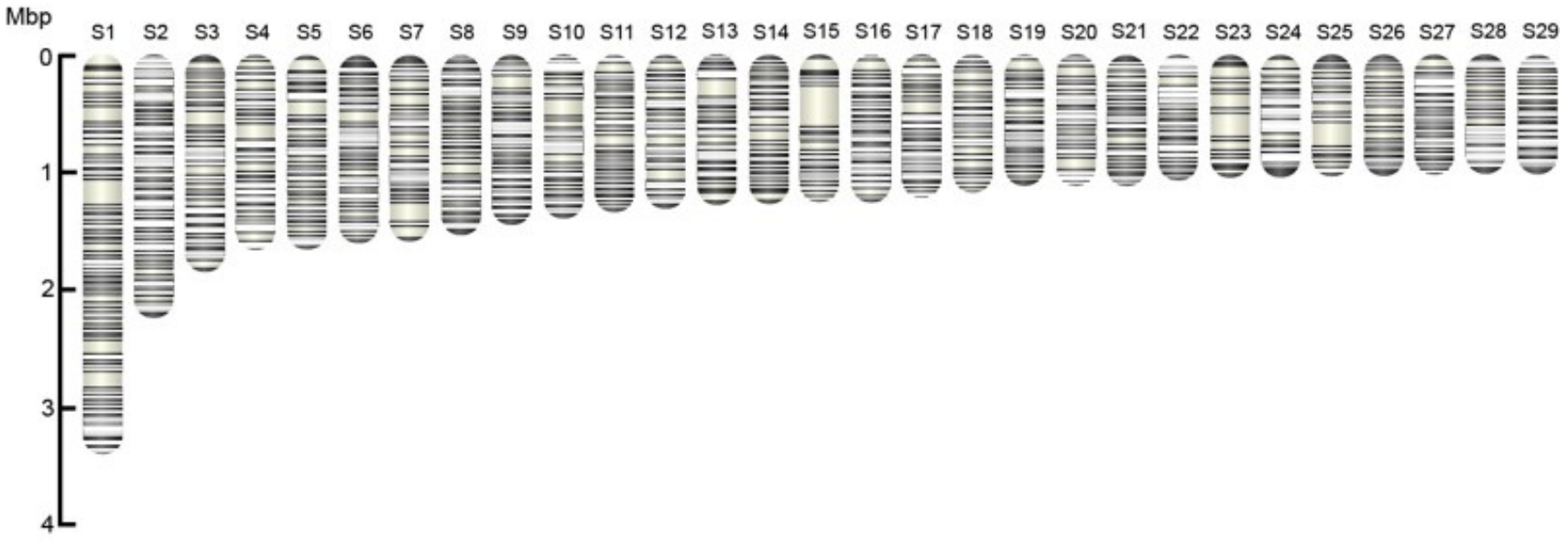

Figure 19. RAD-loci with homozygous and/or heterogyzous SNPs mapped on the top 29 genome scaffolds of *M. notabilis*

The black horizontal bars represent positions (bp) of 2112 RAD-loci mapped on the genome scaffolds. The RAD-loci include homozygous and/or heterozygous SNPs. S1-S29 are the top 29 longest genome scaffolds whose lengths are higher than 1 Mbp. The scaffolds were sorted in descending order of the length.

Figure 20. RAD-loci with homozygous SNPs mapped on the top 29 genome scaffolds of *M. notabilis.*

The black horizontal bars represent positions (bp) of 230 RAD-loci mapped on the genome scaffolds. The RAD-loci include only homozygous SNPs.

3.4 Discussion

With the help of ddRAD-seq analysis, large number of SNP markers, which were randomly and genome-widely distributed, were extracted for the 54 mulberry varieties which belong to the seven mulberry species and one unknown species ('Enbu'). The number of SNP markers in this study was larger than the number of DNA markers used for phylogenetic analysis in previous studies on mulberry (Sharma et al.,2000; Kafkas et al., 2008; Nepal, 2008; Nepal and Ferguson, 2012, Ipek et al., 2012; Zeng et al., 2015).

The phylogenetic analysis using the 2,229 homozygous SNP markers were conducted for the 54 mulberry varieties. The result showed that two species native to Japan, *M. kagayamae* and *M. acidosa,* formed two different clear monophyletic clades (**Fig. 17 and 18**). Since these species have been distributed in two isolated islands of Izu (*M. kagayamae*) and Nansei (*M. acidosa*) as shown in **Fig. 15**, the two clear monophyletic clades may infer that there was no gene flow among these species due to isolation of their geographical regions. Another clear monophyletic clade was observed in the mulberry varieties of native species to Thailand, *M. rotundiloba* (**Fig. 17 and 18**). The clear monophyletic clade might also be due to similar reason.

On the other hand, in *M. bombycis*, another native species to Japan, a clear monophyletic clade of single species was not observed among the varieties. The *M. bombycis* varieties were roughly grouped into three groups. Two of the three groups, B1 and B2, belong to larger monophyletic clades with *M. acidosa* (BC) and *M. kagayamae* (BK), respectively. Whereas another group including the remaining *M. bombycis* varieties was located in nonmonophyletic clades mainly including the mulberry varieties of the two introduced species from China, *M. alba* and *M. latifolia*, and their interspecific hybrids (**Fig. 17 and 18**). Regarding the two introduced species from China and their hybrids, only small monophyletic clades of two varieties were observed in *M. latifolia* while no monophyletic clade was observed in *M. alba*.

The result indicates that the varieties in the nonmonophyletic clades are likely to be highly hybridized. Since the varieties of *M. alba* and *M. latifolia* have been grown and well distributed in many areas of Japan of varied climatic conditions for silkworm rearing, the possibility of ease of crossing among them may have caused the nonmonophyletic clades including them (high level of hybridization).

From the result of the two clades, B1 and B2, of *M. bombycis,* the two mulberry species native to Japan, *M. acidosa* and *M. kagayamae,* may have been evolved from the two *M. bombycis* groups. The origin sites for the two *M. bombycis* varieties belonging to the two clades, B1 and B2, ('Negoya Takasuke' and 'Aizujyujima') are cold snowy regions of Honshu (**Fig. 15**). Since the production of *M. alba* and *M. latifolia* varieties is usually affected by cold snowy conditions resulting to mulberry dieback and difficulties to adapt, this may have made climatic isolation conditions impossible for natural crossing among ancestors of the varieties of the two introduced species and *M. bombycis*, which might lead to low hybridization level of the *M. bombycis* varieties. The classification of the two *M. bombycis* varieties, 'Negoya Takasuke' and 'Aizujyujima', in the monophyletic clade might reflect the hypothesis. However, other two *M. bombycis* varieties in cold snowy region, 'Higashitani' and 'Kenmoshi', were classified in nonmonophyletic clades. Therefore, other factors might affect the level of hybridization for the *M. bombycis* varieties.

The formation of two distinct monophyletic clades I1 and I2 in among *M. indica* varieties may reveal that the ancestral background of the varieties in each clade may be different (**Fig 17** and **18**). Interestingly, there is a possibility that Enbu' (*Morus* sp.) and the *M. indica* varieties in the I1 clade, 'V-1' and 'S-54', have similar genetic backgrounds because they formed a monophyletic clade (ILR) (**Fig. 17** and **18**). Although 'Enbu' has been registered as wild African mulberry, *M. mesozygia*, in the NARO Genebank, the morphologic features are clearly different from those of *M. mesozygia*, and therefore we classified the species of 'Enbu' as

unknown. Since there is a record showing that Japanese mulberry trees were introduced into Kenya by the United Kingdom (Eliot, 1905), ancestors of the two *M. indica* varieties and 'Enbu' might also have been introduced into Kenya in similar way. The other *M. indica* clade (I2) formed a monophyletic clade (ILR) with *M. rotundiloba* (R) and 'Ookaraguwa' (*M. latifolia*) (**Fig. 17** and **18**). 'Ookaraguwa' whose ancestry was derived from China which might suggest that the varieties of *M. rotundiloba* (R) and the two varieties of *M. indica*, 'S-36' and 'K-2', were derived from common ancestors from China.

Evaluated morphological feature results showed relatively high leaf tip ratio (LTR) observed in the two clades B1 and B2 of *M. bombycis* varieties classified close to the two species, *M. kagayamae and M. acidosa*, whereas low leaf tip ratio was observed in *M. bombycis* mixed in other clades having highly hybridized species of *M. alba* and *M. latifolia* and their interspecific hybrids (**Fig. 17**). This morphological feature of LTR may be used as a marker of hybridization among the *M. bombycis* varieties, because the characterization of a high apex length is known as a common feature that has been observed over time (Minamizawa,1984).

Despite acknowledging the probability of using the leaf tip ratio as a morphological marker based on the result in *M. bombycis*, it is not clear on use of the leaf bottom ratio (LBR) which varied among the varieties in monophyletic clades (B1 and B2) and in the non-monophyletic clades. The high LTR and low LBR were observed in the varieties of *M. acidosa* and the variety of *M. bombycis* ('Negoya Takasuke') in the clade BC which consists of the clade C and B1. This might support the hypothesis that the *M. acidosa* varieties might be evolved from *M. bombycis*.

From the comparative result of phylogeny and morphological data, it is difficult to rely on one technology for classification of mulberry, however the combined use of the strategy may play a major role in enabling more morphological features that may be of use as markers.

3.5 Conclusion

The large number of genome-wide SNP markers among the 54 mulberry varieties of the seven and one unknown species were successfully generated by the ddRAD-seq analysis (Muhonja et al., 2020). The phylogenetic analysis using selected homozygous 2,229 SNPs revealed several new insights for the relationships among the mulberry varieties, hybridization among some indigenous mulberry varieties in Japan, and a morphological characteristic of LTR which might be available as a morphological marker. The large quantity of genomic data generated in this study can be useful to one on one tailored marker research while complimenting the information available in the *Morus* database. Additional studies need to be conducted with regard the impact of the identified SNPs for utilization in mulberry crop improvement and breeding strategies. The combination of the discovered SNPs markers and the morphological features will forge a strategy of improving on the characterization and conservation of the mulberry there by narrowing down the classification debate of species identity for ease of management and utilisation of the genetic resource.

