RESEARCH ARTICLE

CrossMark

Mitochondrial DNA sequence variation in Drosophilid species (Diptera: Drosophilidae) along altitudinal gradient from Central Himalayan region of India

MANISHA SARSWAT, SAURABH DEWAN and RAJENDRA SINGH FARTYAL*

Cytogenetics and Molecular Systematics Laboratory, UGC-SAP Department of Zoology and Biotechnology, HNB Garhwal University, Garhwal 246 174, India

Abstract

Central Himalayan region of India encompasses varied ecological habitats ranging from near tropics to the mid-elevation forests dominated by cool-temperate taxa. In past, we have reported several new records and novel species from Uttarakhand state of India. Here, we assessed genetic variations in three mitochondrial genes, namely, 16S rRNA, cytochrome c oxidase subunit I and cytochrome c oxidase subunit II (*COI* and *COII*) in 26 drosophilid species collected along altitudinal transect from 550 to 2700 m above mean sea level. In the present study, overall 543 sequences were generated, 82 for 16S rRNA, 238 for *COI*, 223 for *COII* with 21, 47 and 45 mitochondrial haplotypes for 16S rRNA, *COI* and *COII* genes, respectively. Almost all species were represented by 2–3 unique mitochondrial haplotypes, depicting a significant impact of environmental heterogeneity along altitudinal gradient on genetic diversity. Also for the first time, molecular data of some rare species like *Drosophila mukteshwarensis*, *Liodrosophila nitida*, *Lordiphosa parantillaria*, *Lordiphosa ayarpathaensis*, *Scaptomyza himalayana*, *Scaptomyza tistai*, *Zaprionus grandis* and *Stegana minuta* are provided to public domains through this study.

Introduction

The Himalayan range is among the most intricate and diverse mountain systems in the world. It forms distinct geological and ecological entity, influencing climate and biotic aspects of the region. The varying topography promotes environmental heterogeneity at both temporal and spatial scales affecting diversity and distribution patterns of biodiversity elements. In India, Himalaya amasses around 18–20% of the total geographical area and holds above 50% of the country's forest, with 40% of its endemic species. Uttarakhand state which is located in the central Himalayan region encompasses highly varied tropical to temperate regimes in span of just few hundred kilometres and due to its variable altitudinal terrain, offers an abode for rich, luxuriant flora and fauna. Through our extensive explorations in past decade, we have documented more than 90 species from this region (Sati *et al.* 2013), with significant number new to the world of science, in comparison around 300 drosophilid species recorded so far from varied ecogeographical zones in India (Gupta 2005; Kumar and Ajai 2009). This high ratio of diversity affirms Uttarakhand state as a potential drosophilid diversity hotspot with implication in molecular systematics and evolutionary biological research.

Drosophilidae is a large family of acalyptrate diptera with worldwide distribution. The first catalogue listed more than 2500 species belonging to 55 genera in this family (Wheeler 1981) with two subfamilies, Steganinae and Drosophilinae. Yassin *et al.* (2010) recently documented 4000 described species distributed in 76 genera. However, at present, the database on taxonomy of drosophilidae compiled by Gerhard Bächli (dated March 2015) reports 6633 species of which, *Drosophila* Fallén 1823 is the largest with more than 1150 species recorded so far (Bächli 2015). *Drosophila* is also among the most versatile and best-studied model organisms in modern biology with 12 fully-sequenced genomes now made available in the public domain by *Drosophila* 12 Genomes Consortium (2007). In spite of availability of such robust genetic and genomic resources, meagre is known so far about molecular phylogeny, ecological preferences and

*For correspondence. E-mail: fartyalrs@gmail.com, rs.fartyal@hnbgu.ac.in.

Keywords. genetic diversity; cytochrome c oxidase subunit I; cytochrome c oxidase subunit II; 16S rRNA; Uttarakhand; haplotypes.

evolutionary history of species inhabiting Indian subcontinent, especially biodiversity-rich Uttarakhand region of India. Although researchers in this region have documented several novel species reflecting high degree of endemism, sexual dimorphism and host inclination (Singh *et al.* 1985; Singh and Bhatt 1986, 1988; Singh and Negi 1989, 1992, 1995; Singh and Dash 1998; Singh *et al.* 2000, 2004; Singh and Fartyal 2002; Fartyal and Singh 2002, 2004; Fartyal *et al.* 2005, 2013), no molecular studies have yet been performed using mitochondrial DNA-based markers. These markers can determine the levels of genetic differentiation among various geographically distributed and ecologically isolated populations of drosophilids from this region. This study provides a resolute data of the extent to which these populations differ genetically; permitting interpretation of the evolutionary patterns of reproductive isolation and speciation mechanisms. As an important evolutionary mechanism in speciation is the diversification of genes among populations, mitochondrial DNA-based bar-coding is much acknowledged as an efficient tool not only for rapid species identification, but also for discovery and delimitation of species, i.e. species discovery or molecular taxonomy (Hebert *et al.* 2003b; Jinbo *et al.* 2011; Bast 2015).

Intraspecific molecular analysis is also a reliable approach to decipher how populations have been influenced by historical and existing processes (Emerson *et al.* 2001; Posada and Crandall 2001). mtDNA has been extensively used for reconstructing phylogenies and tracing historical demographies (Avise *et al.* 1987; Moritz *et al.* 1987; Harrison 1989; Avise 1991, 2000, 2004) in closely related taxa and populations within species owing its low recombination rate, maternal inheritance, conserved structure, reduced effective population size and comparatively high evolutionary rate (Brown *et al.* 1979; Moritz *et al.* 1987; Avise 2004). Even though with some drawback, mtDNA study provides an efficient way to detect gene flow, reproductive isolation, species boundaries and population structure (Rubinoff and Holland 2005). Thus, in the present study, we assessed intraspecific variation in three mitochondrial genes, namely, 16S rRNA, cytochrome c oxidase subunit I and cytochrome c oxidase subunit II (*COI* and *COII*) in 26 drosophilid species collected from varying habitats along an altitudinal transect between 550 and 2700 m above mean sea level (table 1). Several known and some novel species were sampled, making this the first molecular analysis to date from the central Himalayan region of India, which would aid in understanding various isolating mechanisms operating and their impact on genetic divergence among genomes as a whole. This also provides a base data to identify emerging reproductive isolating mechanisms and evolutionarily diverging populations before speciation.

