Genomics xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Genetic diversity and natural selection in wild fruit flies revealed by whole-genome resequencing

Ting Lian[a,1], Diyan Li[a,*,1], Xinxin Tan[b,1], Tiandong Che[a], Zhongxian Xu[a], Xiaolan Fan[a], Nan Wu[a], Long Zhang[c], Uma Gaur[a], Boyuan Sun[a], Mingyao Yang[a,*]

[a] Institute of Animal Genetics and Breeding, College of Animal Science and Technology, Sichuan Agricultural University, Chengdu 611130, China
[b] Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China
[c] Institute of Ecology, China West Normal University, Nanchong 637009, China

ARTICLE INFO

Keywords:
Fruit fly
Genetic diversity
Natural selection
Whole-genome resequencing

ABSTRACT

We characterized 26 wild fruit flies comparative population genomics from six different altitude and latitude locations by whole genome resequencing. Genetic diversity was relatively higher in Ganzi and Chongqing populations. We also found 13 genes showing selection signature between different altitude flies and variants related to hypoxia and temperature stimulus, were preferentially selected during the flies evolution. One of the most striking selective sweeps found in all high altitude flies occurred in the region harboring *Hsp70Aa* and *Hsp70Ab* on chromosome 3R. Interestingly, these two genes are involved in GO terms including response to hypoxia, unfolded protein, temperature stimulus, heat, oxygen levels. Mutation in *HPH* gene, a candidate gene in the hypoxia inducible factor pathway, might contributes to hypoxic high-altitude adaptation. Intriguingly, some of the selected genes, primarily utilized in humans, were involved in the response to hypoxia, which could imply a conserved molecular mechanisms underlying high-altitude adaptation between insects and humans.

1. Introduction

An earlier study revealed a remarkable phenotype in adult *Drosophila* which can tolerate low-oxygen environments extremely well [1]. The high-altitude flies have adapted to the naturally extreme environments, such us reduced barometric pressure, extreme temperature and hypoxia for a long time. For survival at the high altitudes, flies must enhance the motility performance, reduce their body weight and size, cell number and cell size in wing, and rates of mitochondrial oxygen consumption and reactive oxygen species production should also be reduced [2–6]. Thus, flies maintain a steady body condition and physiological metabolic performance in order to adapt to the extremely high-altitude environment, which prompted us to explore the genetic evidence of adaptive evolution.

In recent years, many reports suggested the involvement of Notch pathway in hypoxia tolerance, based on genome-wide dissection of hypoxia adaptation over generations in *Drosophila* under laboratory conditions [7,8]. Adaptive evolution studies on various species have identified multiple candidate genes responsible for high-altitude adaptation, including Tibetan wild boar [9,10], Tibetan Mastiffs [11], humans [12], Tibetan antelope [13], birds [14,15], yak [16], tiger and snow leopard [17], which provided strong evidences to enhance our understanding of potential evolutionary mechanisms underlying genetic adaptation from low to high altitude.

Here, we reported whole-genome re-sequencing results of 26 indigenous Chinese wild female flies collected from field, with approximately 27.4 fold genomic coverage, enabling us to comprehensively analyze the adaptive evolution of the wild animals with natural selection.

2. Results

2.1. Genome sequencing, variation calling, genetic diversity and population genetics

We selected 26 wild field *Drosophila* lines representing 6 geographically diverse locations from low to high altitude in China (Table 1). Whole-genome pair-end (2 × 125 bp) sequencing of all the 26 samples were performed on Hiseq2500 platform (Illumina), and a total of 831million reads were generated with an average coverage depth of > 27.4 fold per genome. All reads were aligned to the reference genome assembly of *Drosophila melanogaster* [18], and >

* Corresponding authors.
E-mail addresses: diyanli@sicau.edu.cn (D. Li), yangmingyao@sicau.edu.cn (M. Yang).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ygeno.2017.12.002
Received 11 July 2017; Received in revised form 20 November 2017; Accepted 5 December 2017
0888-7543/ © 2017 Elsevier Inc. All rights reserved.

