PLOS ONE

RESEARCH ARTICLE

Phylogeny of *Morella rubra* and Its Relatives (Myricaceae) and Genetic Resources of Chinese Bayberry Using RAD Sequencing

Luxian Liu[1], Xinjie Jin[1], Nan Chen[1], Xian Li[2], Pan Li[1]*, Chengxin Fu[1]

1 Key Laboratory of Conservation Biology for Endangered Wildlife of the Ministry of Education, and Laboratory of Systematic & Evolutionary Botany and Biodiversity, College of Life Sciences, Zhejiang University, Hangzhou 310058, China, **2** Laboratory of Fruit Quality Biology/The State Agriculture Ministry Laboratory of Horticultural Plant Growth, Development and Quality Improvement, Zhejiang University, Hangzhou 310058, China

* panli_zju@126.com

OPEN ACCESS

Citation: Liu L, Jin X, Chen N, Li X, Li P, Fu C (2015) Phylogeny of *Morella rubra* and Its Relatives (Myricaceae) and Genetic Resources of Chinese Bayberry Using RAD Sequencing. PLoS ONE 10(10): e0139840. doi:10.1371/journal.pone.0139840

Editor: Genlou Sun, Saint Mary's University, CANADA

Received: July 14, 2015

Accepted: September 17, 2015

Published: October 2, 2015

Data Availability Statement: All RAD data are available on Dryad, doi:10.5061/dryad.589g3.

Funding: This work was supported by the Special Fun for Agro-scientific Research in the Public Interest (grant no. 201203089) http://www.moa.gov.cn/zwllm/, C.F. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Phylogenetic relationships among Chinese species of *Morella* (Myricaceae) are unresolved. Here, we use restriction site-associated DNA sequencing (RAD-seq) to identify candidate loci that will help in determining phylogenetic relationships among *Morella rubra*, *M*. *adenophora*, *M*. *nana* and *M*. *esculenta*. Three methods for inferring phylogeny, maximum parsimony (MP), maximum likelihood (ML) and Bayesian concordance, were applied to data sets including as many as 4253 RAD loci with 8360 parsimony informative variable sites. All three methods significantly favored the topology of (((*M*. *rubra*, *M*. *adenophora*), *M*. *nana*), *M*. *esculenta*). Two species from North America (*M*. *cerifera* and *M*. *pensylvanica*) were placed as sister to the four Chinese species. According to BEAST analysis, we deduced speciation of *M*. *rubra* to be at about the Miocene-Pliocene boundary (5.28 Ma). Intraspecific divergence in *M*. *rubra* occurred in the late Pliocene (3.39 Ma). From pooled data, we assembled 29378, 21902 and 23552 de novo contigs with an average length of 229, 234 and 234 bp for *M*. *rubra*, *M*. *nana* and *M*. *esculenta* respectively. The contigs were used to investigate functional classification of RAD tags in a BLASTX search. Additionally, we identified 3808 unlinked SNP sites across the four populations of *M*. *rubra* and discovered genes associated with fruit ripening and senescence, fruit quality and disease/defense metabolism based on KEGG database.

Introduction

Many domesticated fruit trees, such as peach (*Prunus persica* (L.) Batsch), plum (*Prunus salicina* Lindl.), kiwifruit (*Actinidia deliciosa* (A. Chev.) C. F. Liang & A. R. Ferguson and *A. chinensis* Planch) and persimmon (*Diospyros kaki* L. f.) originated in China [1–3]. Among them, some were derived from a single wild species, while others involved multiple species. In addition, some fruits even experienced a much more complicated domestication process. For example, Xu et al. [4] presented evidence to suggest that sweet orange [*Citrus sinensis* (L.) Osbeck] originated from a backcross hybrid between pummelo [*C*. *grandis* (L.) Osbeck] and mandarin (*C*. *reticulata* Blanco). Regardless of how domestication occurred, an unambiguous phylogeny of domesticated species and their potential source is always helpful for understanding the processes involved. Chinese bayberry is one of the most popular fruits in southern China. It is generally thought to have been domesticated from the wild *M*. *rubra* Lour., a subtropical evergreen tree with a wide distribution in China, Japan, Korea and the Philippines [5–6]. The domestication process has remained unresolved because the phylogenetic relationships among *M*. *rubra* and its relatives are still unclear.

The Myricaceae is a small, sub-cosmopolitan family of about 50 species, and the family of predominantly shrubs and trees are characterized by unisexual flowers borne in catkins, peltate glands, entire leaves, a unilocular ovary and single orthotropous ovule. With the exception of two monotypic genera, *Comptonia* L'Hér. ex Ait. and *Canacomyrica* Guillaumin, the species of Myricaceae have traditionally been referred to the Linnaean genus *Myrica* [7]. However, *Myrica* was split into two genera, *Myrica sensu stricto* and *Morella* Lour., based on morphological differences (deciduous or evergreen; dry fruits or fleshy fruits; sunken stoma or not) and phylogenetic analysis of nuclear ITS and chloroplast *trnL-F* sequence data [8]. In Myricaceae, only one genus and four species are in China; *M*. *rubra*, *M*. *nana* (A. Chev.) J. Herb, *M*. *adenophora* (Hance) J. Herb. and *M*. *esculenta* (Buch.-Ham. ex D.Don) I. M. Turner. *M*. *esculenta* and *M*. *adenophora* are easily recognized by their distinct tomentose branchlets and petioles, while *M*. *rubra* and *M*. *nana* are glabrous or sparsely pubescent [9]. Previous phylogenetic studies [8] strongly supported the monophyly of these four species, but the relationships among *M*. *rubra*, *M*. *nana* and *M*. *adenophora* were not resolved. The genetic diversity and population structure of wild *M*. *rubra* populations are poorly known. Insufficient knowledge of phylogenetics and population genetics will hinder the improvement of Chinese bayberry cultivars and the breeding of new ones.

Phylogeny reconstruction within closely related species may be difficult because of incomplete lineage sorting, introgression, short evolutionary scale, and lack of molecular markers in poorly studied taxa [10]. In this circumstance, reduced-representation genome sequencing methods allow us to sequence the regions flanking restriction sites with deep coverage, then to align orthologous sequences across multiple samples to discover thousands of genetic markers for systematics, population genomics and adaptive evolution studies [11–14]. These methods, including restriction site associated DNA sequencing (RAD-seq) and genotyping by sequencing (GBS), are promising and can be easily applied to non-model organisms with no reference genome sequence [15–16]. Hitherto, RAD-seq has been successfully applied to phylogenetic inference in *Pedicularis* [15], temperate bamboos [17], as well as population genomics in *Lagenaria siceraria* (Molina) Standl. [18], *Gasterosteus aculeatus* [19], and adaptive evolution in *Entosphenus tridentatus* [20], *Myodes glareolus* [16].

In Zhejiang, Fujian and Guangdong provinces, Chinese bayberry is one of the most popular and valuable fruits because of its appealing color, delicious taste and essential micronutrients [21]. The fruit is not only eaten fresh, dried and canned, but is also widely used for making wine and juice [22]. It also exhibits a wide range of pharmacological properties due to the high content of anthocyanins, which are reported to have anti-inflammatory, anti-tumor and anti-oxidative properties [23]. According to the literature, Chinese bayberry has been cultivated for more than 2000 years in southern China, and the cultivated area is currently 340,000 ha, with an annual yield of 1.1 million tons valued at 1.5 billion dollars [24]. It has long been a major source of income for farmers in some counties. During the long history of cultivation, more than one hundred cultivars, such as 'Dongkui', 'Biqi' and 'Wandao', were developed [25]. In recent years, Chinese bayberry has been exported to foreign countries and has received international attention due to its extraordinary qualities [26]. However, the development of Chinese bayberry is now confronted by huge challenges. Firstly, there is no consensus on the classification of cultivars, and they are classified only according to ripening date, fruit color, fruit weight and kernel characteristics, resulting in a high frequency of synonyms and homonyms [27]. Secondly, fruit quality declines rapidly at room temperature, which leads to a short shelf life [28–29]. Moreover, with the increase in planting area, Chinese bayberry suffers from a range of pathogens [such as *Phomopsis myricina* Y. J. Huang et P. K. Chi and *Leptographium abietinum* (Peck) M. J. Wingfield] [30–31]. Previous studies attempted to regulate fruit quality during ripening, to control postharvest fruit decay [32–33] and to enhance pathogen resistance [34]. However, knowledge of wild germplasm resources, phylogeny and the molecular basis of fruit quality and defense from disease is limited. The above problems will seriously inhibit the future development of the Chinese bayberry industry.