Table 1. Sampling locations of different drosophilid species from Uttarakhand with their GenBank accession numbers for partial mitochondrial sequences of 16S rRNA, *COI* and *COII*.

| Code | Species | Sampling station | GenBank accession number | | |
			16S rRNA	*COI*	*COII*
BUS	*D. busckii*	SG, AG, UC, MD, KK, CP	KP730750–KP730763	KP730832–KP730842	KP731070–KP731086
IMM	*D. immigrans*	SG, MD, KK, CP	KP730764–KP730768	KP730843–KP730854	KP731087–KP731101
BUZ	*D. buzzatii*	SG, MD	KP730769–KP730772	KP730855–KP730865	KP731102–KP731105
LAC	*D. lacertosa*	SG, MD	KP730773–KP730776	KP730866–KP730876	KP731106–KP731109
REP	*D. repleta*	SG, AG, UC, MD, KK, CP	–	KP730877–KP730886	KP731110–KP731123
BIF	*D. bifasciata*	SG, MD, CP	KP730777–KP730786	KP730887–KP730899	KP731124–KP731139
HUB	*D. hubeiensis*	CP, NT	KP730787–KP730789	KP730900–KP730910	KP731140–KP731143
JAM	*D. jambulina*	SG, CP		KP730911–KP730919	KP731144–KP731153
KIK	*D. kikkawai*	SG, MD, CP	KP730790-KP730792	KP730920–KP730930	KP731154–KP731163
MEL	*D. melanogaster*	SG, AG, UC, MD, KK, CP	KP730793- KP730808	KP730931–KP730943	KP731164–KP731175
NEP	*D. nepalensis*	SG, AG, UC, MD, KK, CP		KP730944–KP730954	KP731176–KP731191
PUL	*D. pulchrella*	SG, CP	–	–	KP731192–KP731201
PUN	*D. punjabiensis*	MD, CP	–	KP730955–KP730965	KP731202–KP731211
SUZ	*D. suzukii*	SG, AG, UC, MD, KK, CP	–	KP730966–KP730981	–
TAK	*D. takahashii*	SG, MD	KP730809–KP730812	KP730982–KP730992	KP731212–KP731222
TRA	*D. trapezifrons*	SG, MD	–	KP730993–KP731003	–
MUK	*D. mukteshwarensis*	NT	KP730813–KP730814	KP731004–KP731006	KP731223–KP731225
NIT	*Liodrosophila nitida*	SG, MD	KP730815–KP730818	KP731007–KP731013	KP731226–KP731235
NEO	*Lordiphosa neokurokawai*	SG, MD, CP	KP730819–KP730825	KP731014–KP731027	KP731236–KP731244
PAR	*Lordiphosa parantillaria*	NT	KP730826–KP730827	KP731028–KP731029	KP731245–KP731247
AYA	*Lordiphosa ayarpathaensis**	NT	KP730828–KP730829	KP731030–KP731031	KP731248–KP731249
HIM	*Scaptomyza himalayana*	SG, AG, UC, MD, KK, CP	–	KP731032–KP731043	KP731250–KP731263
TIS	*Scaptomyza tistai*	CP	–	KP731044–KP731047	–
MIN	*Stegana minuta**	NT	KP730830–KP730831	KP731048–KP731050	KP731264–KP731266
GRA	*Zaprionus grandis*	SG, AG, UC, MD, KK, CP	–	KP731051–KP731053	KP731267–KP731282
IND	*Zaprionus indianus*	SG, AG, UC, MD, KK, CP	–	KP731054–KP731069	KP731283–KP731292

*Novel species, morphological data yet to be published. SG, Srinagar Garhwal; AG, Augustyamuni; UC, Upper Chamoli; MD, Mandal; KK, Kanchula Kharak; CP, Chopta; NT, Nainital.

Materials and methods

Sample collections

Extensive sampling surveys were carried out for a year along altitudinal transect, covering varied ecological habitats ranging from near tropic-like regimes to mid-elevation forests dominated by cool-temperate taxa. Stations sampled were Srinagar-Garhwal (550 m asl, 30°22′N and 78°78′E, Pauri-district), Augustyamuni (800 m asl, 30°39′N and 79°02′E, Rudraprayag-district), Upper Chamoli (1150 m asl, 30°24′N and 79°21′E, Chamoli-district), Mandal (1600 m asl, 30°46′N and 79°26′E, Chamoli-district), Kanchula-Kharak (2100 m asl, 28°43′N and 77°34′E, Chamoli-district) and Chopta (2700 m asl, 30°29′N and 79°10′E, Rudraprayag-district) located in Uttarakhand. The cosmopolitan drosophilid species were collected over rotten fruits, vegetables, fungi and flowers, while rare species were sampled from their natural habitats employing range of techniques which are as follows.

Trap bait: Small containers baited with yeasted banana or other fermenting fruits as oranges, tomato, guava, apples, etc., were suspended through strings from bushes and trees. The method was very effective for collecting species particularly belonging to the two subgenera, *Sophophora* and *Drosophila* of the genus *Drosophila*. The banana trap design of Prof Toda was also used (Toda and Kitching 2002).

Net sweeping: Collection using a modified insect net over natural feeding sites as decaying fruits and leaves, wild grasses, etc. was done to capture flies of other genera of the family drosophilidae which are rarely attracted towards fermenting fruits.

Collection with aspirator: Rare flies showing occasional appearance which could not be collected by above methods was caught directly with the help of an aspirator while they were either courting or resting over the leaves, petals, fungi, etc.

Identification and morphological study

Collected flies were etherized, categorized and subsequently identified using published monographs and guide to species identification (Gupta 2005; Markow and O'Grady 2006) and online identification tools like BioCIS, JDD and FlyBase. External morphologies namely, head, thorax, wing and body, length of adult flies were examined under the stereo zoom microscope (Magnus MS24 model, India, at 3× objective zoom) and metric characters were measured with an ocular micrometer. For further confirmation, the detailed structures of male and female terminalia were observed under trinocular microscope (Magnus, India MLX-DX model, at 10× magnification). The respective genital organs were detached from the adult body and cleared by warming in 10% KOH solution around 100°C for 20–30 min and observed in a droplet of glycerol. The morphological terminology and the definitions of measurements and indices mostly followed were of McAlpine (1981); Zhang and Toda (1992); Hu and Toda (2001). The examined specimens of all species were deposited in the Systematics, Cytogenetics and Molecular Systematics Laboratory, Department of Zoology, H. N. B Garhwal University, Uttarakhand, India.

DNA extraction

DNA isolation procedure was standardized for single fly. Each fly was mashed in 50 μL of fly lysis buffer (20 mM Tris-Cl, pH 8.5, 1.5 mM EDTA, 30 mM NaCl and 250 μg/mL of freshly prepared Proteinase K). The mixture was incubated at 25–37°C for 45 min. Proteinase K was inactivated at 95°C incubation for 5 min. Three μL of this preparation was used per 25 μL PCR reaction mix.

Amplification and sequencing

16S rRNA gene was amplified in 25 μL reaction mix which included 1× *Taq* polymerase buffer, 0.2 mM of each dNTP, 0.2 mM of each forward and reverse primers, 2 mM MgCl$_2$, 1 unit of *Taq* polymerase using Gradient DNA Thermo-Cycler (Techne, UK). The primers used for the amplification of the partial 16S rRNA gene were 16sF (5′-CCGGTTTGAACTCAGATCACGT-3′) and 16sR (5′-CGCCTGTTTAACAAAAACAT-3′) (Simon *et al.* 1994). The thermal cycling conditions were: initial denaturation for 3 min at 94°C, followed by 39 repetitions consisting of denaturation at 94°C for 45 s, annealing at 50°C for 1 min and extension at 72°C for 1.5 min and the final extension at 72°C for 5 min.

