

RESEARCH ARTICLE

Genetic diversity and biogeography of *T. officinale* inferred from multi locus sequence typing approach

Mohammadjavad Jafari^{1,2,3}, Waheed Akram^{1,2,3}, Yanju Pang^{1,2,3}, Aqeel Ahmad^{1,2,3}, Shakeel Ahmed^{1,2,3}, Nasim Ahmad Yasin⁴, Tehmina Anjum⁴, Basharat Ali⁵, Xiangdong Hu^{1,2,3}, Xiaohua Li^{1,2,3}, Shuang Dong⁶, Qian Cai⁶, Matteo Ciprian⁷, Monika Bielec⁸, Sheng Hu⁶, Fatemeh Sefidkon⁹, Xuebo Hu^{1,2,3*}

1 Laboratory of Drug Discovery and Molecular Engineering, Department of Medicinal Plants, College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, **2** National-Regional Joint Engineering Research Center in Hubei for Medicinal Plant Breeding and Cultivation, Huazhong Agricultural University, Wuhan, China, **3** Medicinal Plant Engineering Research Center of Hubei Province, Huazhong Agricultural University, Wuhan, China, **4** Institute of Agricultural Sciences, University of the Punjab, Lahore, Pakistan, **5** Department of Microbiology and Molecular Genetics, University of the Punjab, Lahore, Pakistan, **6** Hubei Cancer Hospital, Wuhan, China, **7** Laboratory of Organometallics, Catalysis and Ordered Materials, State Key Laboratory of Advanced Technology for Materials Synthesis and Processing, Wuhan University of Technology, Wuhan, China, **8** Department of Biomedical Engineering, Wuhan University of Technology, Wuhan, China, **9** Research Institute of Forest and Rangeland, Iran

* xuebohu@mail.hzau.edu.cn



OPEN ACCESS

Citation: Jafari M, Akram W, Pang Y, Ahmad A, Ahmed S, Yasin NA, et al. (2018) Genetic diversity and biogeography of *T. officinale* inferred from multi locus sequence typing approach. PLoS ONE 13(9): e0203275. <https://doi.org/10.1371/journal.pone.0203275>

Editor: Tzen-Yuh Chiang, National Cheng Kung University, TAIWAN

Received: April 16, 2018

Accepted: August 19, 2018

Published: September 18, 2018

Copyright: © 2018 Jafari et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data are all contained within the paper and/or Supporting Information files.

Funding: This research was partially supported by National Key R&D Program of China (No. 2017YFD0501500) and the Fundamental Research Funds for the Central Universities Program (No. 2662017PY104 and No. 2662017PY104).

Competing interests: The authors declare that they have no competing interests. The funders had no role in study design, data collection and analysis,

Abstract

Taraxacum officinale (Asteraceae) is widely distributed weedy plant used as a traditional medicinal herb. The population genetics and historical biogeography of this plant have remained relatively unexplored. This study explores phylogeny, population genetics and ancestral reconstructions adopting multi locus sequence typing (MLST) approach. MLST sequences dataset was generated from genomics and chloroplast DNA sequences obtained from 31 *T. officinale* haplotypes located in 16 different countries. Phylogenetic analysis distributed these haplotypes in well differentiated geographic clades. The study suggested a close relationship between Europe and adjacent Asian countries. Populations of these regions predominantly formed common haplogroups, showed considerable level of gene flow and evidence for recombination events across European and Asian population. Biogeographical inferences obtained by applying statistical dispersal-vicariance analysis (S-DIVA) and Bayesian binary MCMC (BBM) analysis showed that *T. officinale* was putatively originated in Europe. Molecular clock analysis based on ITS dataset suggested that the divergence between Europe and East Asian populations can be dated to 1.07 Mya with subsequent dispersal and vicariance events. Among different spatial process long distance seed dispersal mediated by wind had potentially assisted the population expansion of *T. officinale*.

decision to publish, or preparation of the manuscript.

Introduction

Dandelion (*Taraxacum officinale*; *Compositae-Lactuceae*) is a group of highly invasive wild plants with worldwide distribution. *Taraxacum* species dominate in towns, along roadsides, shores of conduits and in many other rural sites. The representative species of this genus are *Taraxacum officinale* and *T. erythrospermum*. *T. officinale* has been documented particularly in North America, Asia, and Europe [1]. In USA, *T. officinale* is found abundantly in cold regions. This indicates that it is basically acquainted with America through Alaska [2]. It was reported that the spread of *T. officinale* was contributed by the early European pilgrims who introduced different species of dandelion as a fancy [3, 4]. Ecological variation in *T. officinale* is derived from relatively minor morphological modifications, leaving a seemingly limited set of morphological characters for phylogenetic studies. Multiple hybridization events in the evolution process of *Taraxacum* impair great difficulties in the interpretation of specific variation in *Taraxacum* phylogeny and biogeography [5]. A remarkable phenomenon is common existence of geographical parthenogenesis [6]. Another situation worth mentioning is the occurrence of several geographically restricted complexes of sympatric diploid sexuality and closely related agamosperms comprising higher levels of variations and dynamics [7]. These processes are discussed in more detail by Kirschner et al., 2003 [8].

The molecular typing scheme should generate data that can be employed to resolve the structure of the population in addition to provide highly discriminating data and resolving biological relationships among collection of isolates. Molecular data from a single barcode locus is not sufficient for a broader context exploration of phylogeny. A larger number of loci would be preferable to explore population genetics [9]. Multi locus sequence typing (MLST) implies a typing scheme that involves the determination of the alleles at multiple loci by nucleotide sequencing. This method have been employed in studies of the population biology, pathogenicity, and evolution [9]. Phylogenetic analysis based on combined dataset (MLST) from nuclear (nrDNA) and chloroplast DNA (cpDNA) loci has been successfully used for some organisms [10–12].

Utilization of sequence variations of ITS region of genomic DNA (nrDNA) alone is not suitable for phylogenetic analysis of *T. officinale* because of high sequence variations with in agamospermous polyploid individuals [13]. Congruence of homologous characters emerges from the interaction of coding and non-coding molecular barcode datasets that allows overall phylogenetic resolution [14]. Therefore, this combined dataset analysis approach (MLST) can uncover underlying homology and detailed phylogeny to understand *Taraxacum* evolution.

Spatial processes are crucial for determining the structure and dynamics of population and communities [15]. Seed dispersal is the premier spatial demographic process among sessile organisms. The seed-dispersal pattern not only determines the potential area of plant recruitment but also affects the rates of gene flow thus, influences genetic structure of population [16]. Wind mediated long-distance dispersal events play a major role in determining population spread, the flow of individuals between population and the colonization of unoccupied habitats [17, 18]. *Taraxacum* is the most popular example for a wind-dispersed species with a hairy structure aiding aerial transport. The vast majority (99%) of wind-dispersed seeds are predicted to travel up to two meters [19]. Together with the molecular data, the knowledge of spatial processes promises more mechanistic understanding of population dynamics and biogeography.

In this study, samples of *T. officinae* were collected from 16 different countries of the world and molecular data from nrDNA and cpDNA barcode loci combined. We constructed topologies based upon combined datasets of barcode loci and performed haplotype

and recombination analysis to uncover the phylogeny, evolution process and current geographical distribution patterns in the level of species.

Furthermore, a historical biogeography model of *T. officinale* was constructed by applying statistical dispersal-vicariance analysis (S-DIVA) and Bayesian binary MCMC (BBM) analysis implemented in RASP (Reconstruct Ancestral State in Phylogenies) and linking outcomes of this model with spatial process. It is proposed that wind-mediated seed dispersal could have played an important role in the diversification of *T. officinale*. Congruence of *T. officinale* population history with some climatic events illustrate the relevance of spatial process for determining the dynamics of population and historical biogeography of *T. officinale*. Previous studies have mostly focused on phylogeny of *T. officinale* [8, 13, 20]. To best of our knowledge, no any specialized study has described evolution and biogeography of *T. officinale* by using MLST approach and spatial processes.

Materials and methods

Dandelion sample collection

Leaf samples of *T. officinale* were collected from various locations of China or donated by researchers in the respective countries. Taraxacum plants are highly invasive and can be found abundantly on roads sides, gardens and rural areas. No special permissions were required for collection of plant samples. All samples were totally dried to null the propagation capability before shipment. Therefore, no special permissions were required for shipment and transportation process. This study did not involve endangered or protected species. Foreign samples were donated by samplers themselves or by the researchers from Department of Agriculture, Azad University of Gorgan, Iran; Institute of Agricultural Sciences, University of the Punjab, Pakistan. Details of sampling localities are shown in supplemental table A in [S1 Tables](#).

DNA isolation

For fresh leaf samples, 50–100 mg of leaf material was crushed in liquid nitrogen. Genomic DNA was then extracted using the Sangon Biotech Plant DNA extraction kit (Shanghai, China) following manufacturer's instructions. For dried leaf samples, leaf material was superficially cleaned through moderate friction using absorbent paper. DNA was extracted by adopting a modified CTAB method [21]. The quantity and quality of extracted DNA were determined by nano-drop spectrophotometer (NanoDrop 2000, Thermo Scientific).