The objectives of the present study are 1) to resolve the phylogenetic relationships among *M*. *rubra* and its relatives, 2) to discover SNP markers within populations of *M*. *rubra* for future studies on genetic diversity and population structure, and 3) to determine if several genes or pathways are associated with fruit ripening and senescence, fruit quality and disease resistance.

Materials and Methods

Ethics statement

Three individuals from North America were permitted by Harvard University Herbaria (USBH) and NCSU Herbaria (USR). Management Bureau of Mt. Gutian National Nature Reserve issued the permit for Gutian Mountain (ZJGT); Management Bureau of Mt. Leigong National Nature Reserve issued the permit for Leigong Mountain (GZLS); Kunming Forestry Bureau issued the permit for sampling in Kunming (YNFM, YNXW, YNZJ, YNPL, YNHX and YNAL). No specific permissions were required for other locations which are neither privately owned nor protected and the field study did not involve endangered or protected species.

Sample collection and DNA extraction

Eighteen individuals, including six species of *Morella* as well as the closely related outgroup species, *Comptonia peregrine* (L.) Coult., were collected between 2012 and 2014 in China and North America (Table 1, Fig 1). The ingroup samples included one individual of *M*. *cerifera* (L.) Small, one of *M*. *pensylvanica* (Mirb.) Kartesz, four of *M*. *esculenta*, six of *M*. *nana*, one of *M*. *adenophora* and four of *M*. *rubra*. Total genomic DNA was extracted from silica-dried leaf tissue using the modified protocol of Doyle [35].

Acquisition and sequencing of the RAD libraries

Library preparation and sequencing of RAD markers from genomic DNAs were performed by Beijing Genomics Institute (Shenzhen, China) using the restriction enzyme *EcoR*I and sample-specific barcodes. The 18 individuals studied were first pooled and run in a single lane of an Illumina HiSeq 2000 with read length of 100bp, after which one individual of *C. peregrine* was sequenced for quality check, making 19 samples in total.

De novo assembly

To process the raw RAD-seq data for phylogenetic analysis, we utilized the pyRAD software [15]. Given one or more Ilumina sequence files in FASTQ format, pyRAD can de-multiplex the data and create separate files for each sample according to their special barcode. We usually filtered sequences through the following steps: First, sequences containing sequencing errors in the cut site were discarded. Second, reads containing sequencing errors in the sample-specific

Table 1. Details of location and sampling information for species of *Morella* investigated in this study.

Taxon	ID	Location	Latitude	Longitude	Altitude(m)
Morella rubra	ZJGT	Gutianshan, Zhejiang, China	N29°8'41.77"	E118°9'11.18"	455
	GZLS	Leigongshan, Guizhou, China	N26°30'1.04"	E106°44'2.82"	1135
	YNDW	Daweishan, Yunnan, China	N22°55'29.21"	E103°40'46.38"	2257
	YNML	Malipo, Yunnan, China	N23°8'16.67"	E104°41'23.71"	1171
M. adenophora	GXHX	Hengxian, Guangxi, China	N22°31'52.97"	E109°11'17.15"	74
M. nana	YNFM	Fuminxian, Yunnan, China	N25°18'21.31"	E102°41'24.88"	1958
	YNXW	Xuanweishi, Yunnan, China	N27°2'55.94"	E101°68'04.88"	2341
	YNWS	Wenshanxian, Yunnan, China	N23°19'27"	E103°13'34"	2583
	YNPL	Panlong, Kunming, Yunnan, China	N25°10'12.94"	E102°47'2.33"	2050
	YNHK	Haikouzhen, Kunming, Yunnan, China	N24°48'33.77"	E102°37'25.29"	1965
	YNZJ	Zhaojiaocun, Kunming, Yunnan, China	N25°05'13.64"	E102°34'10.37"	1888
M. esculenta	YNAL	Ailaoshan, Yunnan, China	N24°32'25"	E101°01'36"	2504
	YNWS	Wenshanxian, Yunnan, China	N23°19'27"	E103°13'34"	2583
	SCYB	Yanbianxian, Sichuan, China	N27°06'08.62"	E101°46'01.19"	2074
	SCMY	Miyixian, Sichuan, China	N27°02'55.94"	E101°58'04.88"	1952
M. pensylvanica	USBH	Harvard campus, Cambridge, MA, USA	N42°22'37.21"	W71°06'59.98"	8
M. cerifera	USR	NCSU campus, Raleigh, NC, USA	N35°47'04.79"	W78°40'55.54"	127
Comptonia peregrina	USR	NCSU campus, Raleigh NC, USA	N35°47'04.79"	W78°40'55.54"	127

doi:10.1371/journal.pone.0139840.t001

barcode were removed. The restriction site and barcode were then trimmed from each sequence. Bases with a FASTQ quality score below a given value (here, 33) were replaced with N, sequences having more than a given percentage of Ns (here, 1%) were discarded. Paired-end reads of the same species were pooled together and de novo contigs were assembled using Trinity Release v2.0.6 [36], run with a kmer length of 25bp, set the minimum contig size as 150bp and with other parameters set to default.

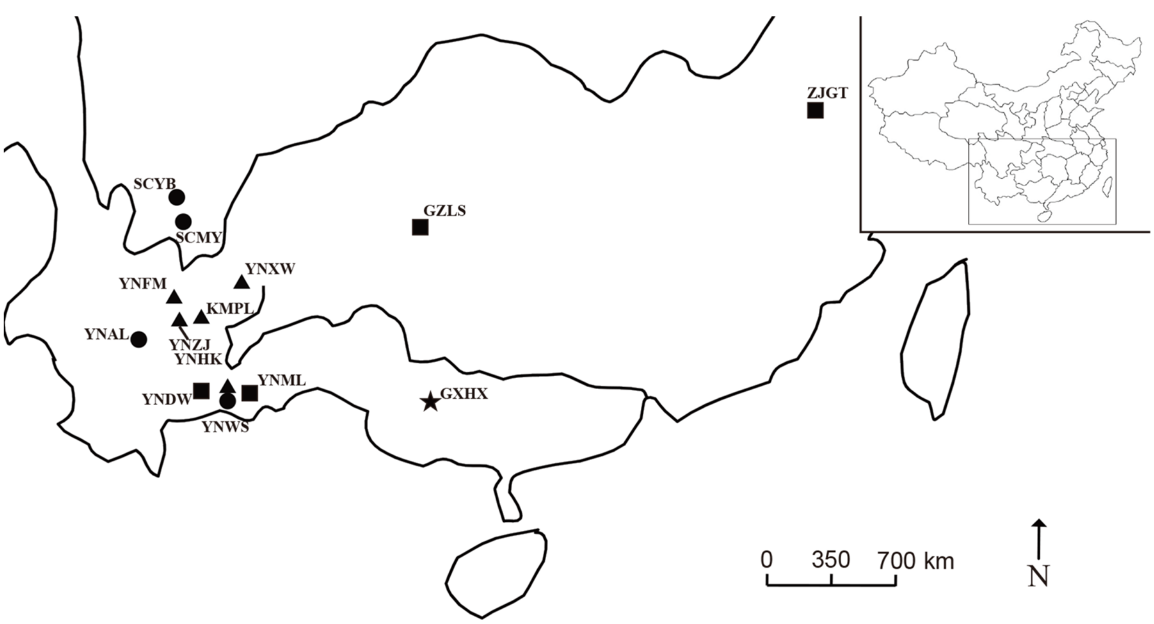

Fig 1. Map of sampling locations of *Morella* in China. Squares indicate *Morella rubra*; triangles indicate *M. nana*; dots indicate *M. esculenta*; star indicates *M. adenophora*.

doi:10.1371/journal.pone.0139840.g001

SNP discovery and phylogeny inference

To explore SNP markers for phylogenetic studies, we employed the pyRAD software, applied only to the single end of the paired-end sequences (R1). Because of the lack of a reference genome, sequence similarity is the simplest way to infer orthology. For each sample, sequences were clustered by similarity (here, 92%) using the uclust function in USEARCH [37] with heuristics turned off, yielding clusters representing putative loci. In order to ensure accurate base calls, clusters of fewer sequences than a set minimum depth of coverage (here, 6) were removed. The remaining clusters were then processed within pyRAD to generate consensus sequences. In pyRAD, the heterozygosity (H) and the error rate (E) are jointly estimated from the observed base counts across all sites in all clusters, by applying the maximum-likelihood equation of [38]. The mean E is then used to assign consensus diploid genotypes for each site in each cluster by calculating the binomial probability that the site is homozygous versus heterozygous given the relative frequencies of observed bases at the site and E [39]. If a base cannot be assigned with more than 95% probability, it is replaced by N in the consensus sequence. Heterozygotic variation is recorded using appropriate ambiguity codes.