COI and *COII* genes were also amplified in a final concentration of 25 μL volumes with 1× *Taq* polymerase buffer, 0.2 mM of each dNTP, 0.2 mM of each forward and reverse primers, 2 mM MgCl$_2$, 1 unit of *Taq* polymerase. The primers used for the amplification were: *COI* (F: 5′-CCAGCTGGAGGAGGAGATCC-3′ and R: 5′-CCAGTAAATAATGGGTATCAGTG-3′) (Palumbi 1996) and *COII* (F: 5′-ATGGCAGATTAGTGCAATGG-3′ and R: 5′-GTTTAAGAGACCAGTACTTG-3′) (Simon *et al.* 1994). For both genes, an initial denaturation step for 3 min at 94°C was performed, followed by 39 repetitions consisting of denaturation at 94°C for 30 s, annealing at 55°C for 45 s and extension at 72°C for 1 min and the final extension at 72°C for 7 min.

The PCR products were visualized on 1.5% agarose gels and the most intense products were selected for sequencing. Products were labelled with the Big-Dye Terminator ver. 3.1 Cycle Sequencing kit (Applied Biosystems, Foster City, USA) and sequenced bidirectionally using an ABI 3730 capillary sequencer following the manufacturer's instructions.

Sequence analysis

Sequences of all the three mitochondrial genes were separately aligned using the program Clustal X 2.0 (Larkin *et al.* 2007). Differences in length were resolved by inserting alignment gaps and positions that could not be aligned, unambiguously were excluded. The degree of sequence disparity between species was calculated by averaging pairwise comparisons of sequence differences across all individuals. For all three genes i.e., 16S rRNA, *COI* and *COII* analysed separately at least four to six individuals of each species per sampling station were sequenced and aligned to yield a final alignment. Overall, base composition, number of transition and transversion from aligned sequences and pairwise evolutionary distance among mitochondrial haplotypes were determined under Kimura 2-parameter (K2P) method using molecular evolutionary genetic analysis (MEGA) software, ver. 6.0 (Tamura *et al.* 2013). This method gives better distance estimates than most other methods when the rates of transitional and transversional substitutions are different. MEGA 6.0 was also used to create a neighbour-joining (NJ) tree based on K2P model as recommended by Hebert *et al.* (2003a). Statistical support for nodes on the tree was evaluated using 1000 bootstrapping replications. The number of polymorphic sites, G+C content, number of haplotypes, haplotype diversity, variance of haplotype diversity, nucleotide diversity and average number of nucleotide differences were determined for each population by DnaSP 5.10.1 (Librado and Rozas 2009).

Different tests of neutrality were applied to molecular data obtained from the altitudinally isolated populations using the DnaSP 5.10.1 (Librado and Rozas 2009), the tests used were Tajima's D test of selective neutrality (Tajima 1989) and Fu and Li's F and D tests (Fu and Li 1993). These tests of neutrality though do not detect departure from neutrality to same degree, yet their combination allows inference to be made on the patterns of selection affecting a specific/selected region of DNA.

Results

mt16S rRNA sequence analysis

We sequenced a total 82 individuals of 15 species for partial sequence analysis of 16S rRNA gene (GenBank accession numbers KP730750–KP730831). Simplicity and unambiguity were observed among all the sequences. Sequencing of the 16S rRNA gene produced an average of 541 bp ranging from 432 to 847 nucleotide base pairs. The analysis revealed nucleotide frequencies as A = 39.27%, T/U = 38.15%, C = 11.97% and G = 10.60%. Pair-wise nucleotide differences and sequence divergences are given in table 1 a&b in electronic supplementary material at http://www.ias.ac.in/jgenet/. Overall 21 unique mitochondrial haplotypes were observed in 82 individuals of 15 drosophilid species. The number of polymorphic sites, G+C content, number of haplotypes, haplotype diversity, variance of haplotype diversity, nucleotide diversity and average number of nucleotide differences are provided in table 2a. All the 21 different haplotypes and their consensus sequences detected in 16S rRNA gene are shown in figure 1a along with their NJ tree in figure 2a. All neutrality tests showed a significant deviation from the null hypothesis of neutrality, however, for *D. busckii*, Fu and Li's F* and D* test indicated departure from neutrality; Tajima's D test did not show any significant departure from neutrality (table 3a). 16S rRNA gene sequence for *D. mukteshwarensis*, *Liodrosophila nitida*, *Lordiphosa parantillaria*, *Lordiphosa ayarpathaensis* and *Stegana minuta* are provided for the first time in public domains through this study.

mtCOI and COII sequence analyses

COI and *COII* genes were also amplified and sequenced for 238 (GenBank accession numbers KP730832–KP731069) and 223 (KP731070–KP731292) individuals, respectively, to determine genetic variation. No insertions, deletions or stop codons were observed in any of the sequence for both the genes. Overall, 25 species were analysed for *mtCOI* gene variation with an average of 598 bp ranging from 407 to 867 base pairs. The nucleotide frequencies observed were A = 28.90%, T/U = 39.52%, C = 14.89% and G = 16.68% with estimated transition/transversion bias (*R*) of 1.26. While for *mtCOII* gene, 23 species were analysed generating an average of 679 bp ranging from 369 to 805 base pairs. The nucleotide frequencies observed were A = 33.01%, T/U = 39.84%, C = 13.51% and G = 13.65% and transition/transversion bias (*R*) of 1.46. Substitution pattern and rates were estimated under K2P method. Pair-wise nucleotide differences and sequence divergences for *COI* and *COII* genes are provided in table 1, c–f in electronic supplementary material.

Table 2(a). 16S rRNA partial gene sequence based diversity analysis.

Species	Number of segregating site	G+C content	Number of haplotype (h)	Haplotype diversity (Hd)	Variance of Hd	Nucleotide diversity per site (π)	Sampling variance of π	Average number of nucleotide difference (k)
D. bifasciata	15	0.242	2	0.556	0.00556	0.01231	0.0000027	8.33333
D. busckii	26	0.221	3	0.692	0.00425	0.02462	0.0000056	11.96703
D. melanogaster	11	0.209	3	0.708	0.00207	0.01206	0.0000009	5.20833
Lordiphosa neokurokawai	16	0.236	2	0.571	0.01428	0.01782	0.0000139	9.14286

Table 2(b). *COI* partial gene sequence based diversity analysis.