Table 1
SNPs annotation and genetic diversity.

Population	Ganzi (2552 m)	Chongqing (2238 m)	Yunnan (1901 m)	Xichang (1542 m)	Yaan (580 m)	Deyang (491 m)	Chengdu (500 m)
Northern latitude	30°03′7.77″	N29°06′6.32″	N24°52′57.82″	N27°53′52.44″	N30°01′4.73″	N31°07′44.21″	N30°34′21.43″
East longitude	E101°57′39.69″	E107°11′47.24″	E102°49′54.47″	E102°15′41.62″	E103°02′16.19″	E104°23′42.77″	E104°03′44.18″
Sample size (n)	3	3	5	3	5	5	2
Average sequencing depth	42.91	26.52	27.88	20.38	20.05	25.07	39.01
Average genome coverage rate (%)	80.01	94.52	93.94	97.21	92.93	77.29	91.14
Total SNPs	4,016,433	1,878,152	1,842,191	1,356,768	1,757,952	1,603,381	1,338,515
Coding SNPs	845,945	442,467	266,382	200,969	269,486	255,716	203,439
Fixed SNPs	44,988	85,899	58,882	80,018	61,222	56,686	93,024
Unique SNPs	399,574	16,087	27,341	11,983	25,926	24,753	9967
Insertions	79	189	45	25	31	59	30
Deletions	1050	1228	1929	1286	1633	1719	1026
Inversions	3	5	10	10	14	8	4
Nucleotide diversity (π)	0.489	0.208	0.082	0.061	0.081	0.078	0.091

Note: 'Fixed SNP' is the conserved SNP in all flies of the same population.

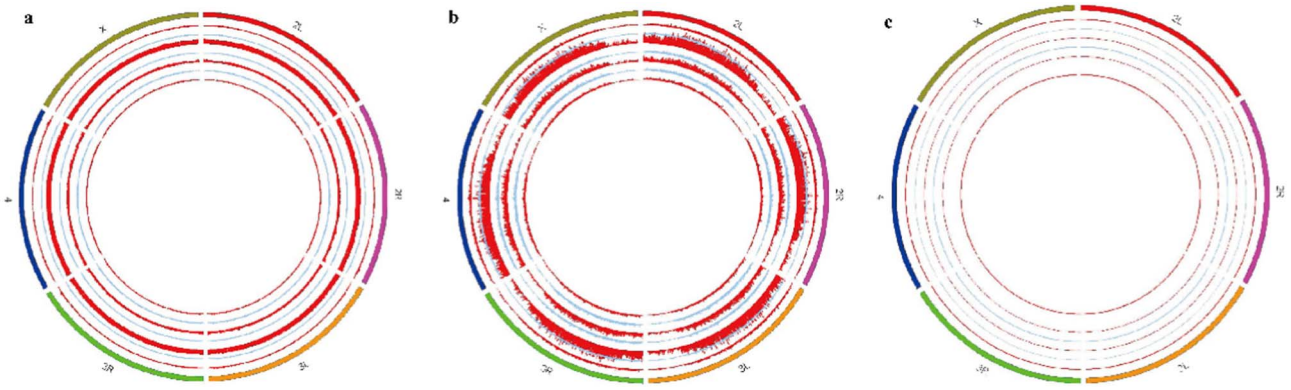

Fig. 1. Circos plot of global distribution of variants along the flies' genomes. (A) Circos plot of global distribution of π in all 26 flies' genomes. (B) Circos plot of global distribution of θ in all 26 flies' genomes. (C) Circos plot of global distribution of F_{ST} in all 26 flies' genomes. The circles from outside to inside shows Yunnan, Xichang, Ganzi, Deyang, Chongqing, Chengdu, Yaan, respectively.

