ISSN 1682-296X (Print)
ISSN 1682-2978 (Online)

Bio Technology

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Biotechnology 12 (5): 189-201, 2013
ISSN 1682-296X / DOI: 10.3923/biotech.2013.189.201

Assessment of Genetic Divergence with Self-Organizing Maps (SOM) in Silkworm, *Bombyx mori* L. (Lepidoptera: Bombycidae) Genotypes

Savarapu Sugnana Kumari, Sunil Misra, Mutheneni Srinivasa Rao and
Upadhyayula Suryanarayana Murty
Division of Biology, CSIR-Indian Institute of Chemical Technology, Tarnaka, Hyderabad,
500607, Andhra Pradesh, India

Abstract: Silkworm, *Bombyx mori*, is a very important economically exploited insect in South East Asia and revealing its genetic variation and diversity among genetic stock and appropriate classification method is needed for identification of potential parents. In the present study, the genetic divergence in 56 silkworm genotypes on 13 important qualitative and quantitative traits using Self Organizing Maps (SOM) was analyzed through cluster analysis. The highest level of genetic diversity were found in polyvoltine breeds by dividing into 8 different clusters and 6 clusters among bivoltine shown in two dimensional images called Kohonen maps. This clustering method allowed to divide the observations into several subclusters in such a way that homogeneity was obtained inside the sub-clusters and heterogeneity among the sub-clusters. Popular genotypes like Pure Mysore (PM), Nistari (NT-M, NT-P) were included in the cluster with 21.875% intra cluster difference among polyvoltine genotypes. Further, NB_4D_2-1 included in the cluster with 16.66% intra cluster differences along with other among bivoltine genotypes. It respectively indicates possibility of their exploitation in the field. There are numerous applications involving in the SOM algorithm but the most widespread employ was the identification and visualization of natural clustering in the data. Results indicated better ascertaining on the genetic diversity and genetic background on these silkworm genotypes provide ample scope to utilize them to achieve the desired objectives of highly heterotic silkworm hybrids to increase silk productivity in this region were discussed for the first time.

Key words: Genetic traits, quantitative, qualitative, polyvoltine, bivoltine, self-organizing maps, clusters

INTRODUCTION

Silkworm, *Bombyx mori* L (Lepidoptera: Bombycidae) is well domesticated Lepidopteron model insect since five thousand years and its genetic diversity is being commercially exploited for hybrid vigor compulsorily as in that of some other animal and plant systems (Nagaraju, 2002). Currently, silkworm germplasm are mainly maintained in Japan, China, India, Russia, South Korea, France etc and it constitute a great resource of silkworm breeding material at global scenario. It is known that about more than 4000 silkworm genotypes of genetically diverged are being preserved in the world (Thangavelu *et al*., 2003; Chauhan *et al*., 2003). As understanding on gene interaction and its phenotypic expression will be foremost in silkworm breeding (Ramesha *et al*., 2009, 2010).

It is difficult to select desirable parents from the germplasm stock solitary based on their performance (Tazima, 1964). Therefore, it is essential to classify them into definite clusters based on genomic divergence. Such a classification will enable in careful selection of potential parental genotype for the hybridization programs. Most of quantitative traits in silkworm are economically noteworthy, measured by various units of different characters to assess the overall performance of a particular breed and analyzed with assistance of biometrical tool. The multivariate biometrical methods employed to measure the genetic divergence in crop plant breeding lucratively (Ceccarelli *et al*., 1987; Roy and Lin, 2000; Arunachalam, 2005) and D^2 analysis (Mahalanobis, 1936) is being one of the methods mostly used in estimating the genetic divergence in silkworm (Kumaresan *et al*., 2007), however more innovative such efficient methods are needed to ascertain complexity of silkworm genotypes.

In the view of above, the Self-Organizing Map (SOM) is a fairly well-known neural network and indeed one of the most popular unsupervised learning algorithms in recent bioinformatics analysis of biological data. Thus

Corresponding Author: Upadhyayula Suryanarayana Murty, Division of Biology, CSIR-Indian Institute of Chemical Technology, Tarnaka, Hyderabad, 500607, Andhra Pradesh, India Tel: +91-040-27193134 Fax: +91-040-27193227

maps may be effectively utilized for genetic analysis (Kohonen, 2001; Behbahani and Nasrabadi, 2009).

In this study, an attempt was made to analyze the SOM or Kohonen maps for clustering of silkworm genotypes to ascertain the magnitude of genetic diversity among 56 silkworm genotypes. It was achieved via to estimate the genetic variation on 13 qualitative and quantitative genetic traits. This is to augment the utilization of silkworm genetic resources for highly heterotic silkworm hybrids to increase silk productivity in South East Asian region are discussed.