3.6 References

Awasthi A.K., Nagaraja G.M., Naik G. V., Kanginakudru S., Thangavelu K., Nagaraju J., (2004). Genetic diversity and relationships in mulberry (genus *Morus*) as revealed by RAPD and ISSR marker assays. *BioMed Central Genetics.* 5, 1–9.

Barchi L., Lanteri S., Portis E., Acquadro A., Valè G., Toppino L., and Rotin, G. L., (2011). Identification of SNP and SSR markers in eggplant using RAD tag sequencing. *BioMed Central Genomics,* 12, 304-313.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., and Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology*, *22*(11), 3124–3140.

Chumchuen S., Kanekatsu R., (2011). AFLP-based transcript profiling for genetic relationships of mulberry (genus *Morus*) germplasm. *Journal of Insect Biotechnology and Sericology.* 80(2), 63–70.

Datta R.K., (2002). Mulberry Cultivation and Utilization in India, in: FAO Electronic Conference on Mulberry for Anımal Productıon (*Morus*-L), 1–24.

Doyle J.J., and Doyle, J.L., (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin,* 19(1), 11–15.

Eliot C., (1905). The East Africa Protectorate. Arnold, London.

Emerson K. J., Merz C. R., Catchen J. M., Hohenlohe P. A., Cresko W. A., Bradshaw W. E., and Holzapfel C. M., (2010). Resolving postglacial phylogeography using high-throughput sequencing. *Proceedings of the National Academy of Sciences*, 107(37), 16196–16200.

Hotta T., (1958). Taxonomical studies on the *Morus* plants and their distributions in Japan and its vicinities. Japanese Society for Promotion of Science, Tokyo.

İpek M., Pirlak L., Kafkas S., (2012). Molecular characterization of mulberry (*Morus* spp.) genotypes via RAPD and ISSR. *Journal of the Science and Food Agriculture.* 92(8), 1633–1637.

Jiang Y., Huang R., Yan X., Jia C., Jiang S., and Long T., (2017). Mulberry for environmental protection. *Pakistan J. Bot.* 49, 781–788.

Kafkas S., Özgen M., Doğan Y., Özcan B., Ercişli S., and Serçe S., (2008). Molecular Characterization of Mulberry Accessions in Turkey by AFLP Markers. *Journal of American Society of Horticultural Science.* 133(4), 593–597.

Kumar S., Stecher G., Li M., Knyaz C., andTamura K., (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution.* 35(6), 1547-1549.

Letunic I., and Bork P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, 47, W256-W259.

Li T., Qi X., Zeng Q., Xiang Z., and He N., (2014). MorusDB: A resource for mulberry genomics and genome biology. *Database(Oxford)*, 2014, bau054.

Mathithumilan B., Sajeevan R. S., Biradar J., Madhuri T., Nataraja K. N., and Sreeman,S. M., (2016). Development and characterization of genic SSR markers from Indian mulberry transcriptome and their transferability to related species of Moraceae. *PLoS ONE* 11(9), e0162909..

Mathithumilan B., Kadam N.N., Biradar J., Reddy S.H., Ankaiah M., Narayanan M.J., Makarla U., Khurana P., Sreeman S.M., (2013). Development and characterization of microsatellite markers for *Morus* spp. and assessment of their transferability to other closely related species. *BioMed Central Plant Biology.* 13(1), 194-215.

Minamizawa K., (1984). Science of mulberry cultivation: fundamental and practice (In Japanese). Tokyo.

Muhonja L., Yamanouchi H., Yang C., Kuwazaki S., Yokoi K., Kameda T., Sezutsu H., and Jouraku A., (2020). Genome-wide SNP marker discovery and phylogenetic analysis of mulberry varieties using double-digest restriction site-associated DNA sequencing. *Gene* 726(2020), 144162.

Nepal M.P., (2008). Systematics and reproductive biology of the genus *Morus* L. (Moraceae). Doctorate thesis, Kansas State University.

Nepal M.P., and Ferguson C.J., (2012). Phylogenetics of *Morus* (Moraceae) inferred from ITS and trnL-trnF sequence data. *Systematic Botany.* 37, 442–450.

Peterson B. K., Weber J. N., Kay E. H., Fisher H. S., and Hoekstra H. E., (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS One*, 7(5), e37135.

Saitou N. and Nei M., (1987). The neighbour-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.,* 4(4), 406–425.

Scaglione D., Acquadro A., Portis E., Tirone M., Knapp S.J., Lanteri S., (2012.) RAD tag sequencing as a source of SNP markers in *Cynara cardunculus* L. *BioMed Central Genomics,* 13, 3-14.

Schneider C.A., Rasband W.S.,and Eliceiri K.W., (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature. Methods,* 9, 671–675.

Sharma A., Sharma R., and Machii H., (2000). Assessment of genetic diversity in a *Morus* germplasm collection using fluorescence-based AFLP markers. *Theoretical and.Applied. Genetics.* 101, 1049–1055.

Vijayan K., Srivastava P.P., and Awasthi A.K., (2004). Analysis of phylogenetic relationship among five mulberry (*Morus*) species using molecular markers. *Genome* 47, 439–448.

Vijayan K.,. (2007). Molecular Markers and their Application in mulberry breeding. *International Journal of Industrial Entomology*, *15*(2), 1-12.

Vijayan, K., (2010). The emerging role of genomic tools in mulberry (*Morus*) genetic improvement. *Tree Genet. Genomes* 6, 613–625.

Vijayan K., Srivastava P. P., Raju P. J., and Saratchandra B. (2012). Breeding for higher productivity in mulberry. *Czech Journal of Genetics and Plant Breeding*, 48(4), 147–156.

Zeng Q., Chen, H., Zhang C., Han M., Li T., Qi X., Xiang Z., and He N., (2015). Definition of Eight Mulberry Species in the Genus *Morus* by Internal Transcribed Spacer-Based Phylogeny. *PLoS One* 10, e0135411.

Zhao W., Zhou Z., Miao X., Zhang Y., Wang S., Huang J., Xiang H., Pan Y., and Huang Y., (2007). A comparison of genetic variation among wild and cultivated *Morus* Species

(Moraceae: *Morus*) as revealed by ISSR and SSR markers. *Biodiversity and Conservation* 16, 275–290.

Chapter 4: Admixture analysis of mulberry varieties of genus *Morus*

4.1 Introduction

Admixture analysis involves estimation of a combination of different ancestral populations among species that may have resulted from the process of hybridization. As such the main impact of admixture aims to boost the fitness and population viability while eliminating the consequences of inbreeding depression. Several causes have been flaunted as causes of admixture in a population that may include the introductions of materials into an area, translocation of genetic material among the populations as well as the removal of isolation barriers (Lavergne and Molofsky, 2007; Shi et al., 2018). Interesting it has been suggested that the higher the gene flow and larger the population size, the higher there may be genetic variations among the populations (Star and Spencer, 2013).

Following the genome wide SNPs analysis of the 54 of the 56 varieties under study (Muhonja et al., 2020), showed key highlights with possibilities of clear monophylecity occurring in *M. kagayamae, M. acidosa* and introduced *M. rotundiloba* in Japan as well as unclear non monophyletic clades in varieties that were introduced. As hypothesized previously that most varieties existing in Japan could be highly hybridised based on introductions as well as the ease of crossing, hence the result of the closeness of the varieties in the non-monophyletic clades providing an indicator of this through the short branch lengths in the phylogeny analysis that could not be easily segregated among the varieties of different species. However, this unclarity could be based on the level of hybridization among them that may have occurred among the varieties over time. Interestingly, Shi et al., (2018) through costs and befits of admixtures on genotypes had highlighted that known hybridizations /admixture may create a great impact of improving species fitness and continuity in populations thus reducing the risk of inbreeding depression while increasing diversity dynamics The above could be true in

mulberry in that the adaptation to different climatic conditions Machii et al., (2000) may provide support on this. Furthermore, with the diversity dynamics could result to divergent populations, where Edmands, (1999) has shown that over time there is usually outbreeding depression causing speciation which may be observed in the initial F1 as a result of the disruption of their location adaptation or from F2 onwards generations after exposure to separation and recombination. The admixture analysis therefore was conducted with aim of estimating the level of hybridization and genetic relatedness that may exist among the seven species of *M. acidosa, M. alba, M. bombycis*, *M. indica*, *M. Kagayamae, M. latifolia, M. rotundiloba* and one unspecified species *Morus* sp. ('Enbu').

4.2 Materials and methods

Admixture analysis was conducted using ADMIXTURE v 1.3 (Alexander et al., 2009) on 54 varieties of mulberry belonging to eight species analysed in the phylogenetic analysis. Evaluation of the admixture analysis used the 47,839 SNPs of both homozygous and heterozygous in nature where estimation of cross validation error of each number of ancestral population (K) value was used in order to estimate the relationship of the ancestral population of each variety. The maximum likelihood estimation method was adapted for ancestral populations with the with cross validation error being estimated. 'K' value range was placed at 1 to 8 based on the expected number of ancestral populations. The result was further deduced using the best fit lowest cross validation error that had the lowest validation error (CV) when compared to other populations (Alexander et al., 2015). Selection of the lowest cross validation error was used as a close estimate for grouping the varieties in standalone species and admixtures.

4.3 Results

Admixture analysis (Alexander et al., 2009, Zhou et al., 2009) was conducted on 54 selected diploid varieties of eight mulberry species that yielded 47,839 genome wide SNPs genotyped in 50 or more varieties containing both homozygous and heterozygous in order to estimate the genetic relatedness among them through the ancestral populations (designated as K). Through mapping the SNPs, the stratification results revealed that there was clustering of the varieties into subgroups with those with over 80% being categorised as independent while the others as admixtures based on the ancestral populations.