PCR amplification and DNA sequencing

A pilot study was performed to select DNA barcode loci showing more than 90% PCR amplification efficacy. Afterwards, 4 candidate DNA barcode loci were amplified and sequenced from the extracted DNA of the leaf samples. These included two coding cpDNA loci 'matK and rbcL', one non-coding cpDNA intergenic spacer loci 'psbA-trnH' and a nuclear DNA locus 'ITS2' ([Table 1](#)). Previously established primers were used to make PCR amplifications of these selected barcode loci. Details of the primers were listed in supplementary table B in [S1 Tables](#). PCR was performed in a reaction volume of 30 μ L, using Sangon Biotech (Shanghai China) nTaq 2X reaction mixture. PCR amplifications were obtained on A100 ThermoCycler (LongGene, China) using 20 ng of template genomic DNA. Amplified products were inspected by electrophoresis on 1% agarose gel prior to DNA sequencing.

Table 1. The profile of DNA barcode loci for the selected *Taraxacum* samples.

Parameter assessed	DNA barcode locus			
	matK	rbcL	ITS2	psbA-trnH
Number of samples	70	70	70	70
PCR Success rate	100%	97%	100%	93%
Sequencing Success rate	100%	100%	100%	100%
Sequence Length (bp)	699–761	529–637	427–472	403–431
Aligned Length (bp)	761	637	472	431
Numbers of Variable sites	86	78	131	56
Parsimony informative sites	43	37	56	29

<https://doi.org/10.1371/journal.pone.0203275.t001>

Sequence variation and population genetic structure

DNA sequences were aligned using Clustal X2 [22] and converted in FASTA format for further analysis. The data was used individually and also in combination of barcode loci (MLST approach) to generate genetic distance, species discrimination and phylogenetic tree construction. Inter- and intra-specific genetic distances were calculated according to Kimura 2-parameter model [23] in MEGA 7 software [24]. The accuracy of species discrimination for each DNA barcodes was calculated using “best match”, “best close match”, and “all species barcodes” criteria as proposed previously [25]. Genetic diversity, recombination analysis and neutrality tests were performed in DNAspV5 [26].

Haplotype networks were generated to depict relationship among *Taraxacum* population belonging to different geographical regions with PopArt (<http://popart.otago.ac.nz>). Networks were constructed under 95% parsimonious branch connections. Fu’s F_s and Tajima’s D approaches were used to collect information regarding demographic history of *Taraxacum* population [27, 28].

Potential recombinants along with parental analysis and approximation of possible recombination breakpoint positions in *T. officinale* haplotypes were analyzed by RDP version 4 [29]. Recombination detection was performed using GENECONV [30], Maximum Chi Square [31] Chimaera [32, 33] and SisterScan [34] methods with default settings. Any potential recombinant event detected by two or more methods were used in this study. A breakpoint map containing the positions of all clearly identifiable breakpoints was compiled after discovery of a set of unique recombination events.

To support our biogeographical model, we estimated pairwise population differentiation using Slatkin’s linearized F_{ST} [35] and the absolute number of migrants (Gene flow) between population using Arlequin ver. 3.0 software [36]. The results were generated from 10,000 random permutations as described [37].

Phylogenetic analyses and divergence time estimation

Phylogenetic inferences between haplotypes were generated using Maximum Likelihood (ML) method in MEGA 7 [24]. Every haplotype in the tree was represented by all barcode loci. The best-fitted model of sequence evolution was tested in TOPALi (<http://www.topali.org/>) and the model parameter values were used to construct ML tree. Clade support was estimated using 1000 bootstrap pseudo replicates [38].

Divergence time was estimated using a Bayesian approach with BEAST software version 1.6.1 [70] using ITS datasets. Two runs of 108 chains were conducted, sampling every 1,000 generations. The settings used were the Yule tree prior, the HKY substitution model, three gamma categories and strict molecular clock. ITS substitution rate for asteraceae is 2.51×10^{-9}

substitutions per site yr^{-1} based on *Eupatorium* spp. [23]. The best-fit model of sequence evolution was chosen by TOPALi as mentioned earlier. Tracer 1.5 (<http://beast.bio.ed.ac.uk/Tracer>) was used to check for convergence of the Markov chains and adequate effective sample size. A maximum clade credibility tree was obtained using Tree-Annotator software.

Biogeographic analyses

The statistical dispersal-vicariance analysis (S-DIVA) and Bayesian binary MCMC (BBM) were implemented to infer ancestral distribution using RASP 3.2 (Yu et al. 2015). The trees previously constructed by BEAST analysis were used as the input data. Four regions were defined according to the sampling and distribution range of *Taraxacum*; Eur = Europe, WA = West Asia (Iran and Iraq), SA = South Asia (Pakistan), EA = East Asia (China, Korea). The MCMC chains were run for 5×10^6 generations in all analyses. The data were sampled every 1000 generations, and the temperature for heating the chains was 0.1.

Results

PCR amplification and DNA sequencing of *Taraxacum* samples

T. officinale samples were collected globally from 33 different sites across two continents. Four DNA barcoding loci were selected after performing pilot studies. These loci were successfully amplified for all plant samples. In accordance with the previous knowledge, all four loci used in this study are considered highly polymorphic in nature [12]. Here 132 barcode sequences were produced. “Genome Workbench” software was used to perform BLAST analysis for molecular identification of plant samples based on these amplified loci. An alignment was generated using CLUSTAL W, covering four barcode genes for different *T. officinale* samples. This alignment was used for detection of recombination, construction of phylogenetic trees and haplotype networks. All the loci were amplified three times to get the best close match. Details of PCR amplifications are provided in [Table 1](#).

Interspecific genetic variations

After sequencing of all four DNA barcode loci, the aligned sequence lengths ranged from 427 to 452 bps. The number of variable sites ranged from 16 to 37 bases in the aligned gene sequences. On average basis, 6.3% nucleotide positions varied across all samples of *T. officinale*. The intra-specific distance ranged between 0.08 to 0.94% with an average value of 0.46% (Supplementary Table C in [S1 Tables](#)). Both gaps and insertions were seen in sequences of these barcode loci when aligned together for further analysis. Generally, ITS2 and psbA-trnH exhibited highest mean intra-specific variations respectively (Supplementary Table C in [S1 Tables](#)).

Species identification and discrimination

Parameters developed by [25] were used to search the best match and the best close match. These loci (ITS2, matK, rbcL and psbA-trnH) showed the best close match of 67.3, 58.6, 47.1 and 63.5% respectively ([Fig 1](#)). While performing the barcode analysis, these individual barcode candidate loci showed relatively less success rate as compared to the best close match and the best match ([Fig 1](#)). ITS2 had the highest success rates in the barcode analysis ([Fig 1](#)).

Phylogenetic analysis

We used the maximum likelihood (ML) approach to obtain a consistent phylogeny based on nrDNA and cpDNA datasets. *P. purpurae* was used as out-group in topology construction. All the gene sequences were combined in the form of one dataset (multi locus sequence typing

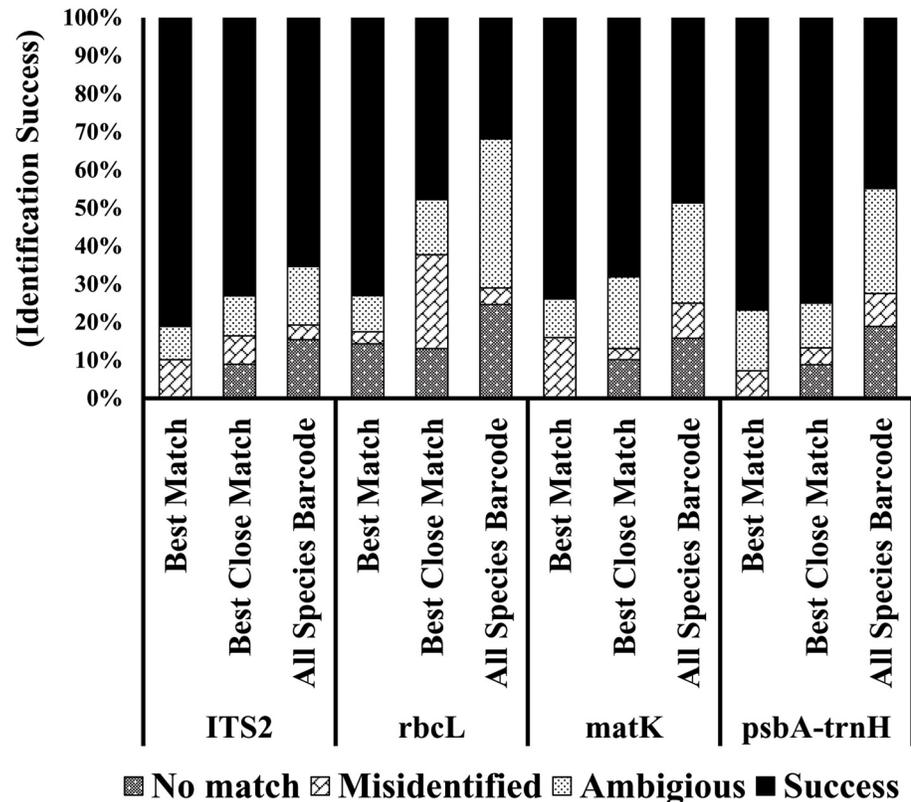


Fig 1. The sequence identification success rates of different barcode loci. Criteria for best match, best close match and all species barcode is based on Meier et al., 2006.