89.16% genome coverage was achieved (Table 1 and Table S1), improving the credibility to call variants. After stringent SNP calling using whole-genome sequencing data, a total of 1.34–4.02 million (M) SNPs for each population have been identified by SAMtools, the highest SNP number were observed in fruit flies from highland (Ganzi), with 4.02 M SNPs, which is more than three times than Xichang and Chengdu flies, and more than two times than Yunnan, Deyang, Yaan and Chongqing flies (Table 1). At the genome level, Chongqing and Ganzi fruit flies exhibited higher genetic diversity than the rest of populations (0.208 in Chongqing and 0.489 in Ganzi flies respectively) (Table 1, Fig. 1 and Fig. S1). As expected, the high altitude flies had a relatively higher heterozygous SNP ratio (0.05–0.26 in Chongqing and 0.09–0.12 in Ganzi flies) compared to the other populations (0.05–0.09) (Table S2). Structural variants (including deletion, insertion and inversion) in each sample were calculated (Table S3 and S4). Total number of structural variants ranged from 146 to 944 (Table S3), and total length of these structural variants ranged from 203,816 to 2,994,007 base pairs (Table S4). Interestingly, deletion accounted for most of the structural variants. Over 80% of the InDels were 1–5 bp in length in CDS region among all the populations and an obvious enrichment of most of InDels were in the multiple of 3 bp in CDS region (Fig. S2).

Phylogenetic analysis based on genome-wide SNPs using the neighbor-joining (NJ) method demonstrated strong clustering of samples according to the three altitude gradients: Ganzi and Chongqing flies comprised the highest altitude flies gradient, whereas Yunnan and Xichang flies clustered in median altitude flies and Yaan, Chengdu and Deyang flies comprised the low altitude flies (Fig. 2). Interestingly, some low altitude flies mixed with median altitude flies. These distinct distribution patterns and expansion signatures suggested that the divergent low altitude flies may have originated from different regions, such as Yunnan, Chongqing and/or surrounding areas.

2.2. Signature of selection in wild fruit flies

To detect the signatures of natural selection associated with high altitude, we searched the fruit fly genome for regions with high F_{ST} (coefficient of nucleotide differentiation) values among the populations living at high (Ganzi and Chongqing), middle (Xichang and Yunnan), and low altitude (Yaan, Chengdu and Deyang) (Table 1). We scanned the chromosomes with a 3-kb window sliding with a step size of 500 bp and calculated the F_{ST} value for each window. In total, 37 unique chromosomal regions containing 13 candidate genes were identified (*CG17600*, *CG33939*, *CG40470*, *CG7556*, *CG9902*, *CR43941*, *CR44833*, *Gprk1*, *Hsp70Aa*, *Hsp70Ab*, *Ir41a*, *SPR* and *Snap25*) by comparison between high and middle altitude flies (Fig. 3). An extended differentiated genomic region was observed on chromosome 3R, embed *Hsp70Ab* gene with the highest F_{ST} value of 0.338 (Table 2).

These genes were involved in "glutamate signaling pathway", "aminopeptidase activity", "DNA binding", "phosphate metabolic process", "response to hypoxia", "response to temperature stimulus", "G-protein coupled receptor protein signaling pathway", "mating behavior" and "regulation of neurotransmitter levels", biological processes (Table S5). Out of these 13 genes, three (*CG9902*, *Hsp70Aa* and *Hsp70Ab*) were also identified by comparison between high and low altitude flies. *Hsp70Aa* and *Hsp70Ab* were involved in "Spliceosome", "Protein processing in endoplasmic reticulum", "Endocytosis" and "Longevity regulating pathway - multiple species" KEGG pathways, whereas *CG9902* is a novel gene with unknown function.

We further focused on mutations in these 13 genes between populations (Table S6 and Table S7). For *Hsp70Aa* and *Hsp70Ab*, only one

Download English Version:

https://daneshyari.com/en/article/8646306

Download Persian Version:

https://daneshyari.com/article/8646306

Daneshyari.com