The admixture analysis with cross validation error estimation using the seven and one unknown species revealed that the ancestral populations of the varieties of *M. acidosa, M. kagayamae, M. rotundiloba, M. bombycis, M. indica,* and 'Enbu' fall in the monophyletic clades along the evaluated eight ancestral populations (K) while other varieties fall in the non-monophyletic clades as shown in **Fig. 21-24**. The cross validation (CV) error values of different K values for estimating the ancestral populations ranged from 0.37194 to 0.46821 and the cross-validation error (CV) value was lowest when K value was set to 3 and 4 in **Fig. 22**. Similar values were observed in **Fig. 23** and **24**. The result obtained is consistent with previously found in the phylogeny analysis (**Fig 17**). The admixture structure (K=4) chosen as optimised among the ordered varieties by consensus tree, distance method and by species category shows that the varieties of *M. bombycis* and *M. kagayamae* in the monophyletic clades were estimated to be derived from the same ancestral population (green coloured in **Fig. 21**) and the same is true for the varieties of *M. rotundiloba*, *M. indica* and 'Enbu' (red coloured in **Fig. 21**) while the varieties of *M. acidosa* were estimated to be derived from a different ancestral population (blue coloured in **Fig. 21**). Similar tendency was observed in each result of different K value and different order of the varieties (**Fig. 22-24**).

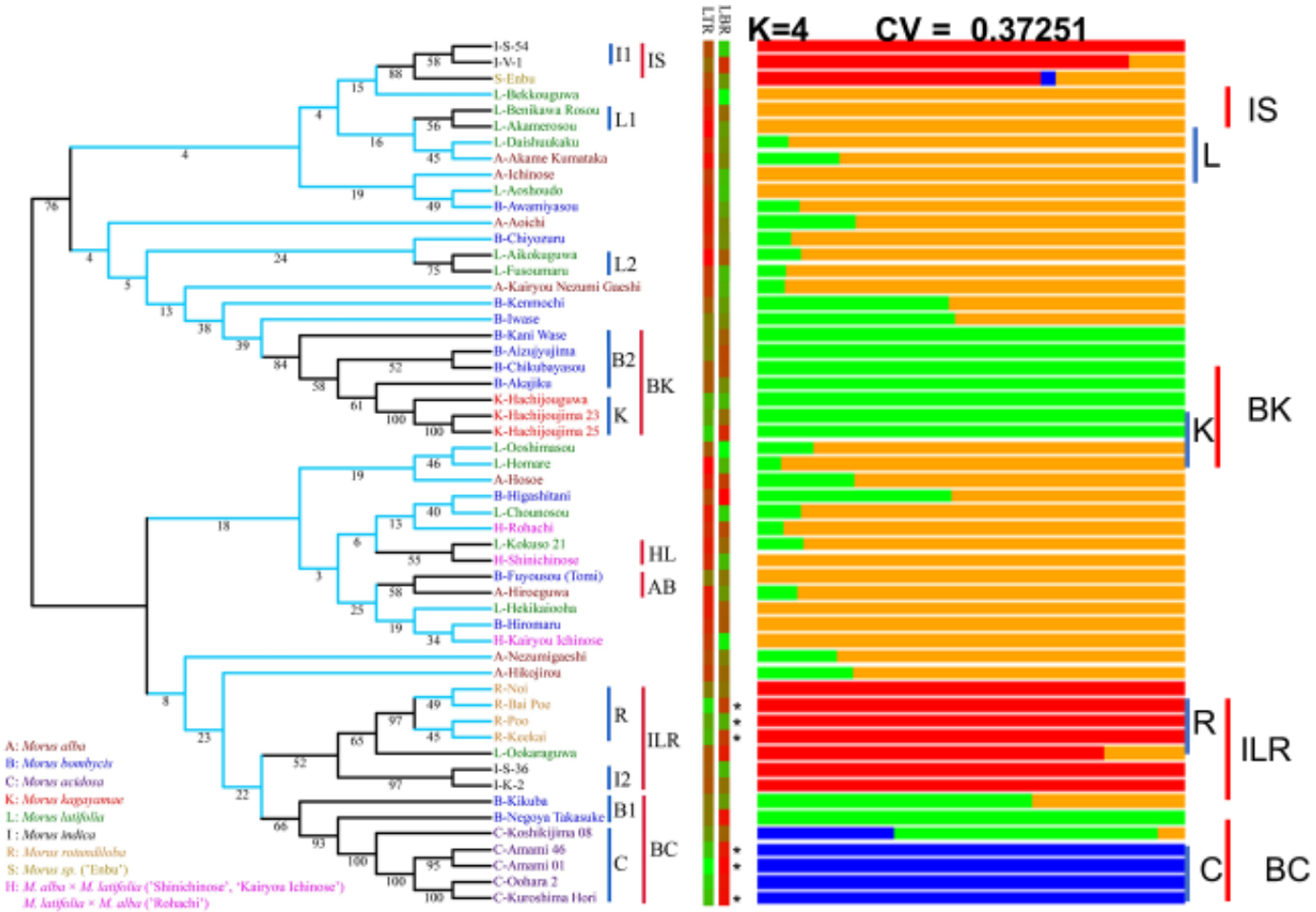

Figure 21. Admixture structure of K=4 based on 47,839 SNPs across the 54 varieties in relation to their neighbour joining tree.

The bands of different colours indicate the clusters of separation within the populations with different varieties among the species subgroups creating clear clades of **C**; *M. acidosa*, **R**; *M. rotundiloba*, **K**; *M. kagayamae* and **L**; *M. latifolia* among the hybridized varieties. The cross-validation error (CV) value was lowest when K value was set to 3 and 4. The detail of the NJ-tree was shown in **Fig. 17**.

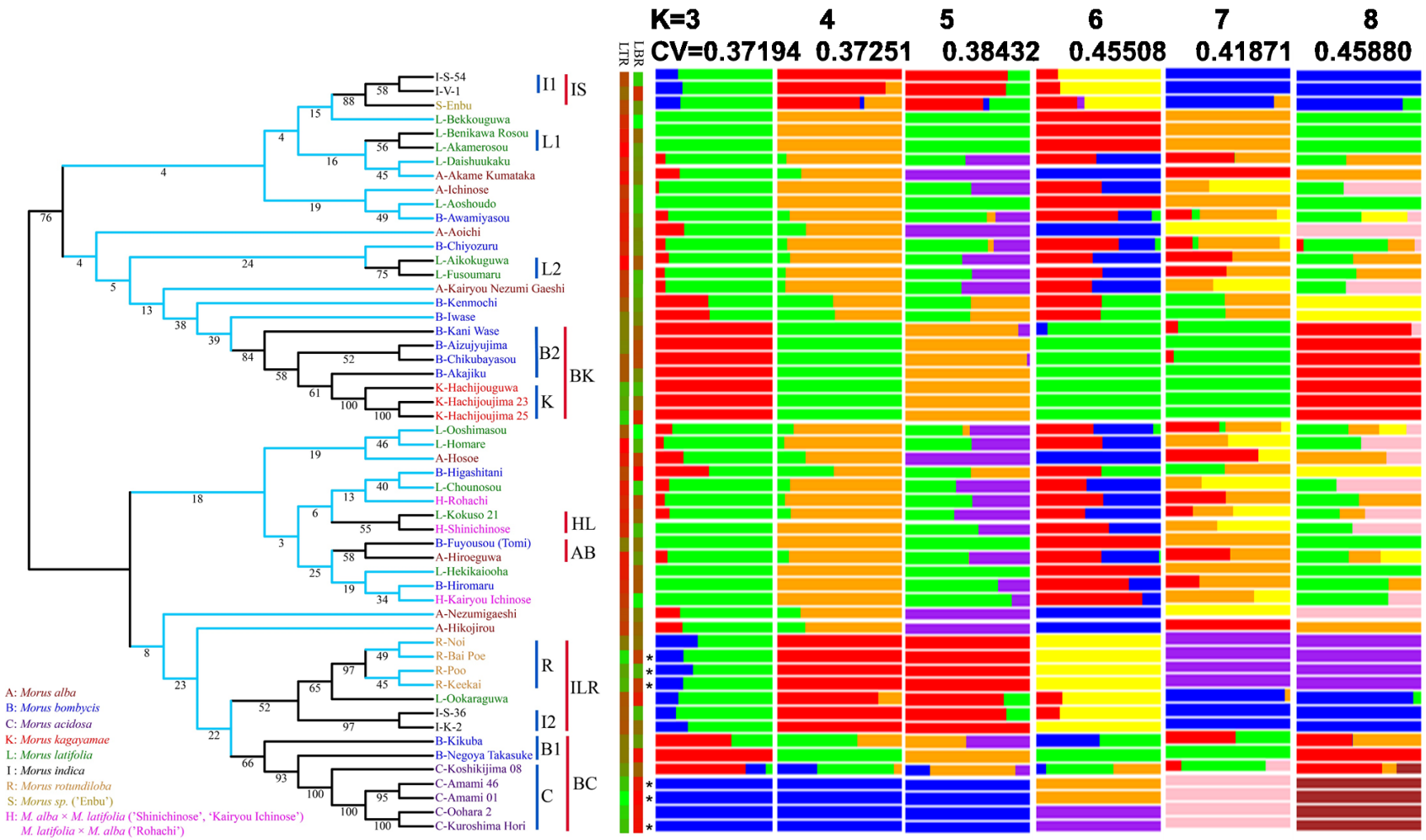

Figure 22. Admixture structures of different K values based on 47,839 SNPs across the 54 varieties in relation to their neighbour joining tree.

The bands of different colours indicate the ancestral population cluster separation. The detail of the NJ-tree was shown in **Fig. 17**.

In addition, the two separated clades of *M. indica* and 'Enbu' share the same clustering characteristics suggesting the common genetic background and indicator of common ancestry with the Indian varieties. 'Enbu' may be closely related to the variety 'V1' of *M. indica* (**Figure 23**). Moreover, the separation of *M. bombycis* varieties into three groups with some being standalone while others having different levels of admixture may suggest the interaction of geneflow among the varieties in their habitat and introduction areas. The characteristic of the interspecific hybrids varieties 'Kairyoichinose', 'Shinichinose' and 'Rohachi' that have been clustered with *M. alba* and *M. latifolia* varieties, respectively, at K=3 shows that they stand alone, however, the presence of admixture in 'Shinichinose' and 'Rohachi' but not 'Kairyoichinose' can be observed by increasing the cross validation value K to 4. Overall, increasing the cross-validation values lead to spread the gene flow further showing more hybridizations that may have happened over time.