Consensus sequences from all samples were clustered according to the sequence similarity using the same similarity threshold as in the previous step of within-sample clustering. For each cluster, sequences were aligned with Muscle v3.8.31 [40] with default parameters set. In order to detect potential paralogs, we set the maximum number (here, 3) of shared polymorphic sites in a locus, under the assumption that a shared heterozygous site across many samples likely represents clustering of paralogs with a fixed difference rather than a true heterozygous site [41]. The remaining clusters were treated as RAD loci and assembled into phylogenetic data matrices. For any given RAD locus, sequences of one or more samples may be missing if substitutions in the restriction site have disrupted recognition, or if the locus did not receive sufficient coverage for confident basecalling. To avoid the effect of missing data or insufficient information sites for phylogeny inference, the minimum taxa coverage (here, 12) of ingroup samples with data for a given locus to be retained in the final data set.

All phylogenetic analyses were conducted using the maximum-parsimony (MP), maximum-likelihood (ML) and Bayesian methods. Maximum-parsimony analyses were executed in PAUP* version 4.0b10 [42] with command files for the parsimony ratchet [43] generated using the program PRAP2 [44]. The following options were implemented: tree bisection-reconnection (TBR) branch swapping, characters treated as equally weighted and unordered, gaps treated as missing characters, and bootstrap analysis was performed with 10000 replicates. Maximum-likelihood method was implemented in RAxML-HPC v8.1.11 on the CIPRES cluster (http://www.phylo.org/) [45] using the general time-reversible (GTR) model of nucleotide substitution with gamma distributed rate heterogeneity. Bayesian inference (BI) implemented in Mrbayes v3.2.3 [46] using the best-fit model (GTR+G) according to the akaike information criterion (Posada & Buckley. 2004). Two independent parallel runs of four Metropolis-coupled Monte Carlo Markov Chains (MCMCs) were run with trees sampling every 1000 generations for five million total generations.

Divergence time estimation

We first determined whether the aligned sequences were saturated for substitutions by performing the saturation test implemented in DAMBE [47]. The results of this test indicated no significant saturation signals. To calibrate our divergence date estimates of *M*. *rubra*, we set two normal priors. Firstly, the *Comptonia* node (node 1) was set to a minimum of 49 Ma (in Myr ago) based on *Comptonia columbiana*, which is a fossil species in the 'Republic' flora NE Washington and appears to be the oldest known Myricaceae fossil record [48]. Its leaves are easily recognized in the fossil record of Republic flora [49] that was dated back to 49 Ma using radiometric techniques [48]. Secondly, we estimated the split time between *M. esculenta* and *M. nana* and *M. rubra* based on data from the study of Myricaceae by Herbert [8]. The resulting time estimation (12.72±0.17 Ma) was set to the stem node of the *M. esculenta* (node 2). According to the two calibration points, divergence time of *M. rubra* was estimated under a Bayesian approach in BEAST v1.8.0 [50]. We implemented a Yule speciation tree prior and a GTR + G substitution model was selected as described above. MCMC runs were performed for 2 x 10^7 generations, with sampling every 5000 generations, following a burn-in of the initial 10% cycles. Tracer v1.5 was used to examine the sampling adequacy and convergence of the chains to a stationary distribution. TreeAnnotator v2.0.2 was used to summarize the post burn-in trees and produce a maximum clade credibility chronogram showing mean divergence time estimates with 95% HPD intervals.

Sequence annotation

In the following analysis, we removed three ingroup species (*M. cerifera*, *M. pensylvanica* and *M. adenophora*), of which only one sample was collected and thus generated insufficient read information. A BLASTX search was implemented against the NCBI non-redundant (Nr) protein database using BLAST, version 2.2.26 [51] with an E-value cut-off of 1e^{-5} for the contigs de novo assembled from each species. According to the results of the Nr protein database annotation, Blast2GO [52] was applied to obtain the functional classification of the contigs by following gene ontology (GO) terms (http://www.geneontology.org) [53], which maps contigs to function according to three principal GO categories: molecular function, cellular component and biological processes [54]. The results of the GO classification plot were obtained by WEGO (http://wego.genomics.org.cn/cgi-bin/wego/index.pl) [55]. The GO annotations of the contigs were mapped to the plant-specific GO slim ontology (http://www.geneontology.org/GO.slims.shtml) and the KEGG (http://www.genome.jp/kegg/) database.

Results

RAD tag generation and de novo assembly

After barcode trimming, cleaning and quality filtering, we obtained a total of 24.47 million paired-end reads (R1 = 78bp and R2 = 90bp). The sequencing quality was high; the Q33 of each sample was above 97%. The mean GC content of each sequence for the three species was c. 38.7% lower than the value of cDNA in the *M. rubra* cultivar 'Biqi' (49.65%) [22]. Detail information of RAD-tags sequencing is given in Table 2. De novo assembly was implemented in Trinity using R2 reads (without cut site), we obtained from 21902 to 29378 contigs with mean size of 229 to 234bp for *M. rubra*, *M. nana* and *M. esculenta*.

Phylogeny inference and divergence time estimation

The single end of the paired end reads (R1) was applied for Phylogeny inference and divergence time estimation. We recovered c. 0.95×10^6 reads from each sample of Illumina sequencing. After filtering and clustering at 92% similarity with coverage greater than 6, we obtained c. 64,129 clusters per sample with a mean depth of 10.44. Around 33631 consensus loci passed filtering for paralogs (Table 3). The sequencing error ($E = 1.63\times10^{-3}$) was lower than heterozygosity ($H = 8.18\times10^{-3}$) by ML estimation. After clustering of consensus sequences across all 19 samples under the minimum taxa data set, we recovered 4253 loci including 8360 parsimony informative variable sites with 3677 unlinked SNP sites, which were applied for phylogeny inference of four species of *Morella* in China. For the Bayesian analysis,

Table 2. Detail information of RAD-tag sequencing.

	M. rubra	*M. nana*	*M. esculenta*
Number of reads (million)	8.73	8.83	6.91
Total length of reads (million bp)	733	741	580
GC Rate %	38.29	39.08	38.72
Number of contigs	29378	21902	23552
Average contig length (bp)	229	234	234
N50	12705	9440	10150
Contig length range (bp)	150–472	150–414	150–442

doi:10.1371/journal.pone.0139840.t002

phylogeny reconstruction using the minimum taxa data set revealed strong support for the four species as a clade (1.00 PP), as well as the monophyly of each species. The topology of (((*M. rubra*, *M. adenophora*), *M. nana*), *M. esculenta*) was significantly favored. Two species from North America (*M. cerifera* and *M. pensylvanica*) were sister to the four Chinese species. The tree topologies from the MP and ML analyses (S1 Fig) were consistent with the results of the Bayesian analysis (Fig 2).

The BEAST-derived RAD-taq chronogram of *Morella* (Fig 2) recovered the four individuals of *M. rubra* as monophyly (posterior probability, PP = 1.00), with an estimated stem and crown age were c. 5.28 Ma (95% HPD: 3.84–7.08 Ma, Node A) and c. 3.39 Ma (95% HPD: 2.20–4.80 Ma, Node B) respectively. For this chronogram, BEAST provided an average substitution rate of 9.40×10^{-9} s/s/y, which is congruent with the mean values reported for plant nuclear DNAs ($5.0–30.0\times10^{-9}$s/s/y) [56]. Divergence time estimation indicated that the origin

Table 3. Results of filtering and clustering of one single end RAD sequences (R1) from 19 samples in this study.