Species	Number of segregating sites	G+C content	Number of haplotypes (h)	Haplotype diversity (Hd)	Variance of Hd	Nucleotide diversity per site (π)	Sampling variance of π	Average number of nucleotide differences (k)
D. busckii	6	0.366	2	0.545	0.00522	0.00804	0.0000011	3.27273
D. immigrans	16	0.304	2	0.545	0.00379	0.01355	0.0000023	8.72727
D. buzzatii	10	0.335	2	0.545	0.00522	0.01082	0.0000021	5.45455
D. lacertosa	20	0.330	2	0.545	0.00522	0.01299	0.0000030	10.90909
D. repleta	20	0.327	2	0.556	0.00556	0.01443	0.0000037	11.11111
D. bifasciata	19	0.312	2	0.538	0.00362	0.01881	0.0000044	10.23077
D. hubeiensis	11	0.308	2	0.545	0.00522	0.01224	0.0000026	6.00000
D. jambulina	19	0.288	2	0.556	0.00813	0.01276	0.0000043	10.55556
D. kikkawai	19	0.297	2	0.545	0.00522	0.01241	0.0000027	10.36364
D. melanogaster	12	0.312	2	0.538	0.00362	0.01177	0.0000017	6.46154
D. nepalensis	17	0.297	2	0.545	0.00522	0.01948	0.0000067	9.27273
Lordiphosa neokurokawai	18	0.293	3	0.714	0.00273	0.01685	0.0000046	8.40659
D. punjabiensis	21	0.307	2	0.545	0.00522	0.01321	0.0000031	11.45455
D. suzukii	17	0.339	3	0.708	0.00207	0.01262	0.0000007	8.12500
D. takahashii	14	0.305	2	0.545	0.00522	0.01108	0.0000022	7.63636
D. trapezifrons	12	0.332	2	0.545	0.00522	0.01604	0.0000045	6.54545
Liodrosophila nitida	15	0.313	2	0.571	0.01428	0.01718	0.0000129	8.57143
Scaptomyza himalayana	15	0.329	2	0.545	0.00379	0.01490	0.0000028	8.18182
Zaprionus indianus	13	0.339	3	0.708	0.00207	0.00994	0.0000009	6.04167

Table 2(c). *COII* partial gene sequence based diversity analysis.

Species	Number of segregating site	G+C content	Number of haplotype (h)	Haplotype diversity (Hd)	Variance of Hd	Nucleotide diversity per site (π)	Sampling variance of π	Average number of nucleotide difference (k)
D. busckii	20	0.293	3	0.706	0.00180	0.01392	0.0000010	9.57353
D. immigrans	6	0.272	2	0.476	0.00847	0.00417	0.0000006	2.85714
D. repleta	35	0.260	3	0.714	0.00273	0.02776	0.0000042	18.79121
D. bifasciata	26	0.250	3	0.708	0.00207	0.01715	0.0000021	12.62500
D. jambulina	17	0.247	2	0.556	0.00556	0.01440	0.0000037	9.44444
D. kikkawai	17	0.250	2	0.556	0.00556	0.01598	0.0000046	9.44444
D. melanogaster	13	0.265	2	0.545	0.00379	0.00913	0.0000011	7.09091
D. nepalensis	22	0.257	3	0.708	0.00207	0.01335	0.0000027	10.41667
Lordiphosa neokurokawai	19	0.268	2	0.556	0.00813	0.01868	0.0000092	10.55556
D. pulchrella	14	0.279	2	0.556	0.00556	0.01250	0.0000028	7.77778
D. punjabiensis	9	0.327	2	0.556	0.00556	0.01355	0.0000033	5.00000
D. takahashii	12	0.279	2	0.545	0.00522	0.00987	0.0000017	6.54545
Liodrosophila nitida	25	0.279	2	0.556	0.00556	0.01732	0.0000054	13.88889
Scaptomyza himalayana	20	0.296	3	0.714	0.00273	0.01195	0.0000012	9.39560
Zaprionus grandis	20	0.264	3	0.708	0.00207	0.01298	0.0000055	9.16667
Zaprionus indianus	16	0.275	2	0.556	0.00556	0.01104	0.0000022	8.88889

Forty-seven unique mitochondrial haplotypes were found for *COI* while 46 for *COII* gene. Table 2, b&c, displays the number of polymorphic sites, G+C content, number of haplotypes, haplotype diversity, variance of haplotype diversity, nucleotide diversity and average number of nucleotide differences. All the mitochondrial haplotypes with their consensus sequences are shown in figure 1, b&c along with their NJ tree in figure 2, b&c. Here also for almost all the species, neutrality tests showed a significant deviation from the null hypothesis of neutrality across all localities.

Table 3 b&c, shows the results of neutrality for both the genes. For *D. busckii mtCOI* region, Fu and Li's D* test did not indicate departure from neutrality. In *mtCOII* gene analysis for *D. immigrans* only Fu and Li's F* test showed significant departure from neutrality, while for *D. repleta*, Tajima's D was the only test that failed to depart from neutrality.