MATERIALS AND METHODS

Silkworm genotypes: Fifty-six included 32 polyvoltine (GFP-C, PAF-G, PALF-G, CFP, CLFP, GLPF, CFP-CR, CFPG, GFM, PAF-BG, GFP-D, GFM-L, PALF-BG, CFM, CFP-BG, CFM-L, CLPF-MS, GDFP, CDFP, GFP, KNT, NT-M, NT-P, PM, IIA, GYP, CYP , GYM, GYM-L, CCM-D, CCM, GCM) and 24 bivoltine (BO_1N, SOF-Br, BO_3-W, SOW, SOC-B, SB-F, BO_3-BL, BO_1-BA, BO_2, SOHW, BO_1-Br, BD_1O, BO_1-S, BO_3, BD_2, BD_3N, BD_2-S, BD_1C-Br, BD_2-G, BD_1LC, BD_2LC, BD_3, BD_4 and NB_4D_2) silkworm genotypes of geographical races and evolved breeds were used for the present study.

Silkworm rearing: The standard rearing techniques and recommended methodology for maintenance of germplasm was followed (Krishnaswami, 1978). Three replications for each season for each genotype with 300 larvae retained after third moult were considered for the study. The rearing for maintenance of silkworm germplasm was followed as 'composite population' to avoid inbreeding depression and genetic erosion. Composite laying is defined as collection of known number of eggs from a known number of individual laying sources that represents the whole population. The composite layings were prepared only after the body pigmentation takes place by taking approximately 50 eggs from each laying of 20 disease free layings in each genotype. All the pieces from 20 laying sources were pasted on a slightly thick brown paper with paste and wrapped in white fine tissue paper after drying. Thus, each composite laying consists of 20 laying sources with about 1,000-1,500 individual eggs. After brushing, the young larvae (1st-3rd instars) were reared at 26-28°C with 80 - 90% relative humidity and the late age larvae (4th-5th instar) were maintained at 24-26°C with a relative humidity of 70-80%. At the end of 5th stage, the spinning or matured larvae were collected manually and mounted on plastic collapsible mountages. After 6 or 7th day, cocoon harvesting was done manually, deflossed, cocoon sorting and cocoon selection was done for preparation of disease free layings for next generation. Total and average filament length for single cocoon was obtained by reeling ten-selected cocoons with the mono cocoon reeling unit called 'Approvate'.

Data analysis by self-organizing maps or kohonen maps: Thirteen qualitative and quantitative traits viz., fecundity (No.), hatching (%), pupation rate (%), shell weight (g), shell ratio (%), denier (d), cocoon weight (g), weight of 10 larvae (g), Effective Rate of Rearing (ERR) by No and weight (g), cocoon weight (g), filament length (m), 5th stage larval duration (h), total larval duration (h) were considered to ascertain the genetic diversity through SOM or Kohonen maps to estimate the genetic variability among the traits were used. The data were generated from the four seasons viz., January-March, April-June, July-September and October-December from 2010 to 2012 and used for application of SOM on computerized software.

Self organizing map is one of the datamining techniques widely used for clustering huge dataset in a visualization pattern (Kohonen, 1995; Vargas *et al*., 2002). It reduces the dimensions of whole data through the use of self-organizing neural networks. The neurons are organized in lattice, either with one or two dimensional array and are associated with weights of the same dimension of the input data. Initially, the data is trained in a systematic pattern and make different units. During the learning stage, the weights of each unit start sharing their values with the other input data and finally occupied the particular position in the lattice. After the learning stage is over, the each unit of dataset will be associated with nearest units in terms of their likelihood association of data. In this process, a map is created in two dimensional manners with different intensity of color depending upon contribution of data of each unit (Kohonen, 2002). The detail steps involved in the algorithm were demonstrated by utilizing of Self-Organizing Maps on the studies of malaria and filariasis endemic zones (Murty *et al*., 2008, 2011).

Data normalization: Summarized data are normalized linearly in such a way that minimum value in each category is 0 and the maximum is 1. This is done to ensure that all the parameters are given equal importance when clustering is done. The neuron weightage was adjusted by the learning rate. The learning rates and distance threshold values for the SOM are generally default values. Unsupervised learning was done using the data learning constant of 0.01 with 5000 iterations that yielded clusters based on the neighborhood distance.