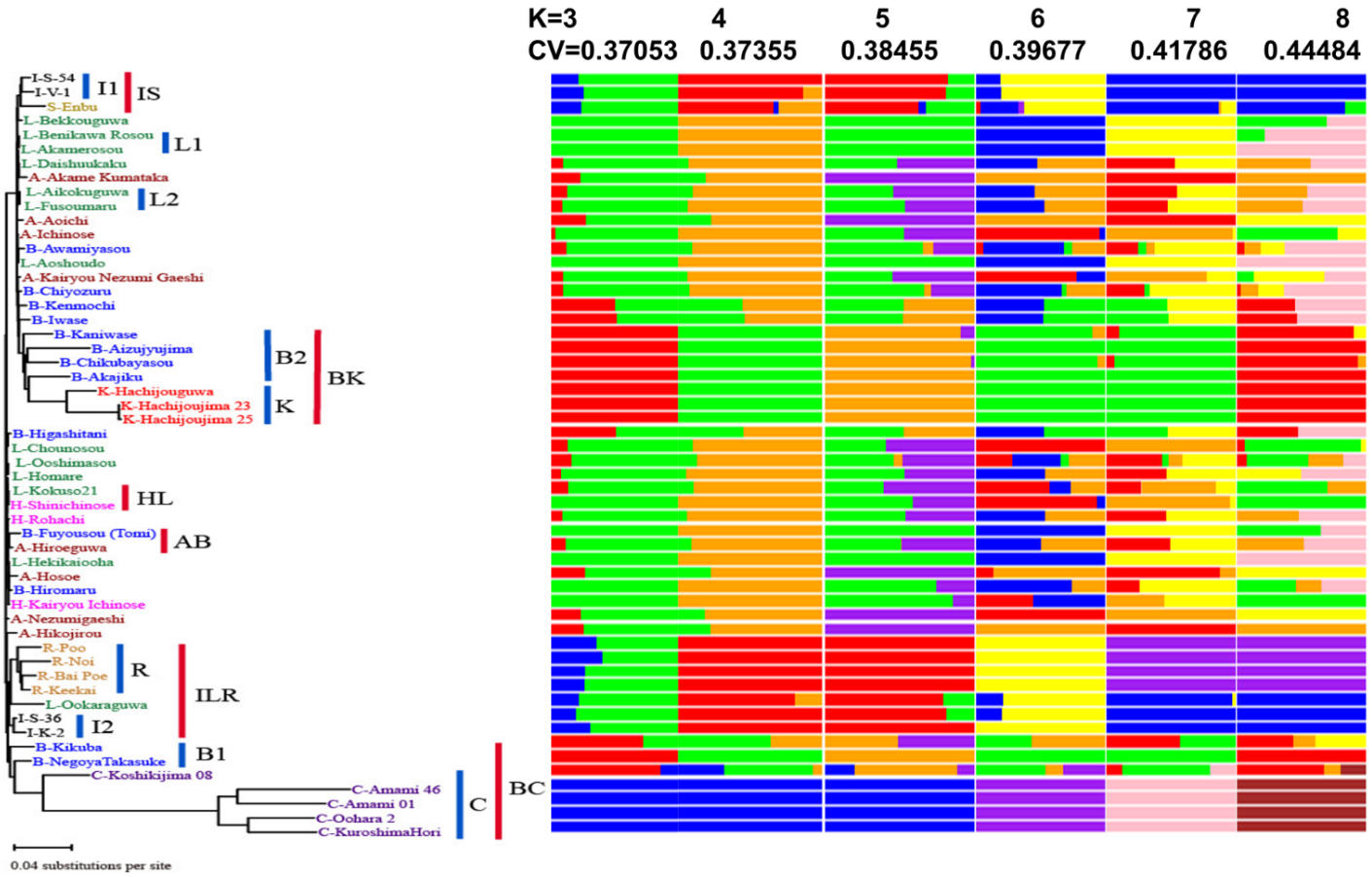

Figure 23. Admixture structures of different K values based on 47,839 SNPs across the 54 varieties in relation to their neighbour joining tree with the distance information.

The bands of different colours indicate the ancestral population cluster separation. The cross validation (CV) error values for the K values ranged from 0.37053 (K=3) to 0.46740 (K=1). The CV was lowest for K=3 while almost the same value was observed for K=4. The detail of the NJ-tree was shown in **Fig. 18**.

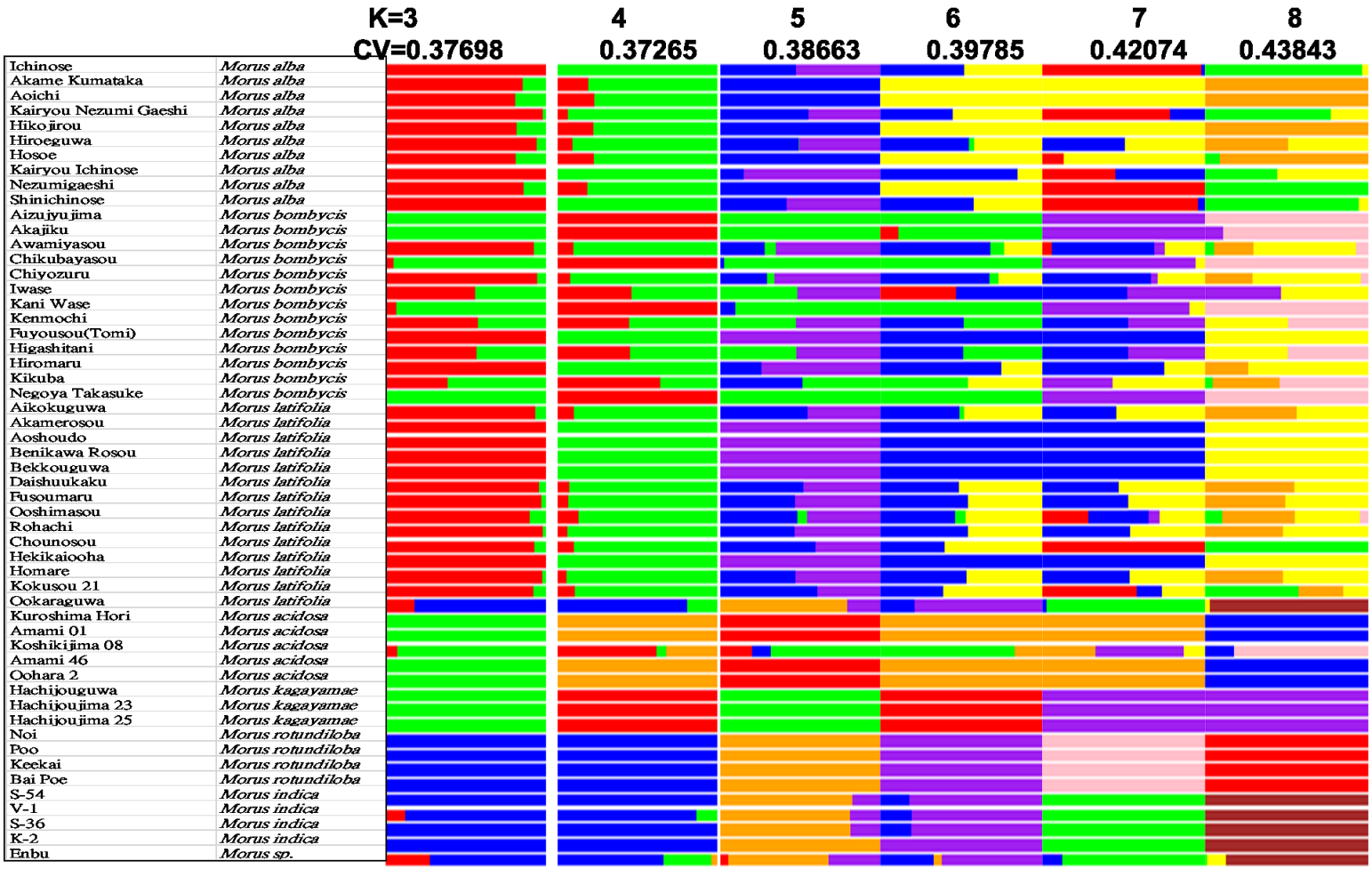

Figure 24. Admixture structures of different K values based on 47839 SNPs across the 54 varieties in relation to the documented species categorization in Japanese Genebank.

The bands of different colours indicate the ancestral population cluster separation. The cross validation (CV) error value was lowest for K=4.

4.4 Discussion of the Admixture analysis results

The result of admixture analysis derived from both heterozygous and homozygous SNPs of the 54 mulberry varieties reveal consistency in the results obtained in the phylogeny analysis. The ordering of the varieties may have been different based on either the consensus tree, distance tree method and by species however what stands out is the ancestry affiliation of the monophyletic clades that do not change irrespective the number of ancestral populations used (thus K=3 to 8) as shown in **Fig. 22-24** while the non-monophyletic clades equally can be observed by the presence of mixed ancestral population by their fractions.

M. acidosa stands out with deep branch length having all the five varieties isolated from the rest although one of them 'Koshikijima 08' seems to be slightly hybridized with either ancestry of *M. bombycis* or introduced ancestry of *M. latifolia* and *M. alba* hence being isolated from the main clade of pure *M. acidosa* (**Fig. 22-24***).*

M. bombycis varieties on the other hand shows some ancestry with the *M. kagayamae* while some of the other groups show combinations of either *M. latifolia* or *M. alba* ancestry possibility of existence of the native species and the effect of hybridization that may have taken place over the time from the introductions from China and India as shown in the interactions among the blocks that could indicate the relationship of common background or ancestry among the mulberry varieties **(Figure 22)**. The above results to categorising the varieties under *M. bombycis* into three categories of pure native varieties not hybridized 'Negoya Takasuke, Kikuba, Akajiku, Chikubayasou, Aizujyujima KaneWase and Iwase' those with high affinity to *M. latifolia* 'Chiyozuru and Awamiyasou'and those towards *M. alba* 'Higashitani, Fuyousou(tomi) Hiromaru and Kenmochi through possible hybridization.

Interesting, there seems to be the isolation island model populations playing a part in the ancestry populations where varieties from the islands of Nansei Islands having *M. acidosa* varieties and Izu Island having *M. kagayamae* varieties standing out with the monophyletic

clades followed by the introduced species of *M. rotundiloba* varieties from Thailand confirming the isolated species category (**Figure 21**). The Island isolation model may have contributed toward preventing the introductions of *M. latifolia, M. alba* and *M. indica* from China and India respectively from crossing with the native species of *M. acidosa ,M. kagayamae* and introduced *M. rotundiloba* from Thailand.