<https://doi.org/10.1371/journal.pone.0203275.g001>

data ‘MLST’ approach) to provide an accurate phylogeny for the construction of phylogenetic trees. Haplotypes of different geographical origin fell into clearly separated clades (Fig 2). Tree branches mostly showed high bootstrap values. Some members with erratic behavior negatively affected the bootstrap value of occupied branches. Data obtained from individual DNA barcode genes was not sufficient to obtain deep phylogeny and varying specific details of relationship and related perspectives were seen in ML based phylogenetic trees based on individual barcode gene sequences (Fig 2).

All the clades on ML tree can be clearly divided into two main groups “1 and 2”. Group 1 contained haplotypes of East and South Asian countries (Fig 2). Group 2 was containing haplotypes from Europe and West Asian countries like Iran and Iraq.

Population biology and phylogeographic structure

From all *T. officinale* samples, 31 haplotypes were identified with some singletons and the remaining contained more than one (Fig 3). Haplotype diversities varied from 0.24–0.54 with nucleotide diversity ranging from 0.023 to 0.35 (Table 2). Furthermore, Tajima’s D and Fu’s Fs tests were employed to detect deviation from neutrality (Table 2). These results demonstrate that population differentiation of *T. officinale* was driven by different selective forces and evolutionary processes. Positive values of these tests provide evidence of balancing selection or of a population subdivision; whereas, negative values indicate positive selection and demographic history of *T. officinale*.

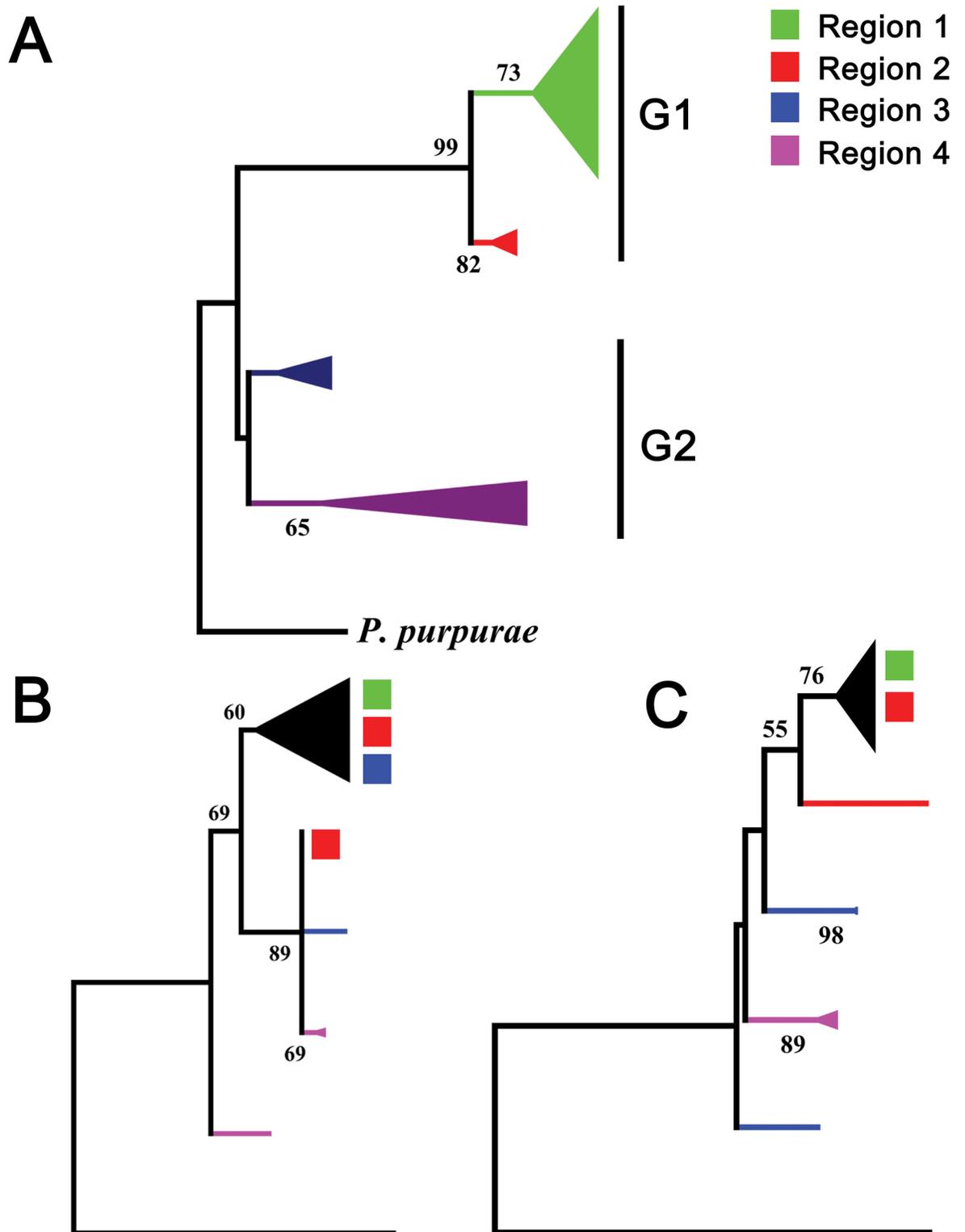


Fig 2. A maximum likelihood phylogenetic tree estimated from four different barcode loci. Phylogenetic trees were inferred from either dataset of all barcode loci (A) or by using ITS (B) and cp DNA (C) datasets alone. Values on nodes indicate boot strap support. Region 1 = East Asia (China and Korea). Region 2 = South Asia (Pakistan), Region 3 = West Asia (Iran and Iraq), Region 4 = Europe. The size of triangle is proportional to the number of haplotypes.