Mitochondrial *COI* and *COII* barcode for some rare species, namely, *D. mukteshwarensis*, *Liodrosophila nitida*, *Lordiphosa parantillaria*, *Lordiphosa ayarpathaensis*,
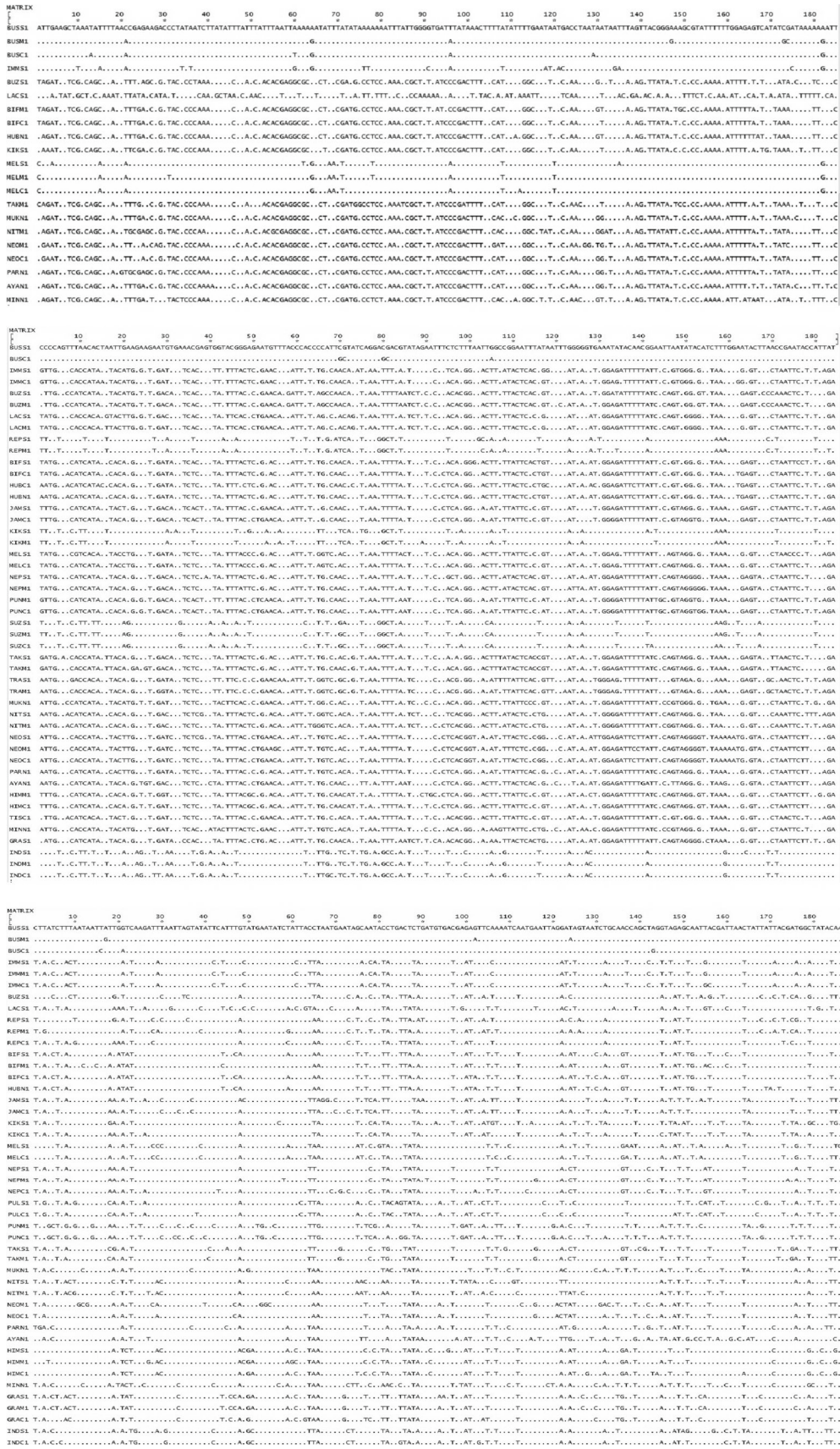
Figure 1. Different haplotypes with their consensus sequences for partial mitochondrial (a) 16S rRNA, (b) *COI* and (c) *COII* genes (top to bottom).
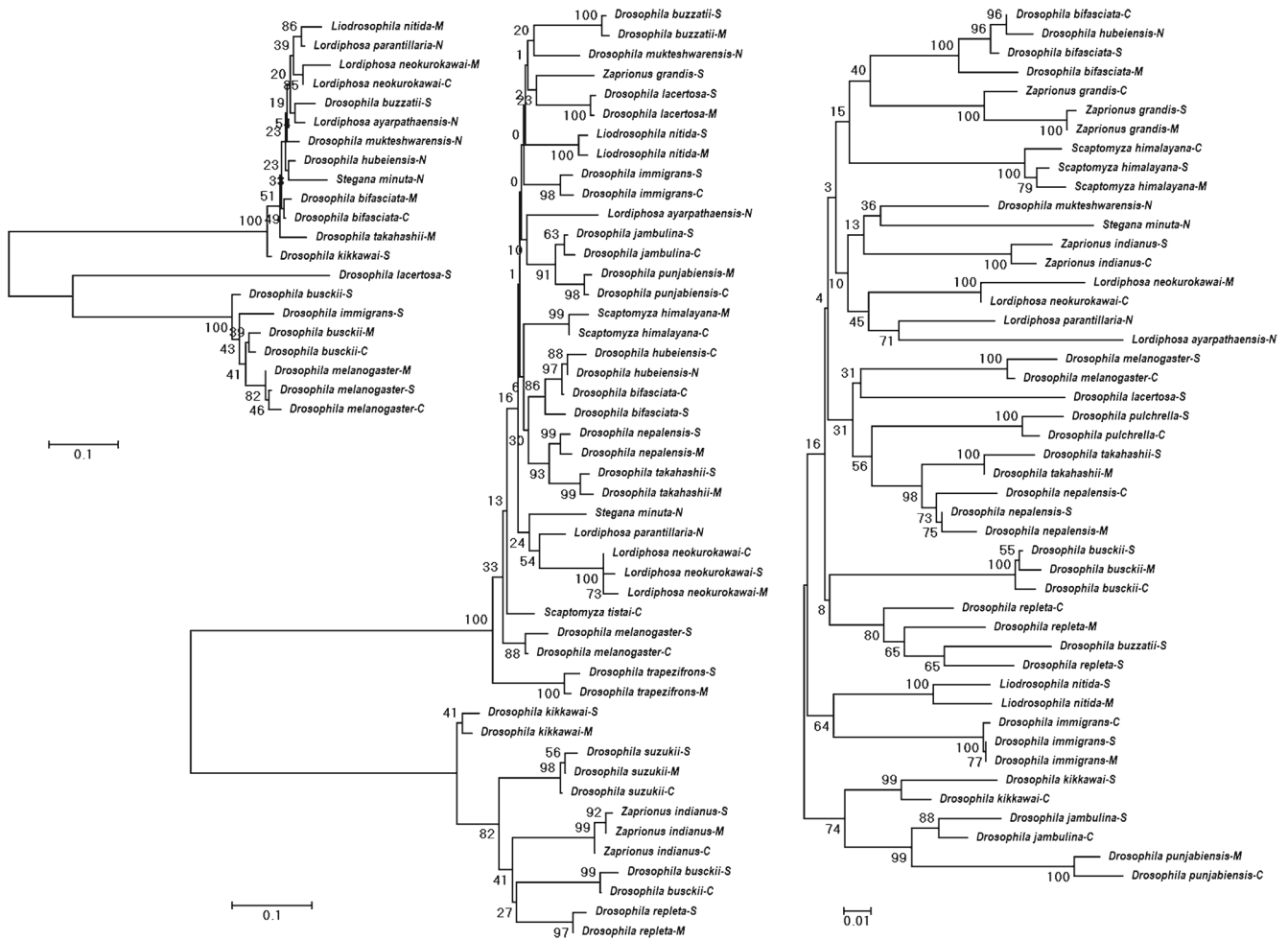
Figure 2. NJ tree based on partial mitochondrial (a) 16S rRNA, (b) *COI* and (c) *COII* genes (left to right). Statistical support for nodes on the tree was evaluated using 1000 bootstrapping replications.

Scaptomyza himalayana, *Scaptomyza tistai*, *Zaprionus grandis* and *Stegana minuta* are released for the first time to public domains through this study.

Discussion

The enormous biodiversity which we witness today is the result of several biological processes, including adaptive radiation (Gavrilets and Losos 2009; Glor 2010; Losos 2010), which is rapid diversification of lineage into different ecological niches. Owing to ecological opportunity in an unutilized environment, the ancestral species can undergo rapid species diversification into new ecological niches available. Central Himalayan region of India, due to its variable altitudinal terrain, encompasses highly varied tropical to temperate-like habitats in span of just few hundred kilometres and thus offers a well-established arena for studying diversification across lineages, which can be useful in, identifying adaptive radiations and understanding processes underlying origin of biological diversity (Losos 2010). Our extensive explorations in the region in past decades unravelled around 95 species (Sati *et al.* 2013), with significant number of them new to science, which could be attributed to the rapid divergence of the lineage as the descendants with time got adapted to range of available ecological resources.

Table 3(a). Neutrality tests based on 16S rRNA partial gene sequence.