RESULTS

Morphological features of silkworm genotype: The morphological differences were found between each

Table 1: Mean values of genetic traits on fecundity, hatching, pupation rate, shell weight, shell ratio and denier for polyvoltine genotypes

Breeds	Fecundity (No.)	Hatching (%)	Pupation rate (%)	Shell weight (g)	Shell ratio (%)	Denier (d)
KNT	382	96.41	90.20	0.134	13.42	2.0
IIA	471	96.32	91.59	0.190	15.41	2.2
GFP-C	506	97.22	90.34	0.197	16.35	1.6
PAF-G	427	95.64	90.47	0.185	15.58	1.8
NT-M	387	97.75	89.11	0.144	13.66	2.0
NT-P	379	96.69	89.32	0.139	13.36	2.0
PALF-G	471	97.64	89.55	0.174	14.81	1.8
CFP	429	96.19	90.28	0.204	16.61	1.9
CLPF	497	97.71	91.33	0.205	17.12	1.9
GLPF	494	97.36	90.45	0.224	16.56	1.7
CFP-CR	491	97.28	90.54	0.190	16.29	1.9
CFP-G	455	96.96	91.65	0.207	16.31	1.7
GYP	434	96.91	91.79	0.161	15.38	2.0
CYP	474	97.52	91.49	0.176	15.07	2.1
CCM-D	486	96.95	89.35	0.196	16.10	2.0
CCM	509	97.22	90.25	0.179	14.99	1.9
GCM	405	97.22	90.63	0.170	15.69	1.9
GFM	384	98.01	89.82	0.161	15.34	1.7
PAF-BG	427	96.96	91.79	0.187	15.53	1.9
GFP-D	480	97.35	90.39	0.176	14.94	1.6
GYM	354	97.23	90.52	0.159	14.79	2.0
GFM-L	423	96.17	91.65	0.160	15.17	1.8
PALF-BG	427	97.39	90.48	0.193	15.41	2.0
GYM-L	388	95.75	90.16	0.162	15.26	2.1
CFM	434	97.39	90.73	0.193	16.17	1.8
CFP-BG	462	97.66	91.65	0.213	17.24	1.8
CFM-L	463	97.21	91.09	0.191	16.06	1.8
CLPF-MS	497	96.90	90.93	0.218	16.97	1.9
GDPF	500	96.92	91.62	0.211	17.19	1.9
CDFP	464	97.69	91.78	0.216	16.37	1.9
GFP	473	97.64	91.49	0.193	15.61	1.6
PM	437	97.25	91.23	0.150	14.20	1.9
Mean	447.14060***	7.07700***	90.74010***	0.1830***	15.59000***	1.88570***
SE	2.70610	0.06020	0.06660	0.0015	0.06720	0.00940
CV	9.68330	0.99200	1.17390	13.2600	6.89270	7.99940
F Ratio	248.30000	4.15000	8.69000	160.6200	55.62000	562.77000
CD	7.65031	0.79755	0.75418	0.0053	0.38315	0.67004

***Significant at $p \leq 0.0001$

genotype with respect to origin, egg, larva and cocoon traits. The all 56 genotypes were from different geographical origin and evolved. All 32 polyvoltine genotype were non-hibernating and 24 bivoltine were of hibernating nature. The larval body build of polyvoltine breeds were slender, whereas bivoltine breeds were stout. Cocoon shape in polyvoltine breeds was oval, 22 genotype spun florescent green colour cocoon (GFP-C, PAF-G, PALF-G, CFP, CLFP, GLPF, CFP-CR, CFPG, GFM, PAF-BG, GFP-D, GFM-L, PALF-BG, CFM, CFP-BG, CFM-L, CLPF-MS, GDFP, CDFP, GFP, PM, IIA), 3 with golden yellow colour (KNT, NT-M, NT-P), 4 with yellow (GYP, CYP , GYM, GYM-L) and 3 with cream colour (CCM-D, CCM, GCM), whereas bivoltine spun 14 oval and 10 peanut shape with white colour. Amongst germplasm, the popular commercial polyvoltine genotypes were PM, NT-M, NT-P and bivoltine NB_4D_2 in India.

Performance of genetic traits: The collected data amongst polyvoltine genotypes on fecundity was ranged from 354 (GYM) to 509 (CCM) and hatching varied between 95.64% (PAF-G) to 98.01% (GFM). Maximum pupation rate in PAF-BG (91.79%) and minimum of 89.11% was observed in NT-M. High shell weight in GLPF (0.224 g) and low in KNT (0.134 g), more shell ratio in CFP-BG (17.24%) and less in NT-P (13.36%) and IIA was found to exhibit thick denier of 2.2 d and thin denier in GFP-D with 1.6 d (Table 1). Ten larval weight ranged between GFP-C (33.54 g) to NT-M (25.35 g) with an average of 30.21 g. A prominently varied in ERR by number from GLPF (9358) to PM (8237) and by weight was observed between GLPF (12.46 kg) to PM (8.54 kg) with an average of 10.73 g. The high cocoon weight in GLPF (1.353 g) and low in KNT (0.995 g), shorter filament length in NT-M with 357 m and longer filament length in GDPF with 704 m, shorter total larval duration was in PALF-G (521 h) and PM (610 h) with longer larval duration was revealed (Table 2).