<https://doi.org/10.1371/journal.pone.0203275.g002>

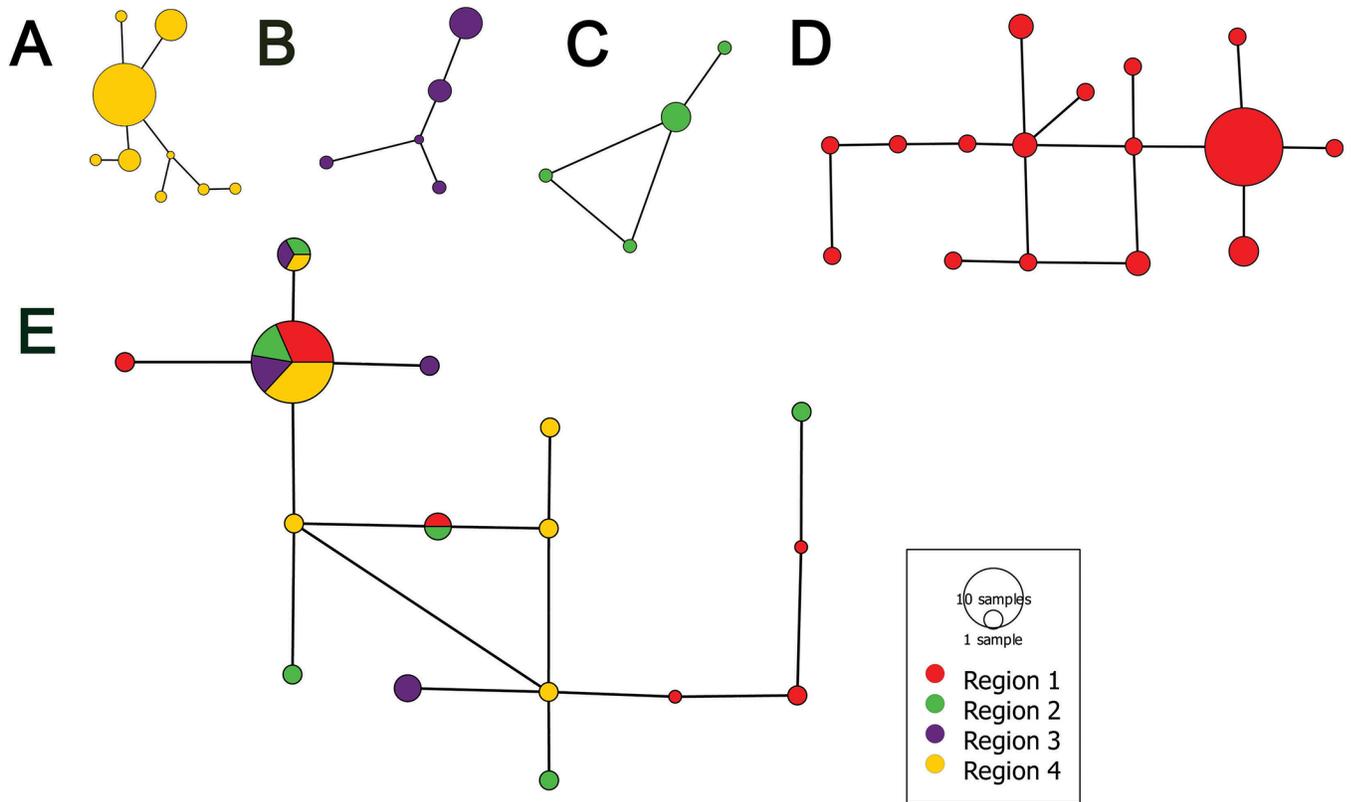


Fig 3. Haplotype analysis of populations of *T. officinales* belonging to different geographical origins. The analysis was performed using “PopART” software based on ‘Minimum Spanning Network’ approach. Haplogroups were made based on datasets of all four barcode loci. A = Haplotype network of samples belonging to region 4 (Europe). B = Haplotype network of samples belonging to region 3 (West Asia). C = Haplotype network of samples belonging to region 2 (South Asia). D = Haplotype network of samples belonging to region 1 (East Asia). E = Network of haplotypes belonging to all four sample regions.

<https://doi.org/10.1371/journal.pone.0203275.g003>

Typing of *T. officinale* plant samples originated from several countries revealed the presence of some recombination events in members of different geographical origins (Fig 4). The permutation test indicated that the distribution of breakpoints was significantly (Fig 4). Different recombination events were detected by the selected recombination detection methods (Fig 4). The recombination analysis showed the exchange of genetic material between members of Europe and Asian lineages. Details of recombination events are provided in supplementary table E in S1 Tables.

The gene flow analysis was performed using sequences datasets of *Taraxacum* species. Relatively lower genetic differentiation (F_{ST}) was seen between population of Europe and Asian regions with some non-significant indices (Table 3). Accordingly, considerable gene flow was

Table 2. The nucleotide polymorphism of *Taraxacum* population.

Region	n	H	nd	hd	Tajima’s D	Fu’s Fs
Europe	28	H1-H09	0.283	0.54	-1.79 ($P < 0.05$)	-4.863 ($P < 0.01$)
West Asia	13	H10-H14	0.078	0.31	-0.58 ($P > 0.10$)	-1.07 ($P < 0.05$)
South Asia	19	H15-H18	0.023	0.24	0.22 ($P > 0.10$)	-0.08 ($P > 0.10$)
East Asia	38	H18-H31	0.358	0.39	-0.82 ($P > 0.10$)	-3.98 ($P < 0.01$)

n = Number of samples, H = Haplotypes, hd = the haplotype diversity, nd = the nucleotide diversity, p -values are indicated in parenthesis

<https://doi.org/10.1371/journal.pone.0203275.t002>

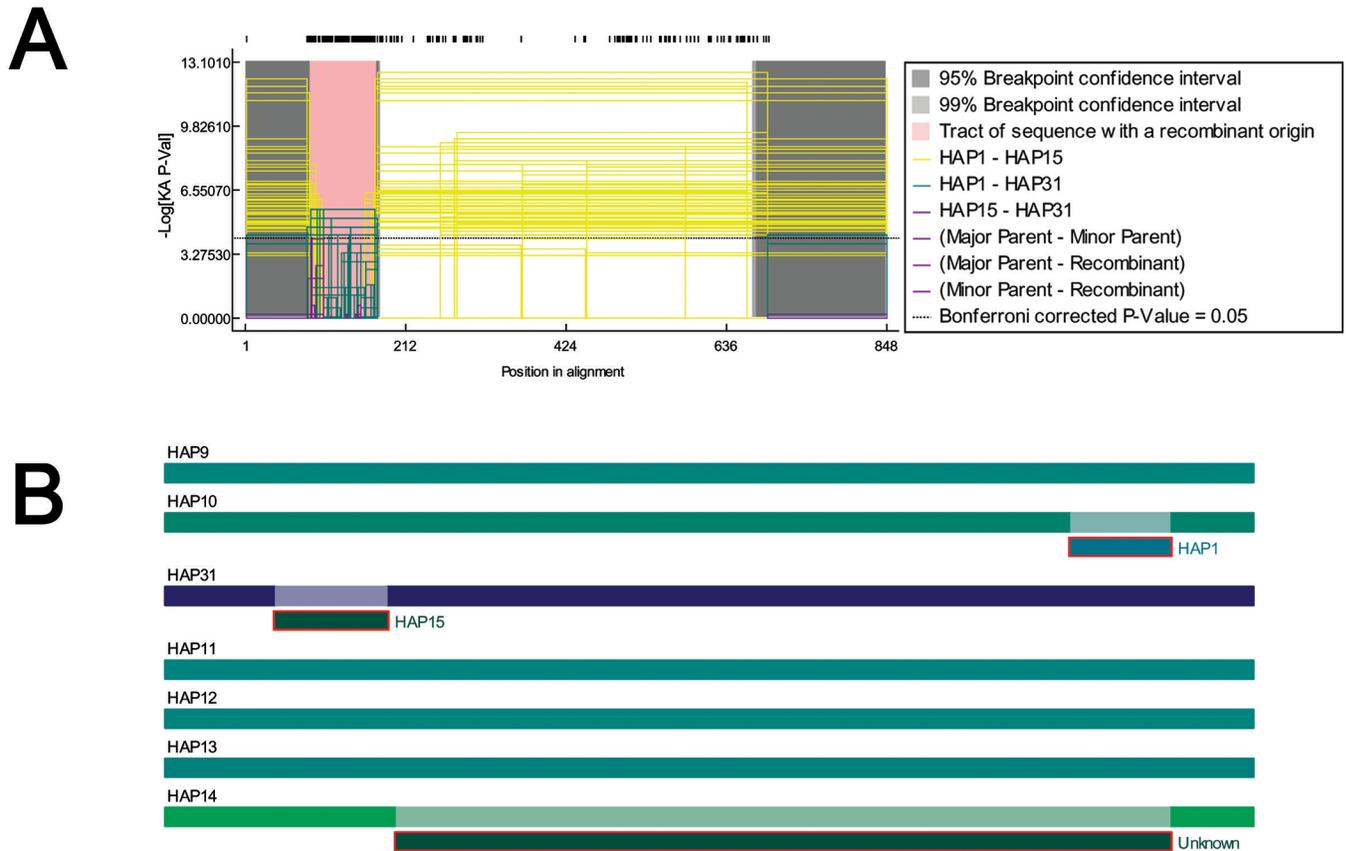


Fig 4. Recombination analysis among different population of *T. officinale*. Analysis was performed in software “RDP version 4”. A = GeneCov plot of potential recombination events along with putative parents. The black blocks above the plot represents positions of informative alleles. KA-P value represents P value based on Blast-Like Karlin-Altschul & Permutation analysis. Colored lines inside plot represents scorings of recombinants. B = An overview of recombination segments among different haplotypes of *T. officinale*.

<https://doi.org/10.1371/journal.pone.0203275.g004>

seen between *Taraxacum* populations of Europe and Asian regions. The number of migrants (gene flow) among these populations ranged from 0.18 to 3.75 (Table 3).

Historical biogeography and divergence time estimates

Lastly, biogeography of *T. officinale* was constructed by performing S-DIVA and Bayesian binary MCMC analyses (Fig 5). The root node of *T. officinale* is situated in Europe with subsequent dispersal and vicariance events between the ancestral and areas of present distribution.

Table 3. Pairwise genetic differentiation (FST) of *T. officinale* populations of different regions (below diagonal) and gene flow analysis among different populations (above diagonal).

Regions	Region 1	Region 2	Region 3	Region 4
Region 1	-	2.64	0.18	0.78
Region 2	0.04	-	1.62	3.75
Region 3	0.18	0.23*	-	1.28
Region 4	0.39*	0.08	0.28*	-

Region 1 = Europe, Region 2 = West Asia, Region 3 = South Asia, Region 4 = East Asia.