Based on the findings of these study, there exist admixtures among the varieties within the species which could have been a result of hybridization in the habitat areas thereby introducing different genes into the genetic pool of the specific species. For example, as you estimate the level of combinations at K=6 to 8 in **Fig. 22-24**, I observe that there is possibility of assumed grouping of the varieties based on morphology in the highly hybridised region of the phylogeny that require either reclassification or renaming due to strong affiliation of the admixture fractions of the ancestral populations among them. There is strong hybridization observed occurring among the *M. alba*, *M. bombycis* and *M. latifolia* at species level based on K= 4 with a few varieties maintaining their purity stands among them (**Figure 21)**. The hybridization characteristic observed in this study should be utilised to enhance conservation strategies in mulberry through increasing population variability as I observed that the isolation by island location as well as the gene flow may have contributed towards the admixtures. The Japanese native species *M. acidosa* and *M. kagayamae* in isolation showed this characteristic while the *M. bombycis* in open interaction habitats showed hybridization effects hence the groupings.

Noted previous studies (Sharma et al., 2000) had indicated diversity however they did not analyse the level of variation among the varieties as well as the study was limited in number of species. This admixture analysis results contribute towards knowledge base of genetic variation and extent of admixture combinations that occur in both open and isolated habitats thus offering grounds for further studies.

4.5 Reference

Alexander D.H., Novembre J., and Lange K., (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research,* 19(9), 1655–1664.

Alexander D.H., Novembre J., and Lange K., (2015). Admixture 1.3 software manual.Genetics software.

Edmands S., (1999). Heterosis and Outbreeding Depression in interpopulation crosses spanning a wide range of divergence..*Evolution*, 53(6), 1757–1768.

Lavergne S., and Molofsky J., (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America*, 104(10), 3883–3888.

Linneaus C.(1753). *Morus* in Species plantarum. Stockholm, Sweden ,2, 986.

Machii H., Koyama A., and Yamanouchi H., (2000). Mulberry breeding, cultivation and utilization in Japan. *FAO Electronic Conference on Mulberry for AnimalProduction* 8, 63-72.

Muhonja L., Hiroaki Y., Ching-chia Y., Seigo K., Kakeru Y., Tsunenori K., Hideki S., and Akiya J., (2019). Genome-wide SNP marker discovery and phylogenetic analysis of mulberry varieties using double-digest restriction site-associated DNA sequencing. *Gene* 726, 144162.

Sharma A, Sharma R, and Machii H., (2000). Assessment of genetic diversity in a *Morus* germplasm collection using fluorescence-based AFLP markers. Theoretical and Applied Genetics, 101(7), 1049–1055.

Shi J., Joshin J., Tielbörger K., Verhoeven K. J. F., and Macel M., (2018). Costs and benefits of admixture between foreign genotypes and local populations in the field. *Ecology and*

Evolution, 8(7), 3675–3684.

Star B., and Spencer,H. G., (2013). Effects of genetic drift and gene flow on the selective maintenance of genetic variation. *Genetics*, 194(1), 235–244.

Zhou H., Alexander D., and Lange K.,(2009). Aquasi-newton acceleration for high-dimensional optimization algorithms.Stat Comput (2011), 21, 261–273.

Chapter 5: Genetic assessment of Enbu, *Morus* sp. through whole genome sequencing

5.1 Introduction

Mulberry has become an economic crop not only for silkworm rearing in the silk production industry but also as a functional plant (Eo et al., 2014; Machii et al., 2000; Niratker et al., 2015; Panche et al., 2016; Sánchez, 2000; Thaipitakwong et al., 2018). Its commonly grown for rearing of silkworm with added advantage for medicinal, forage and utility for livestock, poultry and human. The derived products include use in confectionaries (sweets), condiments (jam) and making soft drinks (juices, teas) and liquor as well as the wood utilised in making soft sports equipment's such as hockey sticks (Kole, 2011). The crop can withstand different growing conditions making it widely distributed all over the world (Machii et al., 2000).

Taxonomists have classified the genus under Moraceae family based on the female style, stigma with hairs or protrusions, supplemented by branch, leaf, flower and fruit shape description (Datwyler and Weiblen,2004) a method involving the morphological and phenological identification that is difficult, ambiguous, time-consuming and subjective. Being subjective, it has created unending debate further due to mulberry's high heterozygosity status (Vijayan et al., 2014a) hence need to continuously invest in understanding the existing varieties over time using latest technologies available.

The advent into genetic relationships assessment in crops has played a major role in plant breeding and genetic conservation resulting in selection of parents for breeding from a genetically divergent group through isolating sexually compatible parents to produce heterosis. Interesting, mulberry as a plant has shown no issues of incompatibility due to high interspecific hybridization that results to viable seed (Aversano et al., 2012; Awasthi et al., 2004; Tikader

and Kamble, 2008). This may have been key reason that retarded the urgency into its genetic research coupled with not being a model plant.

In recent years, the growth from the Sanger sequencing to current third generation sequencing technologies has enhanced exploration of information and reduced the costs of sequencing making it easy for adoption. For example the high throughput sequencing platform has offered an avenue in the field of ecology and evolution enabling the understanding of sciences over time (Ellegren, 2014) and as such accessibility of information has increased interest in genome sequencing of organisms with aim of understanding their evolution patterns, gene families while encompassing both the non-model organisms, domesticated horticultural crops and wild species.

Sequencing of eukaryotic genomes has been slow with less research investment as they presented a clear challenge over time because of their complex repetitive structures, heterozygous positions, insertions/deletions polymorphism, copy number variation as well as small-scale re-arrangements thus necessitating for proper assembly as well as functional deductions through annotations (Appels et al., 2018). Apparently, the complex plant genetic variations in eukaryotes has been as a result of the different transposable elements (microsatellites, SSRs, SNPs and segmented insertions and deletions (Indels) triggered by the cellular mechanisms that allow for regulation of interspecific hybridization, polyploidization and genome change during the process of meiosis and mitosis. To address the above challenges, the high throughput sequencing methods such as hybridization-based markers like RFLP, PCR based markers like RAPD; AFLP and SSR as well as SNPs have been used (Mammadov et al., 2012) because of the abundance of SNPs in the genomes that can be utilized by the high throughput detection platforms yielding results.

Furthermore, the environmental changes have been cited as the main cause of increased speciation with influence on the adaptation and alteration of the ploidy in chromosomes of

some crops and as such key monitoring is needed to check the alterations ((Adams and Wendel, 2005; De Storme and Mason, 2014); This has been demonstrated in mulberry ploidy where chromosome numbers are varied and complex ranging from 28 for most diploids to 308 in Kuromiguwa, a variety of *M. nigra* (Sharma et al., 2000; Koidzumi, 1917).

Previous studies placed the basic number of chromosomes as 14 with species having 28 chromosomes dominating the genus *Morus* (Venkatesh and Munirajappa, 2015). However, a current study suggested that the basic number of chromosomes is seven, prompting the question that how important the whole genome sequencing is in success of mulberry production and identification (He et al., 2013). Furthermore, the first linkage map of mulberry has been developed using randomly amplified polymorphic DNA (RAPD), inter-simple sequence repeat (ISSR), simple sequence repeat (SSR) markers as well as pseudo-testcross mapping strategy (Venkateswarlu et al., 2006). In addition, other comparative studies based on the strategies of using the complete chloroplast genome sequences and inferences have equally played a role to existing databases, although more studies are required to ascertain the phylogenetic relationship among the mulberry varieties (Li et al., 2016; Kong and Yang, 2017).

Strides have been achieved especially with the unveiling of the first draft genome sequence from *Morus notabilis* that has enriched the *Morus* database (Morus DB) and opened new avenues for further study (He et al., 2013; Li et al., 2014). Hence by adopting to genomic sequencing methods there is an opportunity that will lead to clear understanding of this genus in terms of classification and enhancing utilization of the crop better. With the above in mind there was need to understand the lineage of 'Enbu' variety that has been registered as *M. mesozyngia* in Japan yet its characteristics do not match.

5.2 Material and methods

5.2.1 Plant materials

All materials used for the study were collected from the field of NARO Genebank at Tsukuba, Japan. Young leaves of *Morus* sp. 'Enbu' cultivar were extracted and used for the genome sequencing.

5.2.2 Genomic DNA extraction and 'Enbu' sequencing

Young leaves of 'Enbu' were washed with distilled water twice and then drained water by a paper towel. The genomic DNA were extracted from the samples by combined CTAB (cetyltrimethylammonium bromide) method (Lipp et al., 1999) and benzyl chloride method (Zhu et al., 1993) as described below. To the samples, 1000 µL 1.5x CTAB buffer (1.5% CTAB 75mM Tris/HCl (pH 8.0), 15mM EDTA (pH 8.0), and 1.05M NaCl), 20 µL 2 M DTT, and 5 µg polyvinylpolypyrrolidone were added per 100 mg of sample. The samples were grinded with mortar and pestle, and then 500 µL of the samples were transferred to 1.5 mL tubes, respectively. The tubes were incubated at 65˚C for 1 hour. To the tubes, 100 µL benzyl chloride were added and mixed well. The tubes were incubated at 65˚C for 15 minutes (mixed well every 5 minutes), and then centrifuged at 14000 x g, 10 minutes, 10˚C. The upper water layer of each tube was transferred to a new 1.5 mL tube, respectively. Then 50 µL 10% CTAB solution (10% CTAB and 0.7M NaCl) were added to the new tubes and mixed well. To the tubes, 100 µL chloroform were added and mixed well. After centrifuging the tubes at 14000 x g, 10 minutes, 10˚C, the upper water layer of each tube was transferred to a new 1.5 mL tube, respectively. Then 750 µL CTAB precipitation buffer (1% CTAB, 50mM Tris/HCl (pH 8.0), and 10mM EDTA (pH 8.0)) were added to the new tubes and mixed well. The tubes were centrifuged at 14000 x g, 10 minutes, 25˚C, and then supernatant was removed. Added 100 µL 1 M NaCl solution to a pellet and resolved at 65˚C for 1 hour, and then some of the potion (e.g. 500 uL) were combined to each tube, respectively. Then 0.6 volume of isopropanol were added

to the tubes and mixed well. After centrifuging the tubes at 14000 x g, 10 minutes, 25˚C, supernatant was removed. Then ethanol wash was performed as follows. First, 1 mL 75% ethanol were added to the tubes and mixed. Then centrifuged the tubes at 14000 x g, 1min, 25˚C, and supernatant was removed. The ethanol wash was repeated and then all of supernatant was removed with pipet. Added 200 µL TE (10mM Tris/HCl (pH 8.0) and 1mM EDTA (pH 8.0)) (RNaseA about 200ng/mL conc. were included) to the tubes and then resolved at room temperature. The tubes were merged and then DNA concentration and condition were checked by agarose gel electrophoresis. Finally, precipitation of DNA with Isopropanol and DNA pellet was resolved by 10mM Tris HCl buffer (pH 8.5).