Amongst bivoltine genotypes fecundity was ranged from 523 (BD_4) to 619 (BO_2) and hatching varied between 98.23% (SOCB) to 94.58% (BD_4). Maximum pupation rate in BD_1O (92.66%) and minimum of 89.84% was observed

Table 2: Average values of genetic traits on larval weight, ERR, cocoon weight, filament length and larval duration for polyvoltine genotypes

Breeds	Wt of 10 larvae (g)	ERR		Cocoon weight (g)	Filament length (m)	Larval duration (h)	
		No	wt (kg)			5th stage	Total age
KNT	26.01	9245	9.15	0.995	363	127	531
IIA	31.04	9339	11.42	1.232	604	136	560
GFP-C	33.54	9241	11.09	1.207	615	137	527
PAF-G	30.71	9235	10.90	1.190	585	137	527
NT-M	25.35	9156	9.57	1.052	357	127	530
NT-P	27.06	9152	9.47	1.040	368	127	530
PALF-G	27.54	9176	10.62	1.174	584	132	521
CFP	31.40	9226	11.28	1.228	696	135	525
CLPF	30.32	9316	11.07	1.196	642	137	525
GLPF	31.26	9358	12.46	1.353	704	136	525
CFP-CR	30.20	9248	10.73	1.164	614	137	526
CFP-G	32.18	9332	11.66	1.271	621	137	527
GYP	27.29	9336	9.86	1.044	533	135	524
CYP	32.59	9354	10.77	1.169	629	136	527
CCM-D	29.26	9121	11.03	1.215	601	135	524
CCM	31.57	9222	10.95	1.193	586	137	527
GCM	30.20	9244	9.95	1.083	588	138	524
GFM	30.06	9158	9.55	1.049	582	137	524
PAF-BG	29.52	9323	11.19	1.207	569	138	527
GFP-D	32.37	9216	10.79	1.176	600	140	547
GYM	30.15	9226	9.85	1.074	585	139	524
GFM-L	31.18	9294	9.73	1.053	572	137	524
PALF-BG	27.87	9257	11.52	1.250	611	137	522
GYM-L	28.82	9213	9.75	1.064	584	138	526
CFM	30.00	9249	11.00	1.195	638	138	526
CFP-BG	31.66	9348	11.52	1.238	667	137	526
CFM-L	29.62	9308	11.02	1.189	624	138	526
CLPF-MS	32.90	9289	11.72	1.282	656	140	546
GDPF	33.45	9356	11.43	1.227	704	138	530
CDFP	33.38	9343	12.18	1.319	698	137	526
GFP	32.71	9354	11.47	1.236	629	138	528
PM	25.84	8237	8.54	1.053	382	162	610
Mean	30.21***	9248***	10.72***	1.1693***	587.***	136.00***	531.00***
SE	0.868	12.0893	0.8154	0.0057	5.82030	0.36160	1.01810
CV	4.6874	2.0949	2.5782	7.8229	15.85880	4.22960	3.06720
F Ratio	1.23	432.15	68.65	253.87	807.66000	56.31000	2493.42000
CD	13.4992	26.08045	0.57126	0.016	9.21955	0.05343	0.91776

***Significant at $p \leq 0.0001$

in BD_3N. High shell weight in BO_1N (0.343 g) and low in SOF-BR (0.242 g), more shell ratio in BO_1N (21.07%) and less in SOF-BR (17.88%) and BD_1LC was found to exhibit thick denier of 3.2 day and thin denier in BO_1BR with 2.0 day (Table 3).Ten larval weight ranged between BD_2-S (42.65 g) to SOF-BR (36.33 g) with an average of 40.50 g. A prominent varied in ERR by number from BO_3BL (9351) to SOF-BR (9179) and by weight was observed between NB_4D_2-I (15.40 kg) to SOW (12.19 kg) with an average of 13.36 kg. The high cocoon weight in SOW (1.353 g) and low in KNT (0.995g), shorter filament length in BO_3-W with 690 m and longer filament length in BO_1N with 1021 m, shorter total larval duration was in BD_2-S (529 h) and SB-F (554 h) with longer larval duration was revealed (Table 4).