* = P<0.05

<https://doi.org/10.1371/journal.pone.0203275.t003>

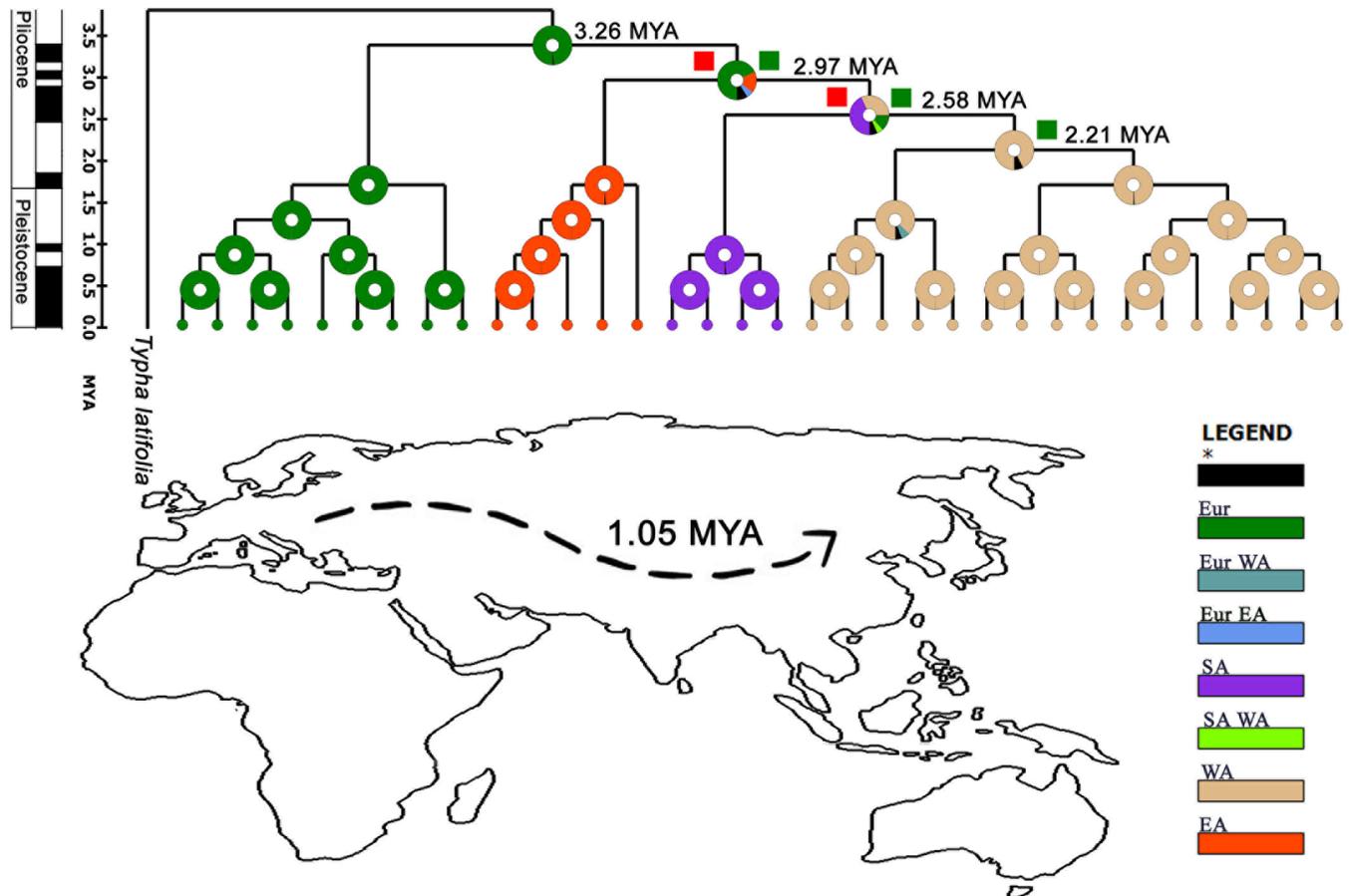


Fig 5. Historical biogeography of *T. officinale* and ancestral area reconstruction. A = Reconstruction of the ancestral area of *T. officinale* and related taxa. The pie charts at each node were obtained using S-DIVA and Bayes MCMC analysis. Upper case letters represent different regions: Eur = Europe. SA = South Asia (Pakistan), WA = West Asia (Iran, Iraq) EA = East Asia (China, Korea), Green square boxes represent dispersal events and red represents vicariance. B = Putative dispersal route of *T. officinale* as inferred from S-DIVA and Bayes MCMC analysis.

<https://doi.org/10.1371/journal.pone.0203275.g005>

Some previous records also confirm that *T. officinale* is originated from Europe [39]. RASP inferred biogeography of *T. officinale* was also supported by rest of the analysis (phylogeny, haplotype, recombination events, parentage and gene flow analysis). Some potential haplotypes were detected with major and minor parents from Europe and Asia. Furthermore, higher gene flow was observed among population of Europe and Asia. These facts support as Europe is likely ancestral area of *T. officinale*. This west-to-east scenario of dispersion across Europe and Asia is also supported by tree topologies inferred from sequences of all four loci. Members of Europe and Asian population were sharing common lineages in phylogenetic trees in case of *T. officinale* (Figs 2 and 3).

The ITS data was used to calculate divergence time among population of Europe and Asian regions of *T. officinale*. Divergence time was calculated based on average substitution rate of population of all identified species. ITS substitution rate for asteraceae is 2.51×10^{-9} substitutions per site yr^{-1} based on *Eupatorium* spp.[40] that will obviously give us a rough estimate. This suggests that East Asian population evolved 1.05 million years ago from European parental population (Fig 5).

Discussion

Dandelions (*T. officinale*) are the plants that have been the topic of scientific and medicinal research for more than a century. In this study, plant samples were collected from 16 different countries of the world. We generated sequence datasets of four different barcode loci and used the data for the construction of species relationship, evaluation of divergence and historical biogeography.

An ideal DNA barcode locus must possess short sequence length, to make it easy for PCR, capable to discriminate up to species level and generate sufficient information to separate genomic dataset [41]. The use of combination of barcode loci is considered as the most suitable approach for identification of genera and genetic divergence analysis [42]. The use of two barcode loci was proposed for phylogenetic studies of different plant families in some earlier researches [43]. It was proved to be inconsistent in later studies [44, 45]. In the current investigation, the use of only chloroplast DNA (cpDNA) or genomic DNA was not sufficient to provide better resolution. Previous studies regarding *T. officinale* [46, 47] also show weaknesses in species resolution by using only cpDNA. The possible reasons can be low substitution rate and high structural mutations with homoplasy [13]. In view of the results of this study combination of all four loci “ITS2 + matK + rbcL + psbA-trnH” is proposed best for barcoding of *T. officinale* (Supplementary table C in S1 Tables). The possible reasons are high percentages of the best match and best close match values, higher PCR success and sequencing rate and ultimately high resolving power of genetic divergence.

The variation of lengths in amplified sequences of different species can be considered possible limitation due to insertion, deletion or genetic re-arrangements [48]. Use of only non-encoding regions are not sufficient in genetic study because of difficulty in logical alignment of obtained genomic sequence loci [49]. Therefore, the sets of barcode loci containing both encoding and non-encoding regions were used in this study. This approach is suitable for both species identification and phylogenetic studies requiring comprehensive sequence alignment. Furthermore, this approach has made it possible to target genetic variations at multiple parts of the genome.

There have been five attempts (including the present study) to construct *Taraxacum* phylogeny with varying barcode loci and sample sizes [8, 13, 46, 47]. Some common findings were the low substitution rates and high length variation with indels presence [13]. Unlike previous studies, plant samples were collected from vast geographical range to uncover maximal variation within the population.

Topological incongruence among genes trees derived from individual gene datasets (SLST approach) was seen at some branches in phylogenetic analysis of *T. officinale*. Phylogeny resolution is often effected by discordant gene events [50]. The possible factors can be gene duplication, recombination and incomplete gene sorting [51]. Some recombination events were detected in *T. officinale*. This can be the likely reason for this incongruence along with the hybridization process. All haplotypes of *T. officinale* were divided into two main groups; Group A and B (Fig 2). The group A comprised of two distinct clades containing haplotypes from China and south Asian regions (Pakistan). Haplotypes from Europe and West Asia region got sister lineage and differentiated in two distinct clades in group B.

In order to infer the connections of *T. officinale* population, haplotype analysis was performed based on dataset of all barcode loci. In most of the cases there was considerable levels of population connectivity across different geographical regions (Fig 3). It seems to be a very weak effect of geographical distance on gene flow process. Secondly, the lack of geographical structuring among haplotypes network of *T. officinale* represents recent demographic events and a population growth. Large haplo-groups containing members from different continents

prove weak population differentiation and large number of recombination events among different population (Fig 3).

Overall, *T. officinale* population connectivity seems to be not limited by distance at the European and Asian regions. Detection of a number of re-combination events (Fig 4) among Europe and Asian population further support this hypothesis.