Taxon	ID	RAD tags ($\times10^6$)	Clusters at 92%[1]	Mean depth	Consensus loci[2]	Number of loci in the minimum-taxa
Morella rubra	ZJGT	1.38	99748	11.19	67569	4100
	GZLS	1.11	91632	9.81	54655	3742
	YNDW	0.90	70869	9.63	33965	3197
	YNML	0.99	78684	9.87	43887	3621
M. adenophora	GXHX	0.71	41305	10.04	17345	2855
M. nana	YNFM	0.88	49315	10.65	22532	3255
	YNXW	1.09	66835	11.26	38145	3852
	YNWS	0.86	57805	9.67	26756	3375
	YNPL	0.85	52608	10.53	25070	3490
	YNHK	0.72	41091	10.86	16550	2781
	YNZJ	1.09	89046	9.12	54454	3831
M. esculenta	YNAL	0.77	52842	9.82	22825	2961
	YNWS	0.90	64529	10.24	32974	3504
	SCYB	0.88	62287	10.44	32872	3591
	SCMY	0.91	65235	10.19	35229	3675
M. pensylvanica	USBH	1.00	69555	10.82	38171	2722
M.cerifera	USR	1.01	60548	10.45	29835	2330
Comptonia Peregrine-1	USR1	0.96	43671	11.61	17602	1065
Comptonia Peregrine-2	USR2	1.11	60849	12.18	28561	1230

Note: [1]Clusters with more than the minimum depth of six reads.

[2]Consensus loci which passed filtering for paralogs.

doi:10.1371/journal.pone.0139840.t003

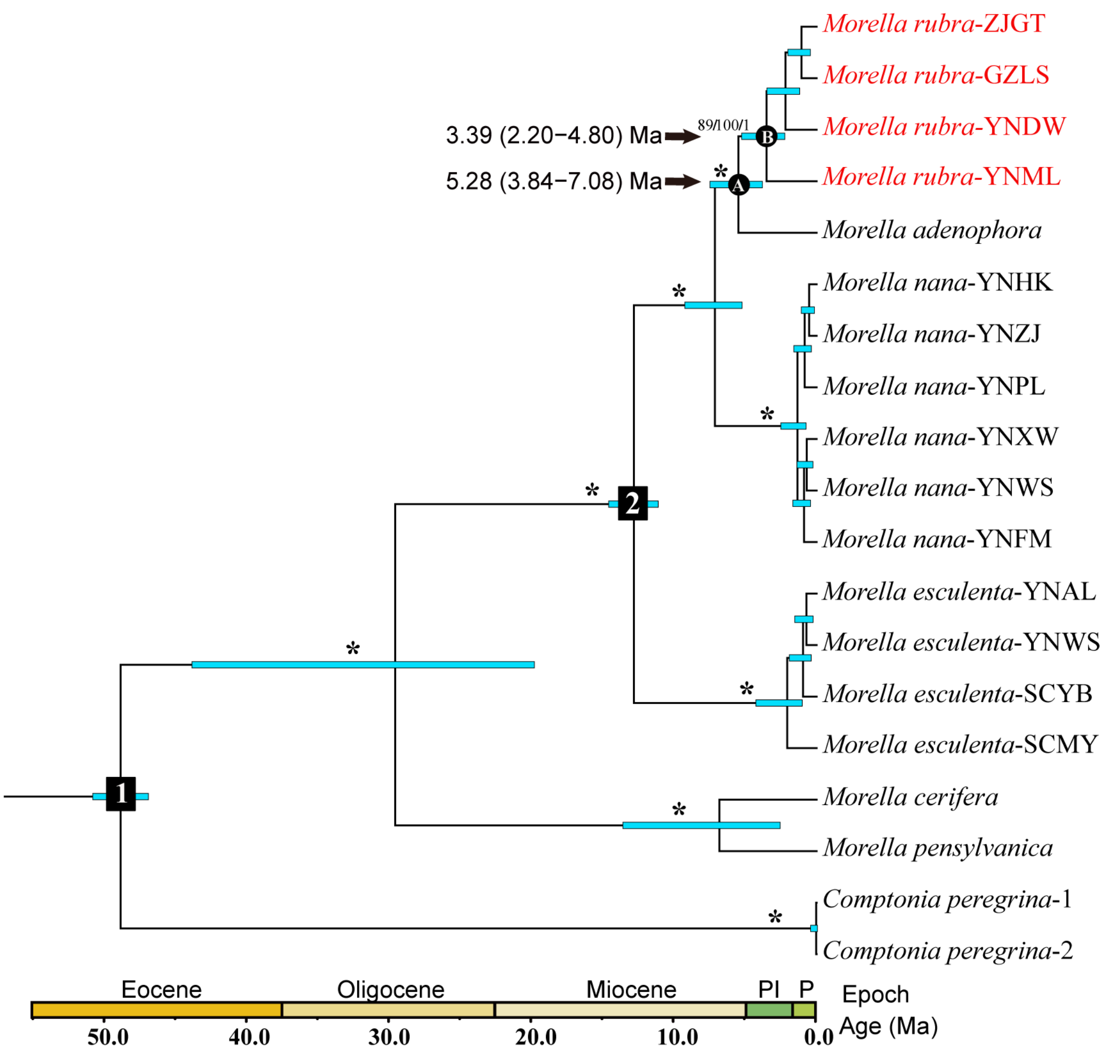

Fig 2. Bayesian phylogeny and divergence time estimation of *Morella*. Node1 and node2 represent two calibration points described in methods above. Blue bars indicate the 95% highest posterior density (HPD) credibility intervals for node ages (Ma). Asterisk indicates that maximum-parsimony bootstrap/maximum-likelihood bootstrap/Bayesian inference posterior probability equal to 100/100/1.

doi:10.1371/journal.pone.0139840.g002

of *M*. *rubra* was near the Miocene-Pliocene boundary; intraspecific divergence in *M*. *rubra* occurred in the late Pliocene.

After clustering of the consensus loci for four individuals (GTS, GZLS, DWS and MLP), we identified 3808 unlinked SNP sites within the populations of *M*. *rubra*. These SNP markers may be applicable in future studies of the population genetics of *M*. *rubra*.

Sequence annotation and GO enrichment analysis

Based on the public Nr databases, 22.9% (6730) of assembled contigs in *M*. *rubra*, 24.1% (5287) in *M*. *nana* and 28.7% (6769) in *M*. *esculenta* were definitely mapped to known genes (Fig 3). The summary of the annotated contigs function is described in S1 Table. The top-hit species distribution of *M*. *rubra* for BLAST results were as follows: *Vitis vinifera*, *Amygdalus persica*, *Populus balsamifera*, *Fragaria vesca*, *Ricinus communis*, *Glycine max* and *Cucumis sativus* (Fig 4). *M*. *nana* and *M*. *esculenta* received nearly the same results (S2 Fig).