Analysis of variance indicates highly significant ($p \leq 0.001$) variation among the silkworm genotypes as well as their interaction with different seasons for all the qualitative and quantitative traits considered for the study. However, from the above results it is very difficult to characterize all the races in terms of their both qualitative and quantitative characteristics. To find out potential races from these germplasm, all these data were subjected for cluster analysis based on percentage contribution using SOM. On the basis of SOM the 56 silkworm genotypes were classified into 14 clusters with substantial inter and intra cluster difference. A classification of data using the SOM is shown in Fig. 1 and 2. The visualized clusters of the SOM clearly distinguished based on the colour intensity such as maximum (red), average (maroon) and minimum (black) for each parameter. Whereas, the qualitative characters like denier and larval duration are considered in reciprocal manner i.e., less denier (fine denier) and less larval duration are preferred in selecting a breed for further development. By observing the percentage of intensity of the colouration of each cluster, one can know the performance of each parameter of a particular breed

Table 3: Mean values of genetic traits viz., fecundity, hatching, pupation rate, shell weight, shell ratio and denier on bivoltine genotypes

Breeds	Fecundity (No.)	Hatching (%)	Pupation rate (%)	Shell weight (g)	Shell ratio (%)	Denier (d)
BD_2	571	97.10	92.64	0.339	20.41	2.5
BO_1N	576	97.22	91.70	0.343	21.07	2.3
BD_3N	542	95.86	89.84	0.307	20.04	2.3
SOF-BR	570	95.40	89.32	0.242	17.88	2.4
BD_2-S	552	96.65	92.31	0.329	20.08	2.6
BD_1C-BR	557	97.34	90.50	0.286	19.74	3.2
BO_3-W	602	96.08	91.01	0.264	18.91	2.1
SOW	596	96.85	90.75	0.248	18.69	2.3
SOCB	590	98.23	89.42	0.258	19.31	2.1
SB-F	588	97.29	90.66	0.250	18.57	2.4
BD_2-G	531	97.46	90.81	0.278	20.20	2.4
BO_3BL	611	96.27	91.12	0.257	18.98	2.3
BO_1BL	551	96.96	91.22	0.280	19.42	2.3
BO_2	619	97.50	90.63	0.285	20.49	2.4
SOHW	578	95.59	89.69	0.260	18.97	2.3
BO_1BR	554	95.98	90.36	0.260	19.49	2.0
BD_1LC	549	97.92	90.38	0.313	20.18	3.2
BD_1O	570	96.20	92.66	0.308	20.00	3.1
BD_2-LC	554	97.37	90.94	0.309	20.76	2.4
BD_3	559	95.81	89.77	0.281	20.14	2.2
BO_1-S	618	96.96	90.86	0.289	20.41	2.1
BD_4	523	94.58	91.99	0.277	19.46	2.3
BO_3	616	96.51	91.49	0.265	19.23	2.1
NB_4D_2 - I	529	96.43	90.13	0.311	18.29	2.5
Mean	570.99***	96.6***	90.8***	0.285***	19.6132***	2.4***
SE	2.3495	0.0975	0.0816	0.0022	0.0695	0.0249
CV	5.7434	1.3982	1.2493	10.6825	4.9122	14.3338
F Ratio	96	5.88	2.87	61.33	16.21	163.6
CD	9.11008	0.98875	0.99343	0.01042	0.56588	0.0751

***Significant at $p \leq 0.0001$

Table 4: Mean values of genetic traits viz., larval weight, ERR, cocoon weight, filament length and larval duration on bivoltine genotypes

		ERR				Larval duration (h)	
Breeds	Wt of 10 larvae (g)	No	wt (kg)	Cocoon weight (g)	Filament length (m)	5th stage	Total age
BD_2	42.0	9308	15.35	1.664	839	135	530
BO_1N	40.6	9310	15.07	1.630	1021	139	533
BD_3N	40.6	9236	14.05	1.533	808	141	534
SOF-BR	36.3	9179	12.36	1.355	720	155	551
BD_2-S	42.6	9335	15.20	1.640	857	134	529
BD_1C-BR	41.3	9257	13.31	1.446	748	157	553
BO_3-W	41.2	9331	12.80	1.394	690	140	533
SOW	37.8	9250	12.19	1.325	728	153	549
SOCB	38.8	9207	12.22	1.336	862	152	548
SB-F	38.0	9268	12.38	1.344	768	155	554
BD_2-G	41.7	9286	12.69	1.375	815	136	532
BO_3BL	40.9	9351	12.54	1.351	720	153	549
BO_1BL	41.2	9305	13.46	1.440	765	139	533
BO_2	41.9	9273	12.73	1.392	799	142	534
SOHW	38.2	9211	12.55	1.370	744	137	532
BO_1BR	39.7	9277	12.34	1.333	835	142	536
BD_1LC	40.9	9261	14.26	1.551	792	140	533
BD_1O	41.2	9311	14.25	1.540	762	159	553
BD_2-LC	42.4	9289	13.63	1.488	827	139	533
BD_3	41.1	9191	12.74	1.395	840	142	536
BO_1-S	41.4	9276	13.03	1.414	872	139	534
BD_4	40.2	9303	13.18	1.425	775	153	549
BO_3	41.0	9286	12.84	1.390	760	137	533
NB_4D_2-I	40.9	9240	15.40	1.703	730	156	551
Mean	40.49***	9272***	13.35***	1.4514***	794***	144***	539***
SE	0.1248	5.5485	0.0789	0.0085	5.7699	0.5985	0.6479
CV	4.2688	0.8291	8.1826	8.0933	10.059	5.733	1.6637
F Ratio	31.91	4.09	68.62	95.24	21.02	273.09	167.75
CD	0.78573	62.5135	0.35564	0.0329	42.62619	1.40559	1.92384