Finally, we implied integrative approach based on all analysis (S-DIVA and MCMC analyses, phylogeny construction, haplotype, recombination and gene flow analysis) in current study and constructed historical biogeography of *T. officinale* and related taxa and lineages dispersal patterns. The analysis infers that *T. officinale* lineages dispersed from the Europe (putative *Taraxacum* origin area) toward West Asia and from the west toward the East Asian region (Fig 5). This study suggests West Asia is the geographic center of the distribution range for most of the *Taraxacum* taxa examined in current study. The presence of less genetic differentiation (F_{ST}) between European and Asian population of *T. officinale* further supports this model. The recombination and gene flow analysis indicate that population currently occurring in east and west Asian countries received a gene flow from Europe. Some potential recombinants of *T. officinale* were detected in the East Asian regions with major and minor parents from Europe and Asia (Fig 4). Such dispersal routes have most likely occurred because of suitable climatic conditions that allowed spatial displacement of the *T. officinale*.

ITS data was used to calculate divergence time among population of European and Asian regions of *T. officinale*. This suggests East Asian population evolved from Europe parental populations nearly 1.05 million years ago from European parental population (Fig 5). As geographical expansion of *T. officinale* is dated to the Pliocene. This range expansion may have been triggered by the existence of East Asian winter monsoons during the Pliocene [52–54]

Tajima's D and Fu's F_S are the most powerful tools to detect deviation from neutrality and population expansion scenarios [55]. Both of these analysis provide an insight that multiple selective forces and evolutionary processes have been the targets of selection in *T. officinale* evolution. Population with positive values of Tajima's D represents excess of intermediate frequency alleles and can result from population bottlenecks, structure and balancing selection [56]. Most of the populations had negative values for Fu's F_S and Tajima's D. Either positive selection or recent population expansion is the most likely explanation for these values [27]. Thus we can conclude that multiple selective forces were involved in evolutionary process of *T. officinale*. Non-significant values of Tajima's D and Fu's F_S were seen in case of some populations. This can be due to low sample size of given population or simply fits with neutrality. Globally, these results demonstrate that the population differentiation of *T. officinale* species was driven by different selective forces and evolutionary processes for different species. This variable range of selection forces may facilitate adaptation to different environments. The findings of Tajima's D and Fu's F_S are also supported by S-DIVA and Bayes MCMC analysis. These analyses inferred both dispersal and vicariance events at different nodes (Fig 5). Regarding vicariance events, the possible explanation can be seed dispersals by waterfowl, likely giving distribution to far isolated areas [57].

The rapid geographic and population expansion of a species is usually facilitated by the changes of landscape and climate factors [58, 59]. Over the time, ecosystem has experienced large variations in environmental conditions linked to trophic interactions and biogeographical shifts of flora and fauna [60–62]. The regimes of climatic change is often reported as in favor of species abundance and biogeographical expansion [63, 64]. Variability and recent changes in environment along with adaptation to the novel local environments may have contributed to the population and biogeographical expansion of *T. officinale*. It would have been almost equally possible that favorable environmental conditions, large seed dispersal rate of *T. officinale* and fast growth of newly established population helped rapid population expansion

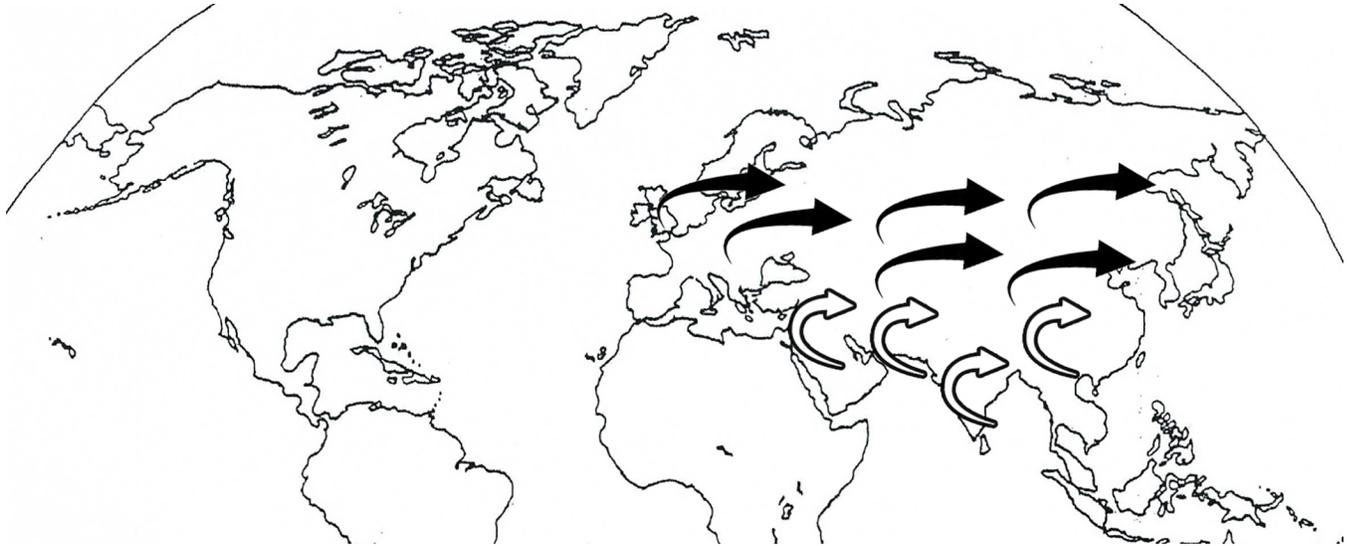


Fig 6. The wind pattern favors *Taraxacum* seed dispersal. Black arrows show westerlies winds whereas white arrows show monsoon wind patterns.

<https://doi.org/10.1371/journal.pone.0203275.g006>

next to Europe and West and East Asian regions. As stated earlier the geographical expansion of *T. officinale* is dated to the Pliocene and may have been triggered by the existence of monsoons weather.

Another important implication of geographic expansion is the introduction of genetic variation by recombination and gene flow process among different species that has the potential to positively affect population dynamics. Recombination is the source of genetic variation upon which other evolutionary forces can then operate. A recombination event can provide a genetic variant that would require many mutation events to be obtained without recombination [65]. Recombination events can thus impact evolution and population dynamics in different ways. They have been associated with the expansion of host ranges and appearance of new species [66]. It shuffles existing mutations to generate new variants that can then be better adapted to the environment and hence speeding up adaptation process [67]. This study provided the evidence of some potential recombination events among different populations of *Taraxacum*. Presence of the recombination events would have improved the adaptability of this plant at new geographical localities. Likewise, recombination in *Taraxacum* can also be associated with the acquisition of novel traits for better adaptation to the new environments. Secondly, recent studies provide a new insight into the adaptation of meiosis during polyploid evolution. Polyploid species can accommodate increased genetic dosage or divergent genomes [68]. The presence of disomic and polysomic inheritance in some *Taraxacum* species reflects positive implications for evolution and genetic improvement to adopt newly inhabited areas.

Spatial processes are important for determining the dynamics of population and biogeography. Seed dispersal is a leading spatial demographic process among sessile organisms setting the spatial template for population expansion [16, 69]. Recent field studies provide an insight of the seed dispersal patterns and their implications in population evolution and genetics [15]. *Taraxacum* seeds can be dispersed over considerable distances, germinate under a variety of conditions and seedlings survive amongst other plants [70, 71]. Previous studies have shown that a close relative of *Taraxacum* (*Hypochaeris radicata*) successfully dispersed via seed in the Netherlands causing substantial levels of gene flow and population expansion [72]. Seasonal differences in wind directionality and speed greatly effects seed dispersal events [73]. Westerlies are the prevailing winds from the west toward east directions (Fig 6). These winds cover

middle latitude where most of the Europe, Middle East and Asian countries are situated. These wind patterns supported the spread of *Taraxacum* population from its putative origin towards East Asian countries. In the same way, the bidirectional wind patterns of Monsoon over Middle East and Asia favor the spatial pattern of seed dispersal between both regions.

Conclusions

In conclusion, our findings suggest that the genus *T. officinale* prevails complex phylogeny and single molecular barcode loci has little power to reveal deep phylogeny. There was considerable incongruence in every gene tree based on individual barcode loci. Use of multi locus sequence typing approach based on datasets of different barcode loci are more efficient in resolving complex *T. officinale* phylogeny and evolutionary events. The results of this study indicate evidence for ongoing gene flow and recombination between European and Asian population of *T. officinale*. Finally, the intra- and intercontinental expansion of *T. officinale* was driven by both biological dispersals and vicariance from Europe to Asian regions. This enigma suggests that environmental factors and evolutionary changes may have been favorable for distribution of *T. officinale*.

Supporting information

S1 Tables. (A) = Details of samples collection sites, sample collector's name and height from the sea. (B) = Details of PCR primers used in this study. (C) = Genetic distance percentage generated using Kimura 2-parameter model analysis for the candidate barcode loci and their combinations. (D) = Details of NCBI Accessions. (E) = A detailed overview of recombination analysis performed by RDP V4 software.
(DOCX)

Acknowledgments

This research was partially supported by National Key R&D Program of China (No. 2017YFD0501500) and the Fundamental Research Funds for the Central Universities Program (No. 2662017PY104 and No. 2662017PY104).