5.2.3 Genome sequencing and assembly

Sequencing of 'Enbu' cultivar using the obtained DNA was outsourced. Sequencing using Illumina HiSeq X for short reads was perfomed by Macrogen Japan and PacBioRS II for long reads was performed by Novogen. Genome assembly using the long reads was done using Canu assembler (v. 1.4) with polishing by Quiver software using long reads and by Pilon software (v. 1.17) using the short reads derived from the HiSeq short reads (Koren et al., 2017). The assembled contigs were then further improved by scaffolding using Chicago method (Chicago library construction, HiSeq X sequencing, scaffolding by HiRise software, and gap filling by long reads) performed by Dovetail Genomics. Redundancy of the improved scaffolds were further removed by Purge Haplotigs software (Roach et al., 2018).

5.2.4 Quantification of completeness of the genome assembly

Quantity of the 'Enbu' genome assembly was evaluated using software of benchmarking universal single copy orthologs (BUSCO) version 2.0.1 (Simão et al., 2015) with the lineage of embryophyta_odb9 that benchmarked 1440 single copy orthologs for

comparison. The result of BUSCO software for the ‘Enbu’ genome assembly was then compared with those of *Morus notabilis* ‘Chuansang’ obtained from MorusDB.

5.2.5 Gene prediction and functional annotation

De novo repeat sequences of the genome assembly was created using RepeatModeller (v. 1.0.11). Masking of repetitive sequence of the genome assembly for hint data of gene prediction was done using RepeatMasker (v. 4.0) using the de novo created repeat sequences by RepeatModeler. I mapped short reads generated by RNA-seq obtained from 10 tissue samples of ‘Enbu’ with STAR (v. 2.5.3) (Doblin et al., 2013) to generate hint data for gene prediction. Gene prediction for the genome assembly using the hint data was done using AUGUSTUS (v. 3.2.3) (Stanke and Morgenstern, 2005). The functional annotation of the predicted genes was obtained by comparing the predicted genes with NCBI-nr protein database by BLASTP software, with Gene Ontology and InterPro databases by InterProScan, with Kyoto encyclopaedia of genes and genomes for KO terms by KAAS (KEGG Automatic Annotation Server). The same functional annotation was done or the ‘Chuansang’ a (*M. notabilis*) variety and the results of ‘Enbu’ and ‘Chuansang’ further compared for analysis.

5.3 Results

5.3.1 Genome sequencing and assembly

Total size of 21.7 Gb (coverage of 62x) of PacBio long reads of 'Enbu' genome were generated. The number of long reads (subreads) was 2,941,707 and average read length was 7,386 bp. Total size of 54.7 Gb (156x) and 82.7 Gb (236x) of HiSeq X short reads (150 bp paired end) of 'Enbu' genome was also generated for insert length 350 bp and 550 bp respectively. Result of genome assembly by Canu assembly followed by polishing scaffolding by Chicago method and removing redundancy by purge haplotigs is shown in **Table 7.** The N50 and maximum scaffold length of 'Enbu' genome assembly were highly improved compared with existing 'Chuansang' mulberry genome assembly (**Table 7**). Average scaffold length and number of gap length were also highly improved (**Table 7**).

Table 7. Statistics of 'Enbu' genome assembly and 'Chuansang' genome assembly

	'Enbu' genome assembly	'Chuansang' genome assembly (He et al., 2013)
Number of sequences	740	110,760
Total length of genome (bp)	349,918,621	330,791,087
N50 (bp)	1,175,103	390,115
Avg. Scaffold Length (bp)	472,863	2,986
Max. Scaffold length (bp)	11,232,081	3,477,367
Gap length (bp)	452,304	16,281,102
GC level %	34.48	35.02
Repeat masked bases (bp)	164,877,816	127,983,338

5.3.2 Quantification of completeness of the genome assembly

Comparison of the completeness of newly sequenced Enbu against the existing *M. notabilis* '*chuansang'* (He et al 2013) was done using bench marking single copy orthologs (BUSCO). 1440 single copy orthogs standard to embryophyta were searched against to establish the

completeness based on complete (C), compete/duplicated (D), fragmented (F)or missing (M) ones. The assembled enbu genome was found to be of high quality assemble with slightly higher complete and duplicated BUSCOs than *M. notabilis* 'Chuansang*'*.

The completeness of the 'Enbu' genome assembly evaluated and compared with ('Chuansang') *Morus notabilis* by BUSCO2 (**Fig. 25**). Showed similarity in the completeness of the data set number of covered singly copy orthologs were shown in 'Enbu' and 'Chuansang' genome assembly. Outcome of the result indicated good completeness of the 'Enbu' genome assembly on a Enbu to Chuansang sequence data comparison while benchmarking on the standard models of *Homo sapiens, Arabidopsis thaliana, Fragaria vesca, Malus domestic*a, and *Bombyx mori*.

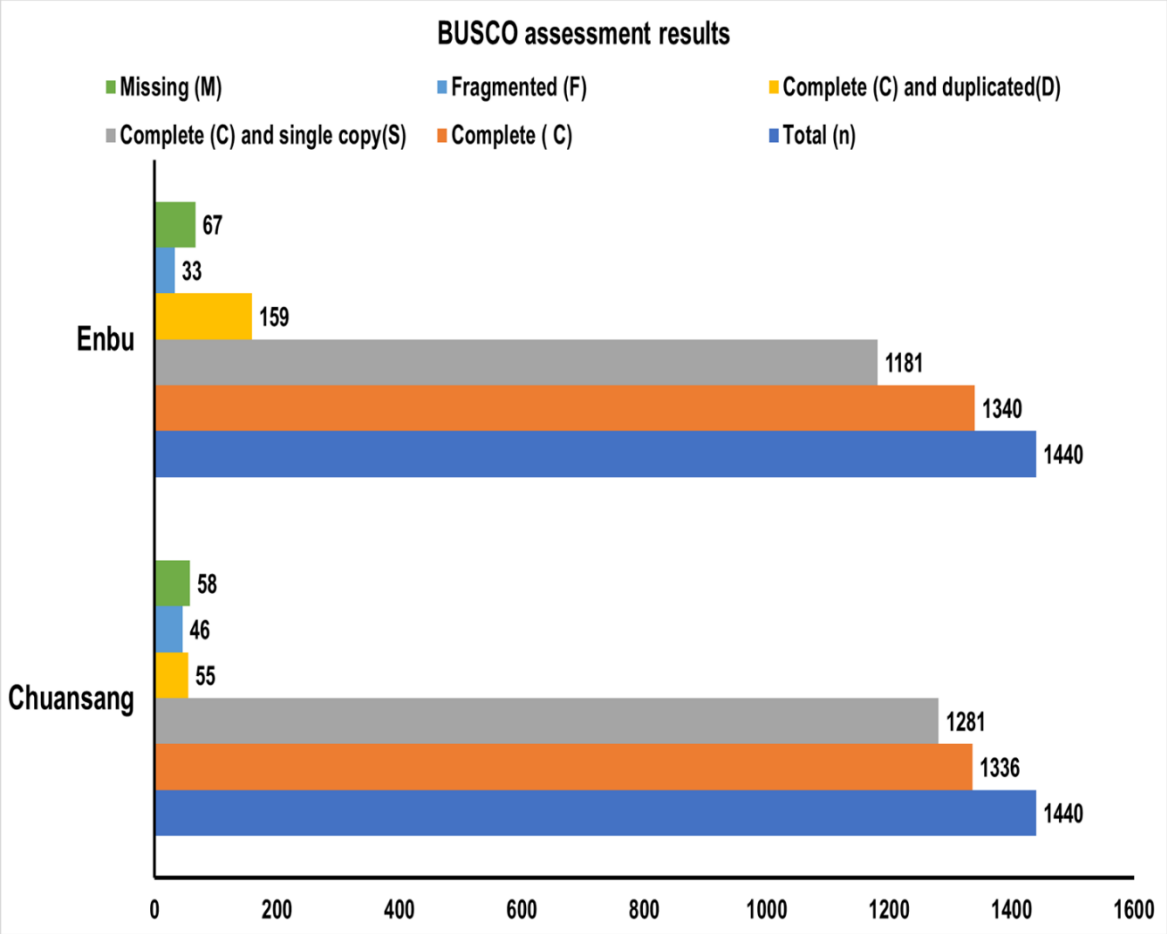

Figure 25. Completeness result obtained by BUSCO2 for orthologs of 'Enbu' and 'Chuansang' genome assembly

5.3.3 Gene prediction

Gene prediction by AUGUSTUS (v 3.2.3) using hint data of RNA-seq data and masked repeat regions of 'Enbu' genome assembly identified 35,483 predicted gene sequences in 32,989 gene loci. The average length of the predicted gene sequences was 1,059 bp.

5.3.4 Gene annotation

Results of functional annotation for NCBI-nr (by BLASTP), GO term and InterPro ID by InterProscan, KO terms by KAAS was shown in **Table. 8**. Assignment of KO terms was done by searching against the genome of ('Chuansang') *M. notabilis*. Total genes assigned with

KO terms in 'Enbu' variety was 8,456genes with 3709 KO terms. Of these 3,626KOs were shared with genes of ('Chuansang') *M. notabilis* and 298 KOs were 'Enbu' specific (**Fig. 26**).