***Significant at $p \leq 0.0001$

(Fig. 1-4). This SOM method allowed the study to divide observations into several sub clusters in such a way that homogeneity was obtained inside the sub-cluster and heterogeneity among the sub-cluster.

Table 5: Clustering of polyvoltine silkworm germplasm breeds

Cluster	Silkworm breeds
1.1	CFP, CLPF, GLPF, CFP-G, CFP-BG, CLPS-MS, GDPF, CDFP, GFP
1.2	GFP-C, CFP-CR, CFM, CFM-L
1.3	PAF-G, PALF-G, CCM, GFP-D
2.2	CYP
2.3	GFM, GFM-L
3.1	IIA, CCM-D, PAF-BG, PALF-BG
3.2	GCM
3.3	KNT, NT-M, NT-P, GYP, GYM, GYM-L, PM

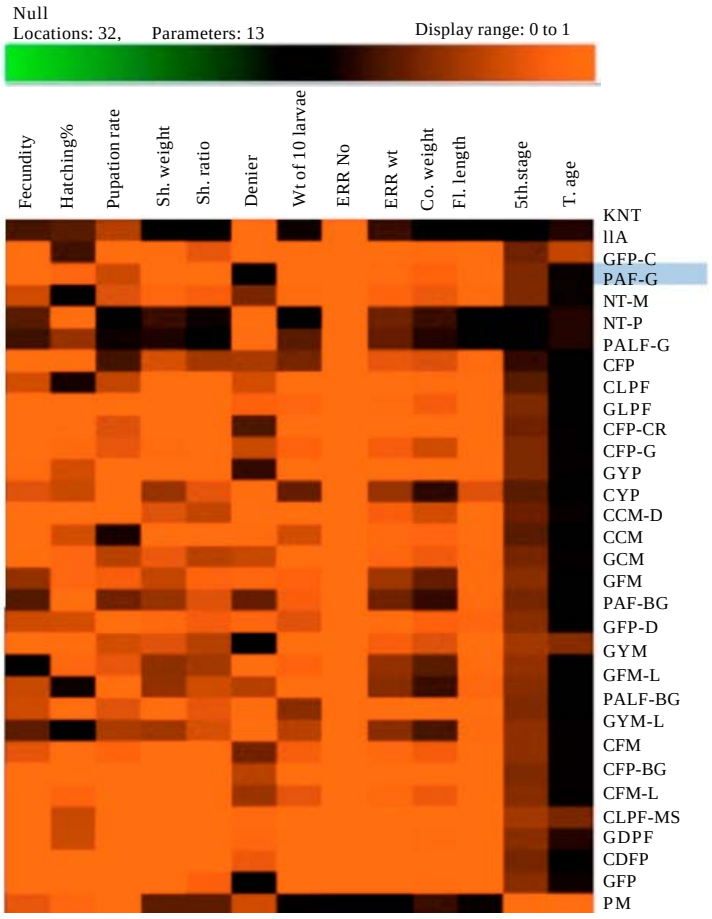

Fig. 1: SOM cluster analysis among 32 polyvoltine silkworm genotypes

Cluster analysis of SOM for polyvoltine breeds: Amongst 32 polyvoltine genotype subjected for SOM analysis resulted in 8 different clusters with substantial genotypic divergence (Table 5).

Cluster 1.1 includes 9 genotypes (CFP, CLPF, GLPF, CFP-G, CFP-BG, CLPS-MS, GDPF, CDFP and GFP). While considering the performance of each race it is found that, qualitative characteristics such as fecundity, hatching%, pupation, shell ratio, shell weight and denier (is in low range) were comparatively high among all the races under this particular cluster. These characters are very much important for considering a race to be a potential parental race. Among these nine genotypes, while considering the comparative qualitative characters analysis, except GFP all other breeds are with not much difference among these characters. However, GFP is occupied in this group may be due to similar type of hatching%, pupation, shell ratio, shell weight and denier while compared to other genotypes.