Author Contributions

Conceptualization: Nasim Ahmad Yasin, Fatemeh Sefidkon, Xuebo Hu.

Data curation: Waheed Akram, Yanju Pang, Aqeel Ahmad, Shakeel Ahmed, Nasim Ahmad Yasin, Tehmina Anjum, Basharat Ali, Xiaohua Li, Fatemeh Sefidkon.

Formal analysis: Nasim Ahmad Yasin, Xiangdong Hu, Fatemeh Sefidkon, Xuebo Hu.

Methodology: Mohammadjavad Jafari, Waheed Akram, Yanju Pang, Aqeel Ahmad, Shakeel Ahmed, Nasim Ahmad Yasin, Xiangdong Hu, Xiaohua Li, Matteo Ciprian, Monika Bielec, Xuebo Hu.

Software: Mohammadjavad Jafari, Waheed Akram, Aqeel Ahmad, Nasim Ahmad Yasin, Fatemeh Sefidkon.

Supervision: Xuebo Hu.

Validation: Nasim Ahmad Yasin.

Writing – original draft: Mohammadjavad Jafari, Waheed Akram, Basharat Ali, Xiaohua Li, Shuang Dong, Qian Cai, Sheng Hu, Xuebo Hu.

Writing – review & editing: Waheed Akram, Nasim Ahmad Yasin, Fatemeh Sefidkon, Xuebo Hu.

References

1. Łuczaj Ł, Pieroni A. Nutritional ethnobotany in Europe: from emergency foods to healthy folk cuisines and contemporary foraging trends. *Mediterranean Wild Edible Plants*: Springer; 2016. p. 33–56.
2. Solbrig OT, Simpson BB. Components of regulation of a population of dandelions in Michigan. *The Journal of Ecology*. 1974;473–86.
3. Solbrig OT. The population biology of dandelions. *American Scientist*. 1971; 59:686–94.
4. Stubbendieck J, Friisoe G, Bolick M. *Weeds of Nebraska and the Great Plains*. (Nebraska Department of Agriculture: Lincoln). 1995.
5. Hughes J, Richards AJ. The genetic structure of populations of sexual and asexual *Taraxacum* (dandelions). *Heredity*. 1988; 60(2):161–71.
6. Hórandl E. The complex causality of geographical parthenogenesis. *New Phytologist*. 2006; 171(3):525–38. <https://doi.org/10.1111/j.1469-8137.2006.01769.x> PMID: 16866956
7. Van der Hulst R, Mes T, Falque M, Stam P, Den Nijs J, Bachmann K. Genetic structure of a population sample of apomictic dandelions. *Heredity*. 2003; 90(4):326–35. <https://doi.org/10.1038/sj.hdy.6800248> PMID: 12692586
8. Kirschner J, Štěpánek J, Mes T, Nijs Jd, Oosterveld P, Štorchová H, et al. Principal features of the cpDNA evolution in *Taraxacum* (Asteraceae, Lactuceae): a conflict with taxonomy. *Plant Systematics and Evolution*. 2003; 239(3):231–55.
9. Maiden MCJ. Multilocus sequence typing of bacteria. *Annual Review of Microbiology*. 2006; 60(60):561.
10. Bonaccorso E, Peterson AT, Navarro-Sigüenza AG, Fleischer RC. Molecular systematics and evolution of the *Cyanocorax* jays. *Molecular phylogenetics and evolution*. 2010; 54(3):897–909. <https://doi.org/10.1016/j.ympev.2009.11.014> PMID: 19931623
11. Grubisha LC, Levens N, Olson MS, Lee Taylor D. Intercontinental divergence in the *Populus*-associated ectomycorrhizal fungus, *Tricholoma populinum*. *New Phytologist*. 2012; 194(2):548–60. <https://doi.org/10.1111/j.1469-8137.2012.04066.x> PMID: 22332931
12. Lee SY, Ng WL, Mahat MN, Nazre M, Mohamed R. DNA Barcoding of the Endangered *Aquilaria* (Thymelaeaceae) and Its Application in Species Authentication of Agarwood Products Traded in the Market. *PloS one*. 2016; 11(4):e0154631. <https://doi.org/10.1371/journal.pone.0154631> PMID: 27128309
13. Kirschner J, Drábková LZ, Štěpánek J, Uhlemann I. Towards a better understanding of the *Taraxacum* evolution (Compositae–Cichorieae) on the basis of nrDNA of sexually reproducing species. *Plant Systematics and Evolution*. 2015; 301(4):1135–56.
14. López-Fernández H, Honeycutt RL, Stiassny ML, Winemiller KO. Morphology, molecules, and character congruence in the phylogeny of South American geophagine cichlids (Perciformes, Labroidae). *Zoologica Scripta*. 2005; 34(6):627–51.
15. Nathan R, Muller-Landau HC. Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in ecology & evolution*. 2000; 15(7):278–85.
16. Levin SA, Muller-Landau HC, Nathan R, Chave J. The ecology and evolution of seed dispersal: a theoretical perspective. *Annual Review of Ecology, Evolution, and Systematics*. 2003; 34(1):575–604.
17. Hammond DS, Brown VK, Newbery DM, Prins HHT, Brown ND. Disturbance, phenology and life-history characteristics: factors influencing distance/density-dependent attack on tropical seeds and seedlings. 1998:51–78.
18. Schupp EW, Fuentes M. Spatial patterns of seed dispersal and the unification of plant population ecology. *écoscience*. 1995; 2(3):267–75.
19. Soons M, Ozinga W. How important is long-distance seed dispersal for the regional survival of plant species? *Diversity & Distributions*. 2005; 11(2):165–72.
20. Van der Hulst R, Mes T, Den Nijs J, Bachmann K. Amplified fragment length polymorphism (AFLP) markers reveal that population structure of triploid dandelions (*Taraxacum officinale*) exhibits both clonality and recombination. *Molecular Ecology*. 2000; 9(1):1–8. PMID: 10652071
21. Agostini G, Lüttke R, Echeverrigaray S, de Souza-Chies TT. Genomic DNA extraction from herbarium samples of *Cunila* D. Royen ex L. (Lamiaceae) and *Polygala* L. (Polygalaceae). *Conservation Genetics Resources*. 2011; 3(1):37–9.

22. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic acids research*. 1997; 25(24):4876–82. PMID: [9396791](#)
23. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of molecular evolution*. 1980; 16(2):111–20. PMID: [7463489](#)
24. Tamura K, Stecher G, Peterson D, Filipiński A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution*. 2013; 30(12):2725–9. <https://doi.org/10.1093/molbev/mst197> PMID: [24132122](#)
25. Meier R, Shiyang K, Vaidya G, Ng PK. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic biology*. 2006; 55(5):715–28. <https://doi.org/10.1080/10635150600969864> PMID: [17060194](#)
26. Librado P, Rozas J. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 2009; 25(11):1451–2. <https://doi.org/10.1093/bioinformatics/btp187> PMID: [19346325](#)
27. Tajima F. The effect of change in population size on DNA polymorphism. *Genetics*. 1989; 123(3):597–601. PMID: [2599369](#)
28. Fu Y-X. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*. 1997; 147(2):915–25. PMID: [9335623](#)
29. Martin D, Rybicki E. RDP: detection of recombination amongst aligned sequences. *Bioinformatics*. 2000; 16(6):562–3. PMID: [10980155](#)
30. Padidam M, Sawyer S, Fauquet CM. Possible emergence of new geminiviruses by frequent recombination. *Virology*. 1999; 265(2):218–25. <https://doi.org/10.1006/viro.1999.0056> PMID: [10600594](#)
31. Smith JM. Analyzing the mosaic structure of genes. *Journal of molecular evolution*. 1992; 34(2):126–9. PMID: [1556748](#)
32. Martin D, Posada D, Crandall K, Williamson C. A modified bootscan algorithm for automated identification of recombinant sequences and recombination breakpoints. *AIDS Research & Human Retroviruses*. 2005; 21(1):98–102.
33. Martin DP, Williamson C, Posada D. RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics*. 2005; 21(2):260–2. <https://doi.org/10.1093/bioinformatics/bth490> PMID: [15377507](#)
34. Gibbs MJ, Armstrong JS, Gibbs AJ. Sister-scanning: a Monte Carlo procedure for assessing signals in recombinant sequences. *Bioinformatics*. 2000; 16(7):573–82. PMID: [11038328](#)
35. Slatkin M. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*. 1995; 139(1):457. PMID: [7705646](#)
36. Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*. 2007; 1(4A):47.
37. Arias DM, Albarrán-Lara AL, González-Rodríguez A, Peñaloza-Ramírez J, Dorado O, Leyva E. Genetic diversity and structure of wild populations of the tropical dry forest tree *Jacaratia mexicana* (Brassicaceae) at a local scale in Mexico. *Revista De Biología Tropical*. 2012; 60(1):1–10. PMID: [22458206](#)
38. Felsenstein J. CONFIDENCE LIMITS ON PHYLOGENIES: AN APPROACH USING THE BOOTSTRAP. *Evolution; international journal of organic evolution*. 1985; 39(4):783. <https://doi.org/10.1111/j.1558-5646.1985.tb00420.x> PMID: [28561359](#)
39. StewartWade SM, Neumann S, Collins LL, Boland GJ. The biology of Canadian weeds. 117. *Taraxacum officinale* G. H. Weber ex Wiggers. *Canadian Journal of Plant Science*. 2002; 82(4):825–53.
40. Schmidt GJ, Schilling EE. Phylogeny and biogeography of *Eupatorium* (Asteraceae: Eupatorieae) based on nuclear ITS sequence data. *American Journal of Botany*. 2000; 87(5):716–26. PMID: [10811796](#)
41. Hebert PD, Cywinska A, Ball SL. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences*. 2003; 270(1512):313–21.
42. Group CPW, Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, et al. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences*. 2009; 106(31):12794–7.
43. Van der Bank M, Fay MF, Chase MW. Molecular phylogenetics of Thymelaeaceae with particular reference to African and Australian genera. *Taxon*. 2002; 51(2):329–39.
44. Herber B. Pollen morphology of the Thymelaeaceae in relation to its taxonomy. *Plant Systematics and Evolution*. 2002; 232(1):107–21.
45. Herber B. Thymelaeaceae. *Flowering Plants· Dicotyledons*: Springer; 2003. p. 373–96.
46. Wittzell H. Chloroplast DNA variation and reticulate evolution in sexual and apomictic sections of dandelions. *Molecular Ecology*. 1999; 8(12):2023–35. PMID: [10632854](#)