Table 8. Assignment of the gene sequences using different databases

Functional gene annotation	No of genes	% hits
Total gene sequences	35,483	
No of sequences with blastp hit (NCBI-nr)	32,404	91.32
No of sequences with InterProID	25,674	74.36
No of sequences with Gene Ontology	18,892	53.24
No of sequences assigned KO terms	8,077	22.76

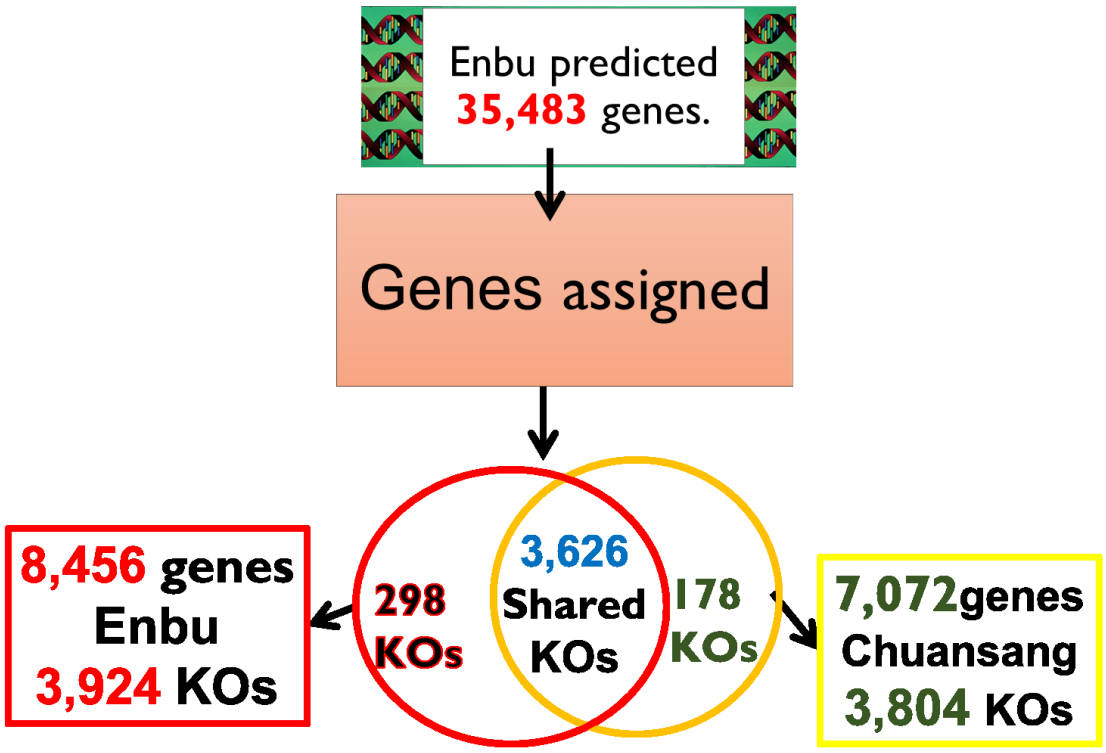

Figure 26. Results of KO terms annotation obtained from both the predicted genes in 'Enbu' and 'Chuansang'.

5.3.5 Discussion on genetic assessment of Enbu

Breeding and development of plant research is limited by availability of information and as such in order to meet the utilization of mulberry bioresource through research acceleration of genomic information is of paramount importance in order to advance production, conservation and increasing adaptable utility. The completeness of the genome was verified using 1,440 BUSCOs of the embryophyte profile of which we compared both the 'Enbu' and 'Chuansang'. The results of completeness show similarity with slight difference in the duplicated single copy orthologs which could be as a result of the gene space variability among the two varieties belonging to two different species. Furthermore, from the assembly, a well assembled genome sequence of 'Enbu' of total size 349,918,621 bp was obtained offering an opportunity to establish markers of importance. The genome further resulted to unveiling of 35,489 genes that were predicted and annotated that could drive the innovations in the 'Enbu' research while offering a platform for comparison in future. Out of these, 8,456 genes of 'Enbu' were assigned the KO terms through the KEGG automated annotation server resulting to 298 specific term being isolated above the 7,072 KO terms assigned to 'Chuansang' and a sharing of 3,626 KO terms among 'Enbu' and 'Chuansang'. Overall, from the predicted 35,483 annotated genes of Enbu and 298 KO terms located in 340 gene that are Enbu specific offers opportunity to decipher which of the genes might be involved in important features offers future continuation of research. The above results reveal that the genome study 'Enbu' therefore may play a key role in establishing the markers for use in the mulberry production. Furthermore, the continued analysis may yield markers that may be used in finding out the relationships existing in the plant insect interactions.

5.3.6 Reference

Adams K. L., and Wendel J F., (2005). Polyploidy and genome evolution in plants. *Current*

Opinion in Plant Biology. 8 (2), 135-141.

Appels R., Eversole K., Feuillet C., Keller B., Rogers J., Stein N., Pozniak C. J., Choulet F., Distelfeld A., Poland J., Ronen G., Barad O., Baruch K., Keeble-Gagnère G., Mascher M., Ben-Zvi G., Josselin A. A., Himmelbach A., Balfourier F.,…and Wang L. (2018). Shifting the limits in wheat research and breeding using a fully annotated 361(6403), eaar7191.

Aversano R., Ercolano M.R., Caruso I., Fasano C., Rosellini D., and Carputo D., (2012). Molecular tools for exploring polyploid genomes in plants. *International Journal of Molecular Sciences.* 13(8), 10316–10335.

Awasthi A.K., Nagaraja G.M., Naik G. V., Kanginakudru S., Thangavelu K., Nagaraju J., (2004). Genetic diversity and relationships in mulberry (genus *Morus*) as revealed by RAPD and ISSR marker assays. *BioMed Central Genetics,* 5(1), 1–9.

Datwyler S. L., and Weiblen G. D., (2004). On the origin of the fig: Phylogenetic relationships of Moraceae from NDHF sequences. *American Journal of Botany*, 91(5), 767–777.

De Strome N.,and Mason A., (2014). Plant speciation through chromosome instability and ploidy change:Cellular mechanisms,molecular factors and evolutionary relevance.Current Plant Biology, 1, 10-33.

Dobin A., Davis C. A., Schlesinger F., Drenkow J., Zaleski C., JhaS., Batut P., Chaisson M., Gingeras T., (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29(1), 15-21.

Ellegren H., (2014). Genome sequencing and population genomics in non-model organisms. In *Trends in Ecology and Evolution*, 29(1), 51–63.

Eo H.J., Park J.H., Park G.H., Lee M.H., Lee J.R., Koo J.S., Jeong J.B., (2014). Anti-inflammatory and anti-cancer activity of mulberry (*Morus alba* L.) root bark. *BioMed*

Central Complementary Alternative Medicine, 14(1), 200-209.

He N., Zhang C., Qi X., Zhao S., Tao Y., Yang G., Lee T., Wang X., Cai Q., Li D., Lu M., Liao S., Luo G., He R., Tan X., Xu Y., Li T., Zhao A., Jia L., … Xiang Z. (2013). Draft genome sequence of the mulberry tree *Morus notabilis*. *Nature Communications*, 4(1), 2445-2454.

Koidzumi C., (1917). Taxonomical discussion on morus plants. .Bull. Imp. sericult.Exp.Stat. Japan 3, 1-62.

Kole, C. (Ed.). (2011). *Wild Crop Relatives:Genomic and breeding resources: Tropical and subtropical fruits.* Springer Heidelberg Dordrecht London New York.

Kong W Q., and Yang J.H., (2017). The complete chloroplast genome sequence of *Morus cathayana and Morus multicaulis a comparative analysis within genus Morus L. Peer Journal* 62(1), 165–172.

Koren S., Walenz B. P., Berlin K., Miller J. R., Bergman N. H., and Phillippy M.. A., (2017). CANU: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, 27, 722–736.

Li Q., Guo J., Yan N., and Li C., (2016). Complete chloroplast genome sequence of cultivated *Morus* L. species. *Genetics and Molecular Research*. 15(4), 60-73.

Li T., Qi X., Zeng Q., Xiang Z., and He N., (2014). MorusDB: A resource for mulberry genomics and genome biology. *Database(Oxford)*, 2014, bau054.

Lipp M., Brodmann R., Pietsch K., Pauwels J Anklam E., Borchers T., Braunschweiger G., Busch U., Eklund E., Eriksen F.D., Fagan J., Fellinger A., Gaugitsch H., Hayes D., Hertel C., Hortner H., Joundrier P., Kruse L., Meyer R., Maraglia M., Muller W., Phillipp P., Popping B., Rentsch R., and Wurtz A., (1999). IUPAC collaborative trial study of a method to detect genetically modified soybean and maize in dried powder. *Journal of*

AOAC international, 82, 923-928.

Machii H., Koyama A., and Yamanouchi,H., (2000). Mulberry breeding, cultivation and utilization in Japan. *FAO Electronic Conference on Mulberry for AnimalProduction* 8, 63-72.

Mammadov J., Aggarwal R., Buyyarapu R., and Kumpatla S., (2012). SNP markers and their impact on plant breeding. *International Journal of plant genomics*, *2012*, 387–413.

Niratker, C. R., Preeti, and Malti. (2015). Antimicrobial activity of leaf extract of Morus indica (Mulberr y) from Chhattisgarh. *Asian Journal of Plant Science and Research*, *5*(1), 28–31.

Panche A. N., Diwan A. D., and Chandra S. R., (2016). Flavonoids: an overview. *Journal of Nutritional Science*, 5(e47), 1-15.

Roach M. J., Schmidt S. A., and Borneman A. R., (2018). Purge Haplotigs: Allelic contig reassignment for third-gen diploid genome assemblies. *BioMed Central Bioinformatics*, 19(1), 460-470.

Sánchez M. D. (2000). Mulberry: an exceptional forage available almost worldwide. *World Animal Review*, 93(1), 1–21.

Simão A., Waterhouse R. M., Ioannidis P., Kriventseva E.V.,and Zdobnov E.M (2015). BUSCO:Assessing genome assembly and annotation completeness with single-copy orthologs, Bioinformatics, 31(19), 3210–3212.