Cluster 1.2 comprises 4 genotypes (GFP-C, CFP-CR, CFM and CFM-L). It is exhibited to observe that all quantitative characters are less when compared to cluster 1.1. These genotypes included in this cluster due to similar type of hatching%, pupation

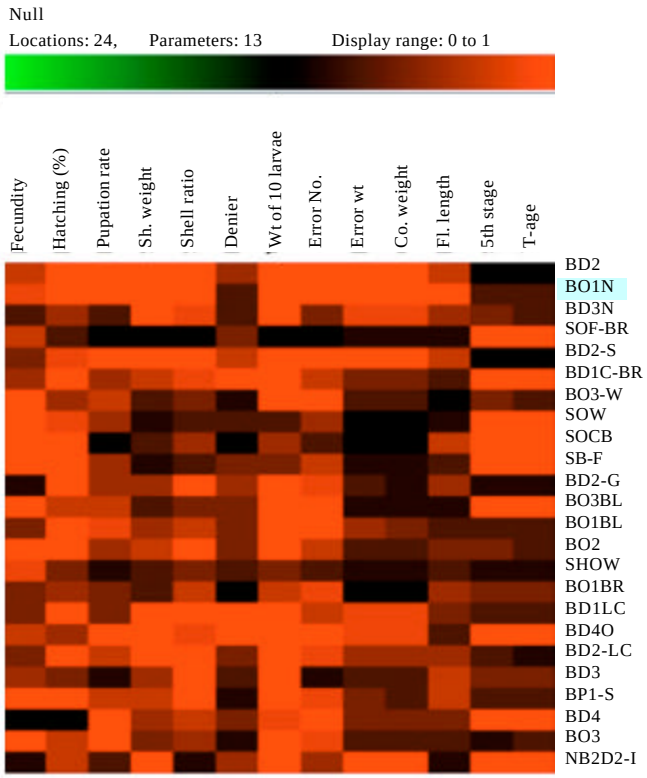

Fig. 2: SOM cluster analysis among 24 bivoltine silkworm genotypes

rate%, shell weight, shell ratio. But the values of fecundity and denier are not taken into consideration.

Cluster 1.3 associated with 4 genotypes (PAF-G, PALF-G, CCM and GFP-D). Here the quantitative performance of pupation rate, shell weight and shell ratio are performing similarly whereas, other characters are largely varied among these genotypes.

Cluster 2.2 accommodated only one genotype i.e., CYP. The performance values of this genotype is better in all qualitative characters except shell weight and shell ratio and among quantitative characters ERR by weight and cocoon weight are with less performance.

Cluster 2.3 included 2 genotypes (GFM and GFM-L). It was observed that all qualitative and quantitative characters are depleted when compared with the races performed in previous clusters except ERR by No and filament length.

Cluster 3.1 included 4 genotypes (IIA, CCM-D, PAF-BG and PALF-BG). It is observed that all the genotypes are performing similarly although these values are significantly showing lower performance when compared to the cluster 1.1 to 1.3. All the qualitative characters such as shell weight, shell ratio are very low and denier is in high.

Cluster 3.3 associated with 7 genotypes (KNT, NT-M, NT-P, GYP, GYM, GYM-L and PM). The genotypes of this cluster are with high hatching% which is a good quantitative characteristic of races however; other qualitative and quantitative characters are significantly lower when compared to races of cluster 1.1. The quality of silk is of very poor with high denier.

Maximum contribution for the genetic diversity in yield by number (ERR NO), followed by filament length (Fl. length), shell ratio (Sh. ratio), minimum in total larval duration (T. age) and high percentage (28.125%) intra cluster of (inter cluster) difference contributed in cluster 1.1, followed by cluster 3.3 (21.875%), clusters 1.2, 1.3, 3.1 (each cluster contributing 12.5%), cluster 2.3 (6.25%) and low percentage intra cluster (inter cluster) differences in clusters 2.2 and 3.2 (each 3.125%) was visualized in the process of SOM genetic divergence among the polyvoltine genotypes (Fig.1, 3 and 5a).

Cluster analysis of SOM for bivoltine breeds: Amongst 24 bivoltine genotype subjected for SOM analysis resulted in 6 different clusters with substantial genotypic divergence (Table 6).