Species	Tajima's D	Fu and Li's D* test statistic	Fu and Li's F* test statistic
D. bifasciata	2.63302**, $P < 0.01$	1.52051**, $P < 0.02$	2.02366**, $P < 0.02$
D. busckii	11.96703 (NS)	1.59120**, $P < 0.02$	1.94487**, $P < 0.02$
D. melanogaster	2.15778*, $P < 0.05$	1.43932*, $P < 0.05$	1.88394**, $P < 0.02$
Lordiphosa neokurokawai	2.21963**, $P < 0.01$	1.60451**, $P < 0.02$	1.91797**, $P < 0.02$

Table 3(b). Neutrality tests based on *COI* partial gene sequence.

Species	Tajima's D	Fu and Li's D* test statistic	Fu and Li's F* test statistic
D. busckii	2.36586**, $P < 0.01$	1.32645 (NS)	1.79091**, $P < 0.02$
D. immigrans	2.79440***, $P < 0.001$	1.51877**, $P < 0.02$	2.10533**, $P < 0.02$
D. buzzatii	2.54398**, $P < 0.01$	1.44217*, $P < 0.05$	1.94596**, $P < 0.02$
D. lacertosa	2.70757***, $P < 0.001$	1.55203**, $P < 0.02$	2.09289**, $P < 0.02$
D. repleta	2.68722***, $P < 0.001$	1.55671**, $P < 0.02$	2.07205**, $P < 0.02$
D. bifasciata	2.85913***, $P < 0.001$	1.54468**, $P < 0.02$	2.16434**, $P < 0.02$
D. hubeiensis	2.57153**, $P < 0.01$	1.46041*, $P < 0.05$	1.97038**, $P < 0.02$
D. jambulina	2.50001**, $P < 0.01$	1.56179**, $P < 0.02$	2.00367**, $P < 0.02$
D. kikkawai	2.69816***, $P < 0.001$	1.54561**, $P < 0.02$	2.08431**, $P < 0.02$
D. melanogaster	2.73944**, $P < 0.01$	1.46498*, $P < 0.05$	2.05539**, $P < 0.02$
D. nepalensis	2.67640***, $P < 0.001$	1.53081**, $P < 0.02$	2.06454**, $P < 0.02$
Lordiphosa neokurokawai	2.01136*, $P < 0.05$	1.53818**, $P < 0.02$	1.90983**, $P < 0.02$
D. punjabiensis	2.71617***, $P < 0.001$	1.55791**, $P < 0.02$	2.10074**, $P < 0.02$
D. suzukii	2.32365*, $P < 0.05$	1.53348**, $P < 0.02$	2.02111**, $P < 0.02$
D. takahashii	2.63369**, $P < 0.01$	1.50196**, $P < 0.02$	2.02596**, $P < 0.02$
D. trapezifrons	2.59517**, $P < 0.01$	1.47616**, $P < 0.02$	1.99144**, $P < 0.02$
Liodrosophila nitida	2.21111**, $P < 0.01$	1.59801**, $P < 0.02$	1.90937**, $P < 0.02$
Scaptomyza himalayana	2.77878***, $P < 0.001$	1.50848**, $P < 0.02$	2.09133**, $P < 0.02$
Zaprionus indianus	2.09075*, $P < 0.05$	1.47818*, $P < 0.05$	1.89719**, $P < 0.02$

Table 3(c). Neutrality tests based on *COII* partial gene sequence.

Species	Tajima's D	Fu and Li's D* test statistic	Fu and Li's F* test statistic
D. busckii	2.15050*, $P < 0.05$	1.57613**, $P < 0.02$	2.01046**, $P < 0.02$
D. immigrans	1.91971 (NS)	1.27929 (NS)	1.66036*, $P < 0.05$
D. repleta	1.56360 (NS)	1.64519**, $P < 0.02$	1.86361**, $P < 0.02$
D. bifasciata	2.50649**, $P < 0.01$	1.60379**, $P < 0.02$	2.14340**, $P < 0.02$
D. jambulina	2.65812***, $P < 0.001$	1.53723**, $P < 0.02$	2.04601**, $P < 0.02$
D. kikkawai	2.65812***, $P < 0.001$	1.53723**, $P < 0.02$	2.04601**, $P < 0.02$
D. melanogaster	2.74144***, $P < 0.001$	1.48403**, $P < 0.02$	2.05802**, $P < 0.02$
D. nepalensis	2.31465*, $P < 0.05$	1.57857**, $P < 0.02$	2.05860**, $P < 0.02$
Lordiphosa neokurokawai	2.50001**, $P < 0.01$	1.56179**, $P < 0.02$	2.00367**, $P < 0.02$
D. pulchrella	2.61813**, $P < 0.01$	1.51062**, $P < 0.02$	2.01045**, $P < 0.02$
D. punjabiensis	2.50312**, $P < 0.01$	1.43508*, $P < 0.05$	1.90955**, $P < 0.02$
D. takahashii	2.59517**, $P < 0.01$	1.47616**, $P < 0.02$	1.99144**, $P < 0.02$
Liodrosophila nitida	2.72140***, $P < 0.001$	1.57971**, $P < 0.02$	2.10281**, $P < 0.02$
Scaptomyza himalayana	2.06486*, $P < 0.05$	1.55483**, $P < 0.02$	1.94118**, $P < 0.02$
Zaprionus grandis	2.09517*, $P < 0.05$	1.56280**, $P < 0.02$	1.97367**, $P < 0.02$
Zaprionus indianus	2.64627***, $P < 0.001$	1.52932*, $P < 0.02$	2.03544**, $P < 0.02$

Thus, we assessed variation in three mitochondrial genes, namely, 16S rRNA, *COI* and *COII* in 26 drosophilid species collected along altitudinal gradient from 550 to 2700 m above mean sea level. In all 543 sequences, we generated: 82 for 16S rRNA, 238 for *COI*, 223 for *COII* with 21, 47 and 45 mitochondrial haplotypes for 16S rRNA, *COI* and *COII* genes, respectively.

Most species were represented by 2–3 unique mitochondrial haplotypes depicting a significant impact of environmental heterogeneity along altitudinal gradient on genetic diversity. Of the six localities surveyed, all individuals sequenced from Srinagar – Garhwal, Augustyamuni and Upper Chamoli showed same haplotype (designated 'S'), Kanchula – Kharak and Chopta formed another haplotype (designated 'C') in some cases, while Mandal formed a discrete haplotype (designated 'M') in most of the species.

Thus, Mandal (1600 m) being a transition area between subtropical and temperate zones with high degree of heterogeneity in climate and vegetation, has an evident impact on genetic diversity of drosophilid species. Large population size and environmental heterogeneity promotes high haplotype diversity within populations (Nei 1987; Avise 1998), while nucleotide diversity (π) specifies the mean number of differences among all pairs of haplotypes in population. Therefore, these values provide insightful inference of genetic diversity for a population (Nei and Li 1979). Almost all species exhibited high haplotype diversity (>0.5) with low nucleotide diversity (<0.02) indicating that the studied populations recently diverged from each other. Although all species represented low π values, they were within the boundaries of the values observed for mtDNA in other *Drosophila* species (Powell 1997). In this

study, we also observed high level of polymorphism as all the flies sequenced were natural populations caught from wild and not the laboratory cultures, leading to high levels of polymorphism.