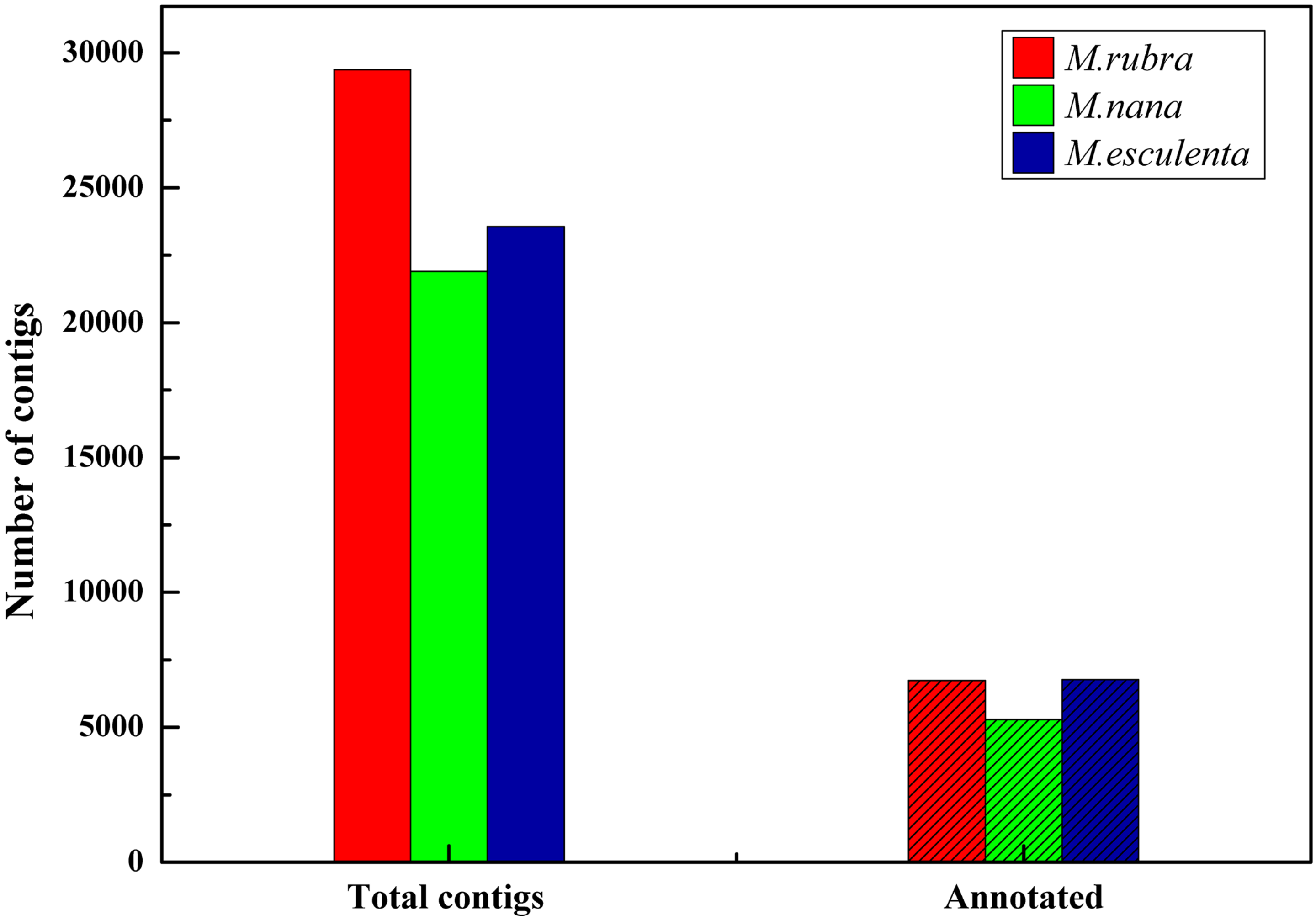

Fig 3. Number of the annotated contigs.

doi:10.1371/journal.pone.0139840.g003

According to the results of Nr protein database annotation, 6730, 5287, and 6769 contigs for *M*. *rubra*, *M*. *nana* and *M*. *esculenta* respectively were implemented in Blast2GO for functional classification. For *M*. *rubra*, genes involved in cellular processes (GO: 0009987) and metabolic processes (GO: 0008152) are the top two most abundant subcategories in the biological process. Cell (GO: 0005623) and Cell part (GO: 0044464) are highly represented under the cellular component category. Binding (GO: 0005488) represents the major proportion of molecular function (Fig 5). *M*. *nana* and *M*. *esculenta* showed similar results (S3 Fig).

Chinese bayberry is generally thought to have been domesticated from the wild *M*. *rubra*. The lack of detailed knowledge of the genetics of wild populations is seriously hindering the industry's ability to improve commercial stocks. To remedy this situation, we obtained annotated contigs of *M*. *rubra* to map to the KEGG database and identified those that may be correlated with fruit ripening, senescence, fruit quality and disease/defense (Table 4, S2 Table).

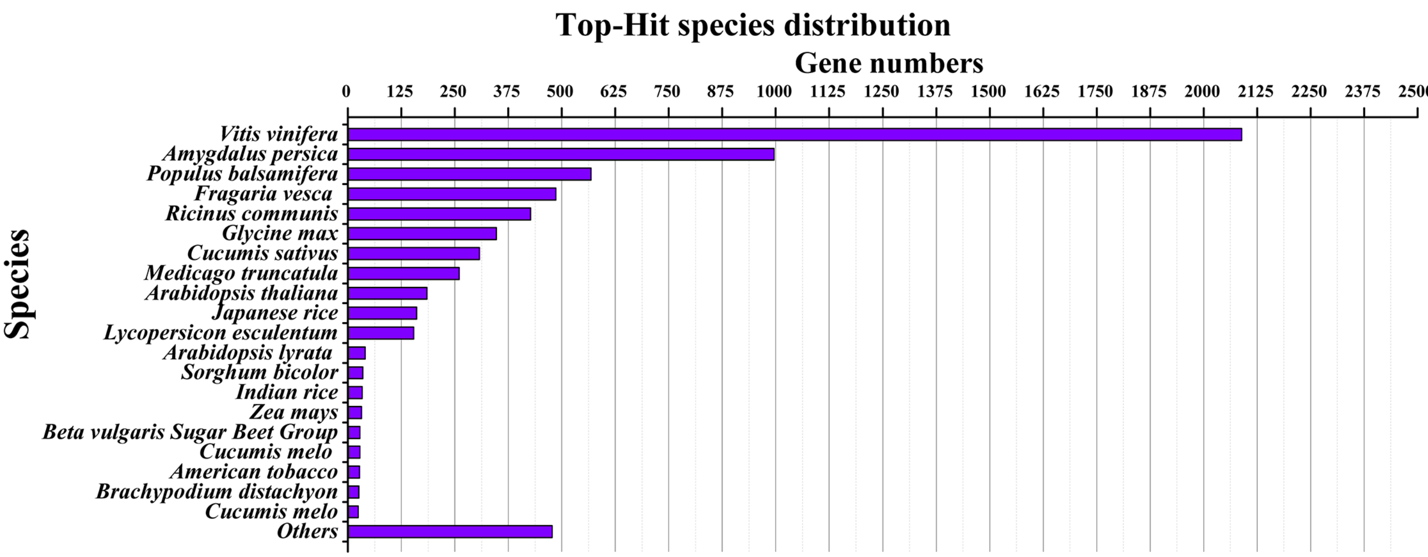

Fig 4. Top hit species distribution of *M. rubra* for BLAST result.