Cluster 1.1 includes six genotypes (BD_2, BO_1N, BD_3N, BD_2S, BD_1LS and BD2-LC). In this cluster all the

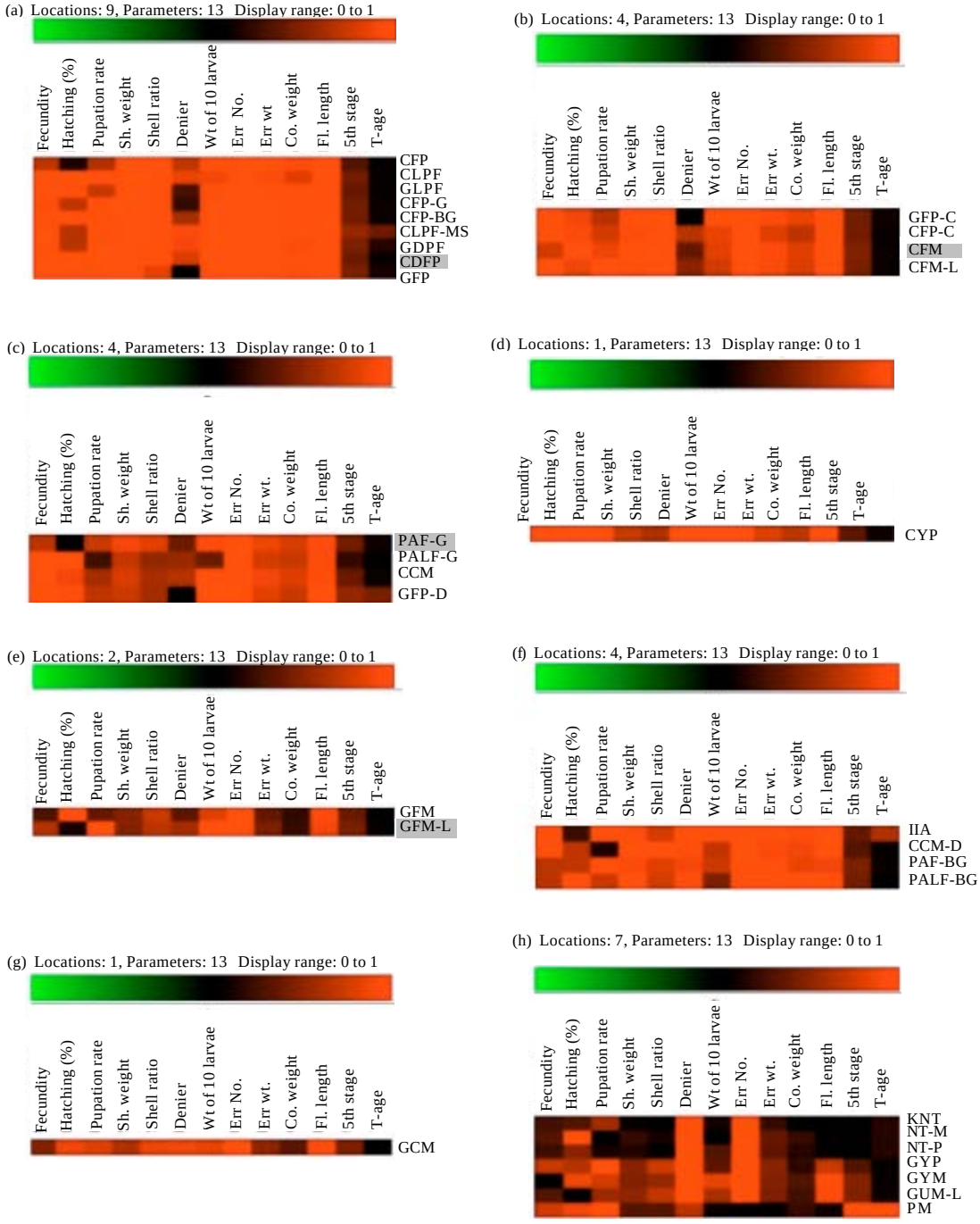

Fig. 3(a-h): Inter cluster divergences among polyvoltine silkworm genotypes (a) SOM cluster 1.1 (b) SOM cluster 1.2 (c) SOM cluster 1.3, (d) SOM cluster 2.2, (e) SOM cluster 2.3 (f) SOM cluster 3.1 (g) SOM cluster 3.2 and (h) SOM cluster 3.3

Table 6: Clustering of bivoltine silkworm germplasm breeds

Cluster	Silkworm breeds
1.1	BD_2, BO_1N, BD_3N, BD_2S, BD_1LS, BD_2-LC
2.1	BD_2-G, BO_1BL, BO_2, BO_1S
1.3	BD_1C-BR, BD_1O, BD_4, NB_4D_2-1
3.1	BO_3W, BD_3, BO_3
3.2	BO_1BR
3.3	SOF-BR, SOW, SOC-B, SB-F, BO_3BL, SOHW