We also obtained significantly positive results for neutrality tests in this study indicating balancing selection or population substructure. Most mitochondrial DNA sequence analysis have indicated that whenever the neutral model is rejected for a region it is due to a significantly negative test (Rand *et al.* 1994; Nachman 1998; Nielsen and Weinreich 1999). It may be due to recent directional selection, population bottleneck, recent population growth or background selection of some deleterious alleles (Tajima 1989). Although the ultimate consequences of these events may be an accumulation of rare polymorphic sites, the difference in the pattern of nucleotide substitution allows differentiation between some of these processes (Fu 1997; Wayne and Simonsen 1998). Our analysis of all the three genes of almost all the species pooled individually indicated significant departure from neutrality. Several different evolutionary forces may cause such departure. All these neutrality tests do not have the same power to detect departure from neutrality due to different evolutionary forces such as hitchhiking, population expansion, background or directional selection (Fu 1997) but their combinatorial application permits inferences on the patterns of selection affecting a specific/selected gene.

Most genetic diversity studies in *Drosophila* and several other taxa primarily utilize closely related species pairs (Coyne and Orr 1997, 1989) rather than population of same species at the early juncture of speciation (Harrison 1998). Although most interspecific studies provide information about the plausible correlation relationship between genetic diversity and the strength of an isolating mechanism after speciation, they do not provide a clearer picture of early stages of speciation, i.e. when particular isolating mechanisms primarily arose or its relationship to degrees of genetic diversity among genomes. Here, six altitudinally isolated populations of 26 drosophilid species, utilizing different ecological niches were analysed. We characterized evolutionarily diverging populations of these drosophilid species based on three mitochondrial genes with respect to both reproductive isolation and degree of genetic differentiation among them. This pattern is generally considered outcome of adaptive radiation as we witnessed rapid origin of species adapted to diversity of ecological niches followed by a slowdown of diversification rate through time as the new niches get occupied (Gavrilets and Losos 2009; Yoder *et al.* 2010).

With advent of plethora of molecular taxonomic techniques, recent sequencing of genomes of several *Drosophila* species and their expanding implication in evolutionary biology, it was important to present before the *Drosophila* research community a better taxonomic framework of the common and novel species reported by us from Uttarakhand region. Considering the high rate of discovery of novel drosophilid species from this region, we for the first time employed molecular characterization tools to speed up the recognition of new units that represent undescribed species. Molecular data of some rare species like *D. mukteshwarensis*, *Liodrosophila nitida*, *Lordiphosa parantillaria*, *Lordiphosa ayarpathaensis*, *Scaptomyza himalayana*, *Scaptomyza tistai*, *Zaprionus grandis* and *Stegana minuta* are now for the first time available to public domains through this effort which would not only strengthen the drosophilid taxonomic database of the region but may also help us to better understand the process of ecological adaptations that occurred over the history of this group.

References

Avise J. C. 1991 Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. *Annu. Rev. Genet.* **25**, 45–69.

Avise J. C. 1998 *Phylogeography*. Harvard University Press, Cambridge, USA.

Avise J. C. 2000 *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge, USA.

Avise J. C. 2004 *Molecular markers, natural history and evolution*, 2nd edition. Sinauer Associates, Sunderland, USA.

Avise J. C., Arnold J., Ball M. R., Bermingham E., Lamb T., Neigel J. E. *et al.* 1987 Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Ann. Rev. Ecol. Syst.* **18**, 489–522.

Bächli G. 2015 *TaxoDros: the database on taxonomy of Drosophilidae*, ver. 3.0. Available online at: http://www.taxodros.unizh.ch (accessed on 8 September 2015).

Bast F. 2015 Phylogenetics: tracing the evolutionary legacy of organisms, metastatic clones, bioactive compounds and languages. *J. Phylogen. Evol. Biol.* **3**, e112.

Brown W. M., George M. and Wilson A. C. 1979 Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* **76**, 1967–1971.

Coyne J. A. and Orr H. A. 1989 Patterns of speciation in *Drosophila*. *Evolution* **43**, 362–381.

Coyne J. A. and Orr H. A. 1997 Patterns of speciation in *Drosophila* revisited. *Evolution* **51**, 295–303.

Drosophila 12 Genomes Consortium 2007 Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* **350**, 203–218.

Emerson B. C., Paradis E. and Thëbaud C. 2001 Revealing the demographic histories of species using DNA sequences. *Trends. Ecol. Evol.* **16**, 707–716.

Fartyal R. S. and Singh B. K. 2002 Taxonomic accounts of new species of Drosophilidae (Insecta: Diptera) of Kumaon region, India. *Entomon* **27**, 355–364.

Fartyal R. S. and Singh B. K. 2004 A new species of genus Paraleucophenga (Diptera: Drosophilidae) from Kumaon region, India. *Senckenb. Biol.* **83**, 177–180.

Fartyal R. S., Singh B. K. and Toda M. J. 2005 Review of the genus *Leucophenga* Mik (Diptera: Drosophilidae) in India, with

descriptions of five new species from northern India. *Entomol. Sci.* **8**, 405–417.

Fartyal R. S., Gao J. J., Toda M. J., Hu Y. G., Takano K. T., Suwito A. *et al.* 2013 *Colocasiomyia* (Diptera: Drosophilidae) revised phylogenetically, with a new species group having peculiar life-cycles on monsteroid (Araceae) host plants. *Syst. Entomol.* **38**, 763–782.

Fu Y.-X. 1997 Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147**, 915–925.

Fu Y.-X. and Li W.-H. 1993 Statistical tests of neutrality of mutations. *Genetics* **133**, 693–709.

Gavrilets S. and Losos J. B. 2009 Adaptive radiation: contrasting theory with data. *Science* **323**, 732–737.

Glor R. E. 2010 Phylogenetic insights on adaptive radiation. *Annu. Rev. Ecol. Evol. Syst.* **41**, 251–270.

Gupta J. P. 2005 A monograph on Indian Drosophilidae. *J. Sci. Res.* **5**, 1–252.

Harrison R. G. 1989 Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends Ecol. Evol.* **4**, 6–11.

Harrison R. G. 1998 Linking evolutionary pattern and process: the relevance of species concepts for the study of speciation. In: *Endless forms: species and speciation* (ed. D. J. Howard and S. H. Berlocher), pp. 19–31. Oxford University Press, Oxford.

Hebert P. D. N., Cywinska A., Ball S. L. and DeWaard J. R. 2003a Biological identifications through DNA barcodes. *Proc. R. Soc. B* **270**, 313–321.

Hebert P. D. N., Ratnasingham S. and deWaard J. R. 2003b Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. B* **270**, 96–99.