47. Mes TH, Kuperus P, Kirschner J, Stepanek J, Oosterveld P, Storchova H, et al. Hairpins involving both inverted and direct repeats are associated with homoplasious indels in non-coding chloroplast DNA of *Taraxacum* (Lactuceae: Asteraceae). *Genome*. 2000; 43(4):634–41. PMID: [10984175](#)
48. Goremykin VV, Holland B, Hirsch-Ernst KI, Hellwig FH. Analysis of *Acorus calamus* chloroplast genome and its phylogenetic implications. *Molecular biology and evolution*. 2005; 22(9):1813–22. <https://doi.org/10.1093/molbev/msi173> PMID: [15930156](#)
49. Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, et al. The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *American journal of botany*. 2005; 92(1):142–66. <https://doi.org/10.3732/ajb.92.1.142> PMID: [21652394](#)
50. Ma X, Sun W, Zhu W, Sun H. Resolving the phylogenetic relationships and evolutionary history of the East Asian endemic genus *Rodgersia* (Saxifragaceae) using multilocus data. *Perspectives in Plant Ecology, Evolution and Systematics*. 2017; 25:20–8.
51. Maddison WP. Gene trees in species trees. *Systematic biology*. 1997; 46(3):523–36.
52. Fan M, Song C, Dettman DL, Fang X, Xu X. Intensification of the Asian winter monsoon after 7.4 Ma: Grain-size evidence from the Linxia Basin, northeastern Tibetan Plateau, 13.1 Ma to 4.3 Ma. *Earth and Planetary Science Letters*. 2006; 248(1–2):186–97.
53. Guo Z, Peng S, Hao Q, Biscaye PE, An Z, Liu T. Late Miocene–Pliocene development of Asian aridification as recorded in the Red-Earth Formation in northern China. *Global and Planetary Change*. 2004; 41(3–4):135–45.
54. Vandenberghe J, Lu H, Sun D, van Huissteden JK, Konert M. The late Miocene and Pliocene climate in East Asia as recorded by grain size and magnetic susceptibility of the Red Clay deposits (Chinese Loess Plateau). *Palaeogeography, Palaeoclimatology, Palaeoecology*. 2004; 204(3–4):239–55.
55. Okello JBA, Nyakaana S, Masembe C, Siegismund HR, Arctander P. Mitochondrial DNA variation of the common hippopotamus: evidence for a recent population expansion. *Heredity*. 2005; 95(3):206–15. <https://doi.org/10.1038/sj.hdy.6800711> PMID: [16030528](#)
56. Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 1989; 123(3):585–95. PMID: [2513255](#)
57. Zhu J, Yu D, Xu X. The phylogeographic structure of *Hydrilla verticillata* (Hydrocharitaceae) in China and its implications for the biogeographic history of this worldwide-distributed submerged macrophyte. *BMC evolutionary biology*. 2015; 15(1):95.
58. Estrada-Peña A. Increasing habitat suitability in the United States for the tick that transmits Lyme disease: a remote sensing approach. *Environmental health perspectives*. 2002; 110(7):635. <https://doi.org/10.1289/ehp.110-1240908> PMID: [12117639](#)
59. Ostfeld RS, Canham CD, Oggenfuss K, Winchcombe RJ, Keesing F. Climate, deer, rodents, and acorns as determinants of variation in Lyme-disease risk. *PLoS Biol*. 2006; 4(6):e145. <https://doi.org/10.1371/journal.pbio.0040145> PMID: [16669698](#)
60. Hurrell J. Decadal trends in the North Atlantic Oscillation: regional temperatures and precipitation. *Oceanographic Literature Review*. 1996; 2(43):116.
61. Beaugrand G, Reid PC, Ibanez F, Lindley JA, Edwards M. Reorganization of North Atlantic marine copepod biodiversity and climate. *Science*. 2002; 296(5573):1692–4. <https://doi.org/10.1126/science.1071329> PMID: [12040196](#)
62. Brander K, Blom G, Borges M, Erzini K, Henderson G, editors. Are we seeing a coherent response to changing temperature? *ICES Marine Science Symposia*; 2003.
63. Blanchard F, Vandermeirsch F. Warming and exponential abundance increase of the subtropical fish *Capros aper* in the Bay of Biscay (1973–2002). *Comptes rendus biologies*. 2005; 328(5):505–9. PMID: [15948639](#)
64. Farina A, Freire J, González-Gurriarán E. Demersal fish assemblages in the Galician continental shelf and upper slope (NW Spain): spatial structure and long-term changes. *Estuarine, Coastal and Shelf Science*. 1997; 44(4):435–54.
65. Arenas M, Araujo NM, Branco C, Castelhana N, Castro-Nallar E, Pérez-Losada M. Mutation and recombination in pathogen evolution: Relevance, methods and controversies. *Infection, Genetics and Evolution*. 2017. <https://doi.org/10.1016/j.meegid.2017.09.029>.
66. Martin DP, Biagini P, Lefevre P, Golden M, Roumagnac P, Varsani A. Recombination in Eukaryotic Single Stranded DNA Viruses. *Viruses*. 2011; 3(9):1699. <https://doi.org/10.3390/v3091699> PMID: [21994803](#)
67. Pérez-Losada M, Arenas M, Galán JC, Palero F, González-Candelas F. Recombination in viruses: Mechanisms, methods of study, and evolutionary consequences. *Infection, Genetics and Evolution*. 2015; 30:296–307. <https://doi.org/10.1016/j.meegid.2014.12.022> PMID: [25541518](#)

68. Liu B, Xu C, Zhao N, Qi B, Kimatu JN, Pang J, et al. Rapid genomic changes in polyploid wheat and related species: implications for genome evolution and genetic improvement. *Journal of Genetics and Genomics*. 2009; 36(9):519–28. [https://doi.org/10.1016/S1673-8527\(08\)60143-5](https://doi.org/10.1016/S1673-8527(08)60143-5) PMID: 19782953
69. Tilman D, Kareiva PM. *Spatial ecology: the role of space in population dynamics and interspecific interactions*: Princeton University Press; 1997.
70. Stewart-Wade S, Neumann S, Collins L, Boland G. The biology of Canadian weeds. 117. *Taraxacum officinale* GH Weber ex Wiggers. *Canadian Journal of Plant Science*. 2002; 82(4):825–53.
71. Isselstein J, Hofmann M. Zur Etablierung und Konkurrenzkraft des Gewöhnlichen Löwenzahns (*Taraxacum officinale* Web.) im Grasland. *VERHANDLUNGEN-GESELLSCHAFT FÜR ÖKOLOGIE*. 1996; 26:519–26.
72. Mix C, Arens P, Rengelink R, Smulders M, Van Groenendael J, Ouborg N. Regional gene flow and population structure of the wind-dispersed plant species *Hypochaeris radicata* (Asteraceae) in an agricultural landscape. *Molecular Ecology*. 2006; 15(7):1749–58. <https://doi.org/10.1111/j.1365-294X.2006.02887.x> PMID: 16689895
73. Wright SJ, Trakhtenbrot A, Bohrer G, Detto M, Katul GG, Horvitz N, et al. Understanding strategies for seed dispersal by wind under contrasting atmospheric conditions. *Proceedings of the National Academy of Sciences*. 2008; 105(49):19084–9.