Sharma A, Sharma R, and Machii H., (2000). Assessment of genetic diversity in a *Morus* germplasm collection using fluorescence-based AFLP markers. Theoretical and Applied Genetics, 101(7), 1049–1055.

Stanke M. and Morgenstern B., (2005). AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research*, 33, W465-W467

Tikader A. and Kamble C.K., (2008). Mulberry wild species in India and their use in crop improvement – A review. *Australian Journal of Crop Science,* 2(2), 64-72.

Thaipitakwong, T., Numhom, S., and Aramwit, P. (2018). Mulberry leaves and their potential effects against cardiometabolic risks: A review of chemical compositions, biological properties and clinical efficacy. *Pharmaceutical Biology*, 56(1), 109–118.

Vijayan, K., Raju, P. J., Tikader, A., and Saratchnadra, B. (2014a). Biotechnology of mulberry (*Morus* L.) - A review. In *Emirates Journal of Food and Agriculture*, 26(6), 472–496.

Venkatesh K. H. and Munirajappa., (2015). Comparative Micro-morphological and Karyomorphological studies in three Mulberry Varieties (*Morus* spp.) *Cytologia* 80(1), 31–35.

Venkateswarlu M., Urs S. R., Nath B. S., Shashidhar H. E., Maheswaran M., Veeraiah T. M., and Sabitha M. G., (2006). A first genetic linkage map of mulberry (*Morus* spp.) using RAPD, ISSR, and SSR markers and pseudotestcross mapping strategy. *Tree Genetics and Genomes*, 3(1), 15–24.

Zhu H., Qu F., and Zhu L.H., (1993). Isolation of genomic DNAs from plants, fungi and bacteria using benzyl chloride, *Nucleic Acid Research*, 21(22), 5279-5280.

Discussions

My finding reveals that despite morphological characterisation being indispensable, it's a challenge in highly heterozygous varieties as ease of crosses results to new breeds with similar phenotypes yet their genetic makeup might be different due to recombination effects. This finding is not far from the ordinary since the classification of mulberry over time has elucidated a great debate on the species and synonyms. The plant phenotype may change depending on climatic conditions and the habitat in which they are exposed to over time. For example, based on the climatic conditions of Brazil that ranges from hot and dry in the arid interior to humid and sticky in the tropical rainforest and rarely falls below 20^0C, it has been observed that the varieties 'Kokuso 20' and 'Kokuso 21' behave differently when compared to their production in Japan. In Japan, 'Kokusou 20' and 'Kokusou 21' have been found to produce either female inflorescence or both male and female respectively while in Brazil they both produce only male inflorescence placing emphasis on environmental effect (Basiolo et al., 2004).The presence of different tier classification strategy is also emerging in the Moraceae family with many varieties being bred from the global perspective of 7 species (Linneaus, 1753) to regional classification and even local where the individual researchers may deem to have discovered new species or merged based on synonyms in the international plant names index. Overall, the leaf morphological features, longest shoot and buds have played a major taxonomical role in the family of Moraceae (Machii et al., 2001).

The present phylogeny and admixture analysis results involving relationships among the populations reveal information on the wild native species where the wild species clearly show separation from the naturalised introduced varieties which may have been a result of either their origin, isolation distance or the combined efforts of hybridization among others (**Fig. 17-18**) as well as (**Fig. 22-24**).These results indicate that molecular markers, and especially the use of SNPs should be embraced as they have great potential in separating the

closely related species although in some cases of highly hybridised varieties, a challenge may arise in clearly separating them resulting to admixtures. My results through phylogenetic analysis have revealed that both monophyletic and non-monophyletic clades exist in mulberry based on the 54 varieties evaluated with clarity being observed in two native species in Japan and an introduced species from Thailand. The above outcome could also be aided by isolation thus avoiding hybridization among the varieties unlike the other three species that formed unclear monophyletic clades with *M. alba* and *M. latifolia* having been considered highly hybridized thus resulting to no clear clades on their own. *M. bombycis* varieties evaluated were grouped into three categories with two forming monophyletic clades with *M. acidosa* and *M. kagayamae* respectively. This could be also an influence of environment and isolation where the native *M. bombycis* has been identified with colder regions and the ones that are segregated could be a result of hybridization making them adaptable to the wide environment of subtropical climate as seen in distribution areas (**Figure 15)**.

Further results showed that two monophyletic clades were formed by *M. indica* an indication of two different ancestry of which 'Enbu' an indigenous variety from Kenya was able to form a monophyletic clade with one of the *M. indica* clades signifying a possible common background based on genetic relationship. As much as the result may not be far-fetched due to the fact that the first introductions of mulberry were introduced in early 1905 by the British administration and trials thereafter conducted on the viability could have Indian origin (Eliot's, 1905),the admixture analysis results still show that there could be more ancestral mixtures within this variety (**Fig. 24)** hence need for further analysis to unravel the mystery.

Genome sequencing having brought a revolution since the first human genome project has gone through a transition from the first generation, second and currently the next generation sequencing (NGS) that is based on the amount of time available, labour input

and the cost implication as factors for choice of technology. Despite genome sequencing having been a complicated technology and out of reach for many upcoming researchers without funding, it has now offered an ease way of protein coding of genes that are sparsely distributed and interrupted by introns with a major problem of identifying the key exons and ability to perform the proper assembly to achieve the quality genome. Adapting the methodology has opened room to study 'Enbu' genome for the target of understanding in-depth makeup of the variety and its traits through its DNA information.

Impact of the research will result to further analysis and development of the molecular markers and SNPs in the above research to facilitate the improvement of mulberry research, through generation of information that has linked the relationship of unknown variety 'Enbu' classification by tracing the likely background. Furthermore, the utilization of the predicted genes in future will result to production, maintenance, breeding varieties with potential to withstand abiotic and biotic stresses thus enhancing the whole complete value chain of mulberry, silkworm rearing and ultimately silk production.

In addition to comparison research of mulberry gene sequences that exist currently in DDBJ for the 54 mulberry varieties belonging to the 8 species used in this study, the information on *M. notabilis* in MorusDB as well as information that is currently with different research groups in the world would enhance milestones in crop protection aspects and research while considering the relationship between the plant and the insect interactions. The detection of various SNPs can be used to enhance monitoring and management of varieties and cultivars thus improving the conservation strategies. The results obtained based on the proper allocation of biological information will result to proper allocation of the conservation resource towards preserving habitats and species while breeding new varieties.

References

Biasiolo M., Teresa M., Canal D. A., and Tornadore N., (2004). Micromorphological characterization of ten mulberry cultivars *(Morus* Spp*.)*, *Economic Botany,* 58(4), 639–646.

Eliots C., (1905). The East Africa Protectorate.Arnold, London.

Linneaus C.(1753). *Morus* in Species plantarum. Stockholm, Sweden, 2, 986.

Machii H, Koyama A., Yamanouchi H., Matsumoto K., Kobayashi S., and Katagiri K., (2001). A list of morphological and agronomic traits of mulberry genetic resources. Miscellaneous Publications of the National Institute of Sericulture and Entomological Science, 29, 1-307 ISSB 0915-2679.

Acknowledgments

To God be the glory and honour for the opportunity and blessings he has accorded me in terms of my health, my studies, my research environment and all the wonderful people I have interacted with during my three years study program at Tokyo University and research station at NARO.

It has been a long journey that has taken a lot of planning , preparations and organization by Professor Kameda Tsunenori who is the team leader of the knowledge co-creation and development project of Science and Technology Research Partnership for sustainable development program referred to as (SATREPS), that has seen me through by funding for my study and research in collaboration with the Government of Japan through the Japan science and technology agency (JST) and Japan international cooperation agency (JICA).

Secondly, I am indebted to my advisor and Supervisor Professor Hideki Sezutsu for the personalised guidance, training and leadership till this juncture I finish my studies. The above could not be accomplished without the support of my Committee members Professor Shoji Kawamura and Associate Professor Kazuhiro Nakayama both from The University Tokyo for their great inputs and guidance in many aspects of population genetics and especially admixture analysis, views on phenotypic diversity and human movement as well as colour expression and colour blind test for presentation for made me to understand the way to improve on my work and especially on the aspects of population genomics. Furthermore, my special gratitude goes to Professor Rensuke Kanekatsu of Shinshu university for extended guidance during my review as well as highlighting some of his work that is related to my research for wider understanding and hopefully future collaboration.

My objectives were met through collaboration and able guidance of Senior researchers of NARO; Dr. Yamanouchi Hiroaki, Dr. Jouraku Akiya, Dr. Yang chia-cha, Dr. Yokoi Kakeru, Dr. Kuwazaki Seigo, Dr. Yamamoto Kimiko who mentored and moulded me in skills ranging from ex-situ gene bank of mulberry, Laboratory skills, Genomics, bioinformatics and leadership. I am really indebted to their hard work and sacrifice that has made me a better person and may the work of your hands be blessed.

In the daily routine, I have been supported by a team that made sure my stay and mental health and need were catered for. They made me smile and lift the weight off my shoulders and they will remain in my circles for collaborations and friendship.

Further acknowledgements go to different organisations, research institutions and laboratories under which my research was supported; the University of Tokyo (UTOKYO) for enabling my study programme, my employer Kenya Agricultural livestock and research organisation (KALRO-Kenya) for allowing me to advance my career and National Agricultural

Research Organisation (NARO-Japan) for offering me their open-door research organisation to conduct my research under their gene bank, Laboratory of Insect Genomics Research, and Laboratory of silk and biomaterials.

Special thanks go to Dr. Lusike Wasilwa and Dr. Muo Kasina for their mentorship support and drive towards my study. The Horticulture research Institute Director, Dr. Waturu, Dr. Mbaka and my colleagues for enabling and guiding me in my research growth.

To Dad, Mr. Francis Omega, may God rest you in eternal peace, you make me learn to be strong and stand on my feet. To my Mother Mrs Alice Miraho Isaya. who believed in letting me grow wings and fly, by sacrificing all for me in prayer, financially and advice among others just to make sure I had a better life. She is the strength that pushes me today. To my seven brother's and family thank you for the greatest support and making me a leader God bless you. To my friends and especially Sheila, Jackie and Winnie,you kept encouraging and handling my every situation for me when I was about to break down.