doi:10.1371/journal.pone.0139840.g004

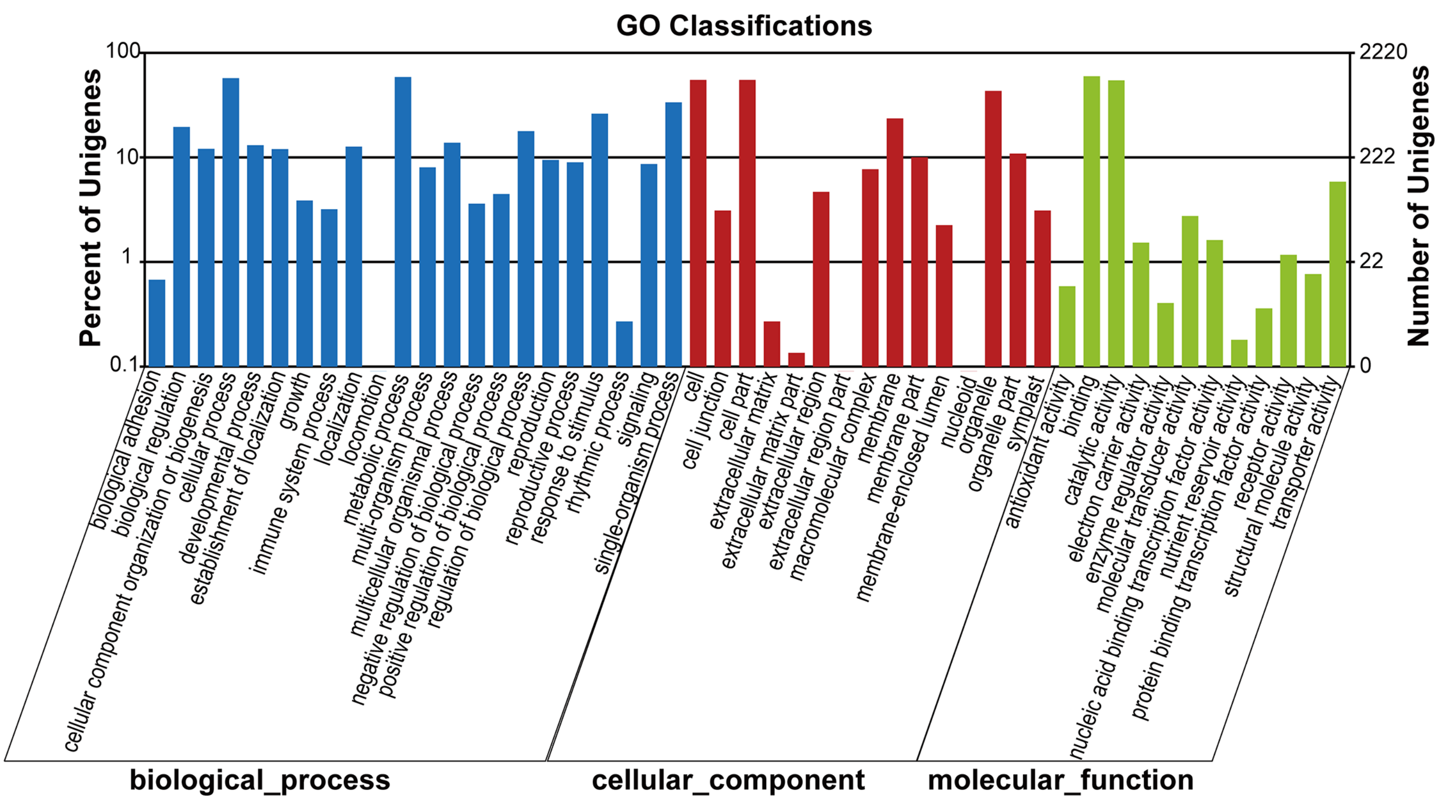

Fig 5. GO classifications of annotated contigs of *M. rubra*.

doi:10.1371/journal.pone.0139840.g005

Table 4. Annotated contigs associated with fruit ripening and senescence, fruit quality formation and disease/defense metabolism in *Morella rubra*.

	Kegg ID	Deduced protein	Pathway involved
Ripening and senescence			
	K14516	ethylene-responsive transcription factor 1	Plant hormone signal transduction
	K14517	ethylene-responsive transcription factor 2	Plant hormone signal transduction
	K14514	ethylene-insensitive protein 3	Plant hormone signal transduction
	K05933	aminocyclopropanecarboxylate oxidase	Cysteine and methionine metabolism
	K00434	L-ascorbate peroxidase	Ascorbate and aldarate metabolism
	K01365	cathepsin L	Lysosome and Phagosome
Quality formation			
Color			
	K05280	flavonoid 3'-monooxygenase	Flavonoid biosynthesis
	K09422	myb proto-oncogene protein, plant	
	K12930	anthocyanidin 3-O-glucosyltransferase	Anthocyanin biosynthesis
	K12338	anthocyanidin 3-O-glucoside 5-O-glucosyltransferase	Anthocyanin biosynthesis
Texture			
	K01213	galacturan 1,4-alpha-galacturonidase	Pentose and glucuronate interconversions
	K01184	polygalacturonase	Pentose and glucuronate interconversions
	K00430	peroxidase	Phenylpropanoid biosynthesis
Aroma			
	K00001	alcohol dehydrogenase	Degradation of aromatic compounds
	K15086	(3S)-linalool synthase	Monoterpenoid biosynthesis
	K14174	beta-farnesene synthase	Sesquiterpenoid and triterpenoid biosynthesi
	K15797	(S)-beta-macrocarpene synthase	Sesquiterpenoid and triterpenoid biosynthesi
	K15799	(+)-alpha-barbatene/beta-chamigrene/thujopsene synthase	Sesquiterpenoid and triterpenoid biosynthesis
Disease/defense			
	K13456	RPM1-interacting protein 4	Plant-pathogen interaction
	K13457	disease resistance protein RPM1	Plant-pathogen interaction

doi:10.1371/journal.pone.0139840.t004

Discussion

Phylogenetic relationships and taxonomy of *Morella rubra* and its close relatives

Previous phylogenetic studies have shown that *Myrica gale* and *M*. *hartwegii* are distinct from other species in the Linnaean genus *Myrica* [8, 57], therefore requiring recognition of two genera, *Myrica sensu stricto* (2 spp.) and *Morella* Lour.. Such a treatment is also supported by morphology [58] and cytology [59]. The four Chinese species were found to form a monophyletic clade in the *Morella*, but the phylogenetic relationships between them were not resolved [8], probably due to insufficient informative loci. In this study we used RAD-seq data to resolve the phylogenetic relationships of four species of *Morella* in China. In comparison to ITS and cpDNA markers, RAD-seq is excellent because more than 4253 loci including 8360 parsimony informative sites (the minimum-taxa data set) were generated. The phylogenetic relationships of the four Chinese species of *Morella* are well resolved with highest support. *M*. *esculenta* is basal in the Chinese *Morella* clade, which is congruent with the study by Hebert [8]. *M*. *nana* formed a strongly supported clade that is sister to the (*M*. *adenophora* + *M*. *rubra)* clade. Our result showed RAD-seq to be an effective approach for resolving phylogenetic relationships among closely related species.

In regards to the taxonomy, the four *Morella* species in China are easily distinguished with each other by the habit, indumentum, inflorescence type, fruit shape, flowering season, leaf, and flower morphology [9]. This agrees with our RAD-seq based phylogeny. So far, there are still many remaining controversies in *Myrica* species in the Indo-China region. In addition, many of these species are actually *Morella*, but most of the names have not been transferred to the right genus yet. For example, one (*Myrica esculenta* Buch.-Ham. ex D. Don) to five species (*M*. *esculenta*, *M*. *farquhariana* Wall., *M*. *sapida* Wall., *M*. *nagi* Thunb., *M*. *integrifolia* Roxb.) are recognized in India by different authors. Yanthan et al. [60] tried to resolve this dispute using the 18S-26S rDNA ITS sequences and proposed that *M*. *nagi* and *M*. *esculenta* should be treated as two separate species. This effort helps us to understand the complexity of *Myrica* species in Indo-China region. However, we believe that a phylogenetic study based on next-generation sequencing (such as RAD-seq in this study) and more comprehensive sampling will easily solve the mystery that can finally provide us a natural classification system for Myricaceae.

Intraspecific divergence in *M. rubra* driven by the third uplift of the QTP

The origin and evolution of biodiversity is always linked with a range of geological or climatic processes, such as continental drift, the uplift of mountain chains and climatic fluctuation associated with ice ages. These processes can create new habitats and provide opportunities for speciation by interacting with each other [61]. Historical orogenesis and climatic oscillations can also cause fragmentation of species distributions and isolation of populations, leading to reduced gene flow and allopatric divergence [62]. To date, a series of studies have shown that ecological factors, such as temperature and precipitation, can drive speciation or infraspecific divergence [63–65]. Based on RAD-seq data, we determined that speciation of *M*. *rubra* was at c. 5.28 Ma (95% HPD: 3.84–7.08 Ma), and infraspecific divergence in *M*. *rubra* occurred in the late Pliocene (3.39Ma, 95% HPD: 2.20–4.80 Ma). The latter timescale is consistent with the third intense uplift of the Qinghai-Tibet Plateau (QTP) and the formation of the Hengduan Mountains (c. 3.6 Ma) [66–67]. Our ongoing study on the domestication of *M*. *rubra* also reveals that wild populations from Yunnan are the basal ones, indicating that it might have originated in the Hengduan Mountains area (unpublished data). Therefore, it is likely that intraspecific divergence in *M*. *rubra* occurred in the late Pliocene, and was driven by the uplift of the QTP and the formation of the Hengduan Mountains.