Hu Y. G. and Toda M. J. 2001 Polyphyly of *Lordiphosa* and its relationships in Drosophilinae (Diptera: Drosophilidae). *Syst. Entomol.* **26**, 15–31.

Jinbo U., Kato T. and Ito M. 2011 Current progress in DNA barcoding and future implications for entomology. *Ento. Sci.* **14**, 107–124.

Kumar R. and Ajai K. 2009 Studies on Drosophilidae (Diptera: Drosophilidae) of Gujarat state in India. *Dros. Inf. Serv.* **92**, 106.

Larkin M. A., Blackshields G., Brown N. P., Chenna R., McGettigan, P. A., McWilliam H. *et al.* 2007 Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948.

Librado P. and Rozas J. 2009 DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**, 1451–1452.

Losos J. B. 2010 Adaptive radiation, ecological opportunity, and evolutionary determinism. American Society of Naturalists E. O. Wilson award address. *Am. Nat.* **175**, 623–639.

Markow T. and O'Grady P. M. 2006 *Drosophila: a guide to species identification and use*. London Academic Press, London, UK.

McAlpine J. F. 1981 Morphology and terminology—adults. In: *Manual of nearctic diptera, agriculture canada monograph no. 27* (eds. J. F. McAlpine, B. V. Peterson, G. E. Shewell, H. J. Teskey, J. R. Vockeroth and D. M. Wood), vol. **1**, pp. 9–63. Ontario Biosystematics Research Institute, Ottawa, Canada.

Moritz C., Dowling T. E. and Brown W. M. 1987 Evolution of animal mitochondrial DNA: relevance for population biology and systematics. *Ann. Rev. Ecol. Syst.* **18**, 269–292.

Nachman M. W. 1998 Deleterious mutations in animal mitochondrial DNA. *Genetica* **102/103**, 61–69.

Nei M. 1987 *Molecular evolutionary genetics*. Columbia University Press, New York, USA.

Nei M. and Li W. H. 1979 Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* **76**, 5269–5273.

Nielsen R. and Weinreich D. M. 1999 The age of nonsynonymous and synonymous mutations in animal mtDNA and implications for the mildly deleterious theory. *Genetics* **153**, 497–506.

Palumbi S. R. 1996 Nucleic acids II: the polymerase chain reaction. In: *Molecular systematics*, 2nd edition (ed. D. M. Hillis, C. Moritz and B. K. Mable), pp. 205–247. Sinauer, Sunderland, USA.

Posada D. and Crandall K. A. 2001 Intraspecific gene genealogies: trees grafting into networks. *Trends Ecol. Evol.* **16**, 37–45.

Powell J. R. 1997 *Progress and prospects in evolutionary biology: the drosophila model*, 1st edition. Oxford University Press, Oxford.

Rand D. M., Dorfsman M. and Kann L. M. 1994 Neutral and non-neutral evolution of drosophila mitochondrial DNA. *Genetics* **138**, 741–756.

Rubinoff D. and Holland B. S. 2005 Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Syst. Biol.* **54**, 952–961.

Sati P. C., Fartyal R. S., Sarswat M., Dewan S., Kandpal M., Kumar K. *et al.* 2013 Check-list of drosophilid species so far described and recorded from Uttarakhand state, India. *Dros. Inf. Serv.* **96**, 79–83.

Simon C., Frati F., Beckenbach A., Crespi B., Liu H. and Flook P. 1994 Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann. Entomol. Soc. Am.* **87**, 651–701.

Singh B. K., Bhatt M. and Negi N. S. 1985 Drosophildae of Uttarakhand. *Dros. Inf. Serv.*, 159–160.

Singh B. K. and Bhatt M. 1986 Drosophildae of Kumaon region. *Him. Res. Grind* 264–274.

Singh B. K. and Bhatt M. 1988 A preliminary report on the Drosophildae of Kumaon region with description of two new species and three new records. *Orient Insects* **22**, 147–161.

Singh B. K. and Negi N. S. 1989 Drosophilidae of Garhwal region with the description of one new species. *Proc. Zool. Soc. Calcutta* **40**, 19–26.

Singh B. K. and Negi N. S. 1992 Two new and one unrecoered species of Drosophilidae from Uttarakhand, India (Insecta: Diptera). *Senckenb. Biol.* **72**, 321–327.

Singh B. K. and Negi N. S. 1995 Further addition to the Drosophilid fauna of Uttarakhand, India (Diptera: Drosophilidae). *Entomol. Z.* **105**, 428–434.

Singh B. K. and Dash S. 1998 Drosophilidae of Kumaon region, India with the description of four new species (Insecta: Diptera). *Proc. Zool. Soc. Calcutta* **51**, 45–56.

Singh B. K. and Fartyal R. S. 2002 Family Drosophilidae (Insecta: -Diptera) in Kumaon region. India with the description of one new species and three new records. *Proc. Zool. Soc. Calcutta* **55**, 11–18.

Singh B. K., Dash S. and Fartyal R. S. 2000 Further additions to the genus *Leucophenga* from Kumaon, India (Insecta: Diptera: Drosophilidae). *Senckenb. Biol.* **80**, 149–154.

Singh B. K., Dash S. and Fartyal R. S. 2004 Revision of the species of the subgenus *Drosophila* (Drosophila) of the Kumaun region, India with the description of eight new species (Insecta, Diptera, Drosophilidae). *Senckenb. Biol.* **83**, 163–176.

Tajima F. 1989 Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.

Tamura K., Stecher G., Peterson D., Filipski A. and Kumar S. 2013 MEGA6: molecular evolutionary genetics analysis, version 6.0. *Mol. Biol. Evol.* **30**, 2725–2729.

Toda M. J. and Kitching R. L. 2002 Forest ecosystems. In: *Biodiversity research methods, IBOY in western Pacific and Asia* (ed. T. Nakashizuka and N. Storks), pp. 216, Kyoto University Press, Kyoto, Japan.

Wayne M. L. and Simonsen K. L. 1998 Statistical tests of neutrality in the age of weak selection. *Trends Ecol. Evol.* **13**, 236–240.

Wheeler M. R. 1981 The Drosophilidae: a taxonomic overview. In: *The genetics and biology of Drosophila* (ed. M. Ashburner, H. L. Carson and J. N. Thompson Jr), **3a**, 1–97, London Academic press, London, UK.

Yassin A., Markow T. A., Narechania A., O'Grady P. M. and DeSalle R. 2010 The genus *Drosophila* as a model for testing tree- and character-based methods of species identification using DNA barcoding. *Mol. Phylogenet. Evol.* **57**, 509–517.

Yoder J. B., Clancey E., Des Roches S., Eastman J. M., Gentry L. *et al.* 2010 Ecological opportunity and the origin of adaptive radiations. *J. Evol. Biol.* **23**, 581–596.

Zhang W. and Toda M. J. 1992 A new species-subgroup of the *Drosophila immigrans* species-group (Diptera, Drosophilidae), with description of two new species from China and revision of taxonomic terminology. *Jpn. J. Ento.* **60**, 839–850.