Valuable genetic resources of Chinese bayberry based on RAD-seq

Chinese bayberry is a specialty fruit of China, and grown commercially in eastern and western China. It is one of the most popular fruit crops because of its food, medicinal and landscape value and has become an important export product in China [68]. However, it is highly perishable and susceptible to mechanical injury, physiological deterioration and fungal decay, resulting in a postharvest life of only 1 to 2 days under ambient temperature [31].

Feng et al. [22] analyzed the RNA-seq of Chinese bayberry to determine the molecular mechanisms for change in fruit color and taste during ripening. Zhu et al. [26] analyzed 2000 EST sequences from the cDNA libraries of Chinese bayberry cultivar 'Biqi', and identified several genes associated with disease/defense and anthocyanin accumulation, gene encoding elements correlated with ethylene biosynthesis and signal transductions, and proteins linked to senescence regulation and quality during fruit ripening. In this study, based on de novo assembly using R2 reads, we identified annotated contigs which are thought to be correlated with fruit ripening and senescence, fruit quality, disease/defense metabolism, and other important pathways.

Chinese bayberry has been cultivated for more than 2000 years, but detailed studies of its biology started only three decades ago [25]. There are approximately 305 recorded accessions, of which 268 are named cultivars [69]. Zhang et al. [25] used amplified fragment length polymorphism (AFLP) to reveal genetic diversity of 100 accessions of Chinese bayberry, and showed that the subgroups were somewhat related to the region of origin of the accessions, but accessions from the same region did not necessarily belong to the same group or subgroup due to extensive gene flow among different regions. Jiao et al. [24] developed simple sequence repeat (SSR) markers for Chinese bayberry and came to a similar conclusion. However, none of the previous studies was able to reveal the relationships among cultivars with high resolution, probably because of an insufficient number of informative sites. Besides, studies on genetic diversity and population structure of wild *M*. *rubra* populations are limited. The origin of domesticated Chinese bayberry has never been resolved. In our study, however, 3808 SNPs were identified within wild *M*. *rubra* populations, which will be a valuable resource for subsequent studies on the population genetics of *M*. *rubra* and the domestication of Chinese bayberry.

Supporting Information

S1 Fig. Phylogeny inference using Maximum parsimony (MP) method and Maximum likelihood (ML) methods.
(TIFF)

S2 Fig. The top hit species distribution of *M. nana* and *M. esculenta* for BLAST result.
(TIFF)

S3 Fig. GO classifications of annotated contigs of *M. nana* and *M. esculenta*.
(TIFF)

S1 Table. Top BLAST hits from public databases. Lists of the top results from BLASTING *M*. *rubra*, *M*. *nana* and *M*. *esculenta* contigs against public databases (E-value cut-off of 10^{-3}).
(XLSX)

S2 Table. The result of the annotated contigs against KEGG database for *M. rubra*.
(XLS)

Acknowledgments

We would like to thank Lihui Gong, Yonghua Zhang, Yihan Wang and Yunrui Mao for assisting in collecting samples, as well as Guoyun Wang, Zehuang Zhang and Chunming Xu for their various support. We are grateful to David E. Boufford for revising and improving the manuscript. This work was supported by the Special Fund for Agro-scientific Research in the Public Interest (grant no. 201203089).

Author Contributions

Conceived and designed the experiments: PL XL CF. Performed the experiments: LL. Analyzed the data: LL. Contributed reagents/materials/analysis tools: LL NC XJ. Wrote the paper: LL PL.

References

1. Janick J The origins of fruits, fruit growing, and fruit breeding. Plant Breeding Reviews 25: 5–320.
2. Ferguson AR, Huang H (2007) Genetic resources of kiwifruit: domestication and breeding. Horticultural Reviews 33: 1–121.

3. Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, et al. (2013) The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nature Genetics 45: 487–494. doi: 10.1038/ng.2586 PMID: 23525075
4. Xu Q, Chen LL, Ruan X, Chen D, Zhu A, Chen CL, et al. (2013) The draft genome of sweet orange (*Citrus sinensis*). Nature Genetics 45: 59–66. doi: 10.1038/ng.2472 PMID: 23179022
5. Fang Z, Zhang Y, Lü Y, Ma G, Chen J, Liu D, et al. (2009) Phenolic compounds and antioxidant capacities of bayberry juices. Food Chemistry 113: 884–888.
6. Kang W, Li Y, Xu Y, Jiang W, Tao Y (2012) Characterization of Aroma Compounds in Chinese Bayberry (*Myrica rubra* Sieb. et Zucc.) by Gas Chromatography Mass Spectrometry (GC-MS) and Olfactometry (GC-O). Journal of Food Science 77: C1030–C1035. doi: 10.1111/j.1750-3841.2012.02747.x PMID: 23009608
7. Cronquist A (1981) An integrated system of classification of flowering plants: Columbia University Press. pp. 214–217
8. Herbert J (2005) Systematics and biogeography of Myricaceae. Unpublished PhD dissertation, University of St Andrews St Andrews.
9. Lu AM, Bornstein AJ (1999) Myricaceae. pp. 275–276 in Wu Z, Raven P Flora of China, Volume 4, Cycadaceae through Fagaceae. Science Press and Missouri Botanical Garden, Beijing and St. Louis.
10. Cariou M, Duret L, Charlat S (2013) Is RAD-seq suitable for phylogenetic inference? An in silico assessment and optimization. Ecology and Evolution 3: 846–852. doi: 10.1002/ece3.512 PMID: 23610629
11. Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007) Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. Genome Research 17: 240–248. PMID: 17189378
12. Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, et al. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. PloS One 3: e3376. doi: 10.1371/journal.pone.0003376 PMID: 18852878
13. Emerson KJ, Merz CR, Catchen JM, Hohenlohe PA, Cresko WA, Bradshaw WE, et al. (2010) Resolving postglacial phylogeography using high-throughput sequencing. Proceedings of the National Academy of Sciences 107: 16196–16200.
14. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PloS One 6: e19379. doi: 10.1371/journal.pone.0019379 PMID: 21573248
15. Eaton DA, Ree RH (2013) Inferring phylogeny and introgression using RADseq data: an example from flowering plants (Pedicularis: Orobanchaceae). Systematic Biology 62: 689–706. doi: 10.1093/sysbio/syt032 PMID: 23652346
16. White TA, Perkins SE, Heckel G, Searle JB (2013) Adaptive evolution during an ongoing range expansion: the invasive bank vole (*Myodes glareolus*) in Ireland. Molecular Ecology 22: 2971–2985. doi: 10.1111/mec.12343 PMID: 23701376
17. Wang X, Zhao L, Eaton D, Li D, Guo Z (2013) Identification of SNP markers for inferring phylogeny in temperate bamboos (Poaceae: Bambusoideae) using RAD sequencing. Molecular Ecology Resources 13: 938–945. doi: 10.1111/1755-0998.12136 PMID: 23848836
18. Xu P, Xu S, Wu X, Tao Y, Wang B, Wang S, et al. (2014) Population genomic analyses from low-coverage RAD-Seq data: a case study on the non-model cucurbit bottle gourd. The Plant Journal 77: 430–442. doi: 10.1111/tpj.12370 PMID: 24320550
19. Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA, et al. (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. PLoS Genetics 6: e1000862. doi: 10.1371/journal.pgen.1000862 PMID: 20195501
20. Hess JE, Campbell NR, Close DA, Docker MF, Narum SR (2013) Population genomics of Pacific lamprey: adaptive variation in a highly dispersive species. Molecular Ecology 22: 2898–2916. doi: 10.1111/mec.12150 PMID: 23205767
21. Cheng H, Chen J, Chen S, Wu D, Liu D, Ye X, et al. (2015) Characterization of aroma-active volatiles in three Chinese bayberry (*Myrica rubra*) cultivars using GC–MS–olfactometry and an electronic nose combined with principal component analysis. Food Research International 72: 8–15.
22. Feng C, Chen M, Xu CJ, Bai L, Yin XR, Li X, et al. (2012) Transcriptomic analysis of Chinese bayberry (*Myrica rubra*) fruit development and ripening using RNA-Seq. BMC Genomics 13: 19. doi: 10.1186/1471-2164-13-19 PMID: 22244270
23. Xie X, Qiu Y, Ke L, Zheng X, Wu G, Chen J, et al. (2011) Microsatellite primers in red bayberry, *Myrica rubra* (Myricaceae). American Journal of Botany 98: e93–e95. doi: 10.3732/ajb.1000271 PMID: 21613157

24. Jiao Y, Jia HM, Li XW, Chai ML, Jia HJ, Chen Z, et al. (2012) Development of simple sequence repeat (SSR) markers from a genome survey of Chinese bayberry (*Myrica rubra*). BMC Genomics 13: 201. doi: 10.1186/1471-2164-13-201 PMID: 22621340
25. Zhang S, Gao Z, Xu C, Chen K, Wang G, Zheng J, et al. (2009) Genetic diversity of Chinese bayberry (*Myrica rubra* Sieb. et Zucc.) accessions revealed by amplified fragment length polymorphism. Hortscience 44: 487–491.
26. Zhu C, Feng C, Li X, Xu C, Sun C, Chen K (2013) Analysis of expressed sequence tags from Chinese bayberry fruit (*Myrica rubra* sieb. And zucc.) at different ripening stages and their association with fruit quality development. International Journal of Molecular Sciences 14: 3110–3123. doi: 10.3390/ijms14023110 PMID: 23377019
27. Chen K, Xu C, Zhang B, Ferguson IB (2004) Red bayberry: botany and horticulture. Horticultural Reviews 30: 83–114.
28. Zhang W, Chen K, Zhang B, Sun C, Cai C, Zhou C, et al. (2005) Postharvest responses of Chinese bayberry fruit. Postharvest Biology and Technology 37: 241–251.
29. Zhang W, Li X, Zheng J, Wang G, Sun C, Ferguson IB, et al. (2008) Bioactive components and antioxidant capacity of Chinese bayberry (*Myrica rubra* Sieb. and Zucc.) fruit in relation to fruit maturity and postharvest storage. European Food Research and Technology 227: 1091–1097.
30. Guo L, Deng X (2004) Identification of the pathogen of bayberry (*Myrica rubra*) leaf blight and its fungicides sensitivity. Chinese Agricultural Science Bulletin 21: 359–362.
31. Wang K, Cao S, Jin P, Rui H, Zheng Y (2010) Effect of hot air treatment on postharvest mould decay in Chinese bayberry fruit and the possible mechanisms. International Journal of Food Microbiology 141: 11–16. doi: 10.1016/j.ijfoodmicro.2010.05.004 PMID: 20510474
32. Zhang W, Li X, Wang X, Wang G, Zheng J, Abeysinghe DC, et al. (2007) Ethanol vapour treatment alleviates postharvest decay and maintains fruit quality in Chinese bayberry. Postharvest Biology and Technology 46: 195–198.
33. Wang K, Jin P, Tang S, Shang H, Rui H, Di H, et al. (2011) Improved control of postharvest decay in Chinese bayberries by a combination treatment of ethanol vapor with hot air. Food Control 22: 82–87.
34. Chen XY, Zheng YR (1988) A preliminary report on path of invading of pathogens and morphogenesis of bacterial gall of red bayberry. Journal of Zhejiang University (Agriculture & Life Sciences) 14:239–241.
35. Doyle J (1991) DNA protocols for plants. Molecular techniques in taxonomy: Springer. pp. 283–293.
36. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology 29: 644–652. doi: 10.1038/nbt.1883 PMID: 21572440
37. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26: 2460–2461. doi: 10.1093/bioinformatics/btq461 PMID: 20709691
38. Lynch M (2008) Estimation of nucleotide diversity, disequilibrium coefficients, and mutation rates from high-coverage genome-sequencing projects. Molecular Biology and Evolution 25: 2409–2419. doi: 10.1093/molbev/msn185 PMID: 18725384
39. Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Research 18: 1851–1858. doi: 10.1101/gr.078212.108 PMID: 18714091
40. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32: 1792–1797. PMID: 15034147
41. Hohenlohe PA, Amish SJ, Catchen JM, Allendorf FW, Luikart G (2011) Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout. Molecular Ecology Resources 11: 117–122. doi: 10.1111/j.1755-0998.2010.02967.x PMID: 21429168
42. Swofford DL (2002) Phylogenetic analysis using parsimony (* and other methods). Version 4. Sinauer.
43. Nixon KC (1999) The parsimony ratchet, a new method for rapid parsimony analysis. Cladistics 15: 407–414.
44. Wall PK, Leebens-Mack J, Müller KF, Field D, Altman NS (2008) PlantTribes: a gene and gene family resource for comparative genomics in plants. Nucleic Acids Research 36: D970–D976. PMID: 18073194
45. Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. IEEE. pp. 1–8.
46. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574. PMID: 12912839

47. Xia X, Xie Z (2001) DAMBE: software package for data analysis in molecular biology and evolution. Journal of Heredity 92: 371–373. PMID: 11535656
48. Wolfe JA, Wehr W (1987) Middle Eocene dicotyledonous plants from Republic, northeastern Washington.
49. Manchester SR (1999) Biogeographical relationships of North American tertiary floras. Annals of the Missouri Botanical Garden: 472–522.
50. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7: 214. PMID: 17996036
51. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. Journal of Molecular Biology 215: 403–410. PMID: 2231712
52. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3674–3676. PMID: 16081474
53. Ashburner J, Friston KJ (2000) Voxel-based morphometry—the methods. Neuroimage 11: 805–821. PMID: 10860804
54. Harris TW, Chen N, Cunningham F, Tello-Ruiz M, Antoshechkin I, Bastiani C, et al. (2004) WormBase: a multi-species resource for nematode biology and genomics. Nucleic Acids Research 32: D411–D417. PMID: 14681445
55. Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, et al. (2006) WEGO: a web tool for plotting GO annotations. Nucleic Acids Research 34: W293–W297. PMID: 16845012
56. Wolfe KH, Li W-H, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proceedings of the National Academy of Sciences 84: 9054–9058.
57. Huguet V, Gouy M, Normand P, Zimpfer JF, Fernandez MP (2005) Molecular phylogeny of Myricaceae: a reexamination of host–symbiont specificity. Molecular Phylogenetics and Evolution 34: 557–568. PMID: 15683929
58. Herbert J (2005) New combinations and a new species in *Morella* (Myricaceae). Novon: 293–295.
59. Baird JR (1970) A taxonomic revision of the plant family Myricaceae of North America, North of Mexico. Unpublished PhD dissertation, University of North Carolina.
60. Yanthan M, Biate D, Misra AK (2011) Taxonomic resolution of actinorhizal *Myrica* species from Meghalaya (India) through nuclear rDNA sequence analyses. Functional Plant Biology 38: 738–746.
61. Liu J, Möller M, Provan J, Gao LM, Poudel RC, Li DZ (2013) Geological and ecological factors drive cryptic speciation of yews in a biodiversity hotspot. New Phytologist 199: 1093–1108. doi: 10.1111/nph.12336 PMID: 23718262
62. Coyne JA (1992) Genetics and speciation. Nature 355: 511–515. PMID: 1741030
63. Givnish TJ (2010) Ecology of plant speciation. Taxon: 1326–1366.
64. Keller I, Seehausen O (2012) Thermal adaptation and ecological speciation. Molecular Ecology 21: 782–799. doi: 10.1111/j.1365-294X.2011.05397.x PMID: 22182048
65. Li L, Abbott RJ, Liu B, Sun Y, Li L, Zou J, et al. (2013) Pliocene intraspecific divergence and Plio-Pleistocene range expansions within Picea likiangensis (Lijiang spruce), a dominant forest tree of the Qinghai-Tibet Plateau. Molecular Ecology 22: 5237–5255. doi: 10.1111/mec.12466 PMID: 24118118
66. Li J, Fang X (1999) Uplift of the Tibetan Plateau and environmental changes. Chinese Science Bulletin 44: 2117–2124.
67. Shi Y (1998) Uplift and environmental changes of Qinghai-Xizang (Tibetan) Plateau in the late Cenozoic: Guangdong Scirence & Technology Press, Guangzhou.
68. Li X, He Y (2006) Non-destructive measurement of acidity of Chinese bayberry using Vis/NIRS techniques. European Food Research and Technology 223: 731–736.
69. Zhang Y, Miao S (1999) Resources of red bayberry and its utilization in China. S China Fruits 28: 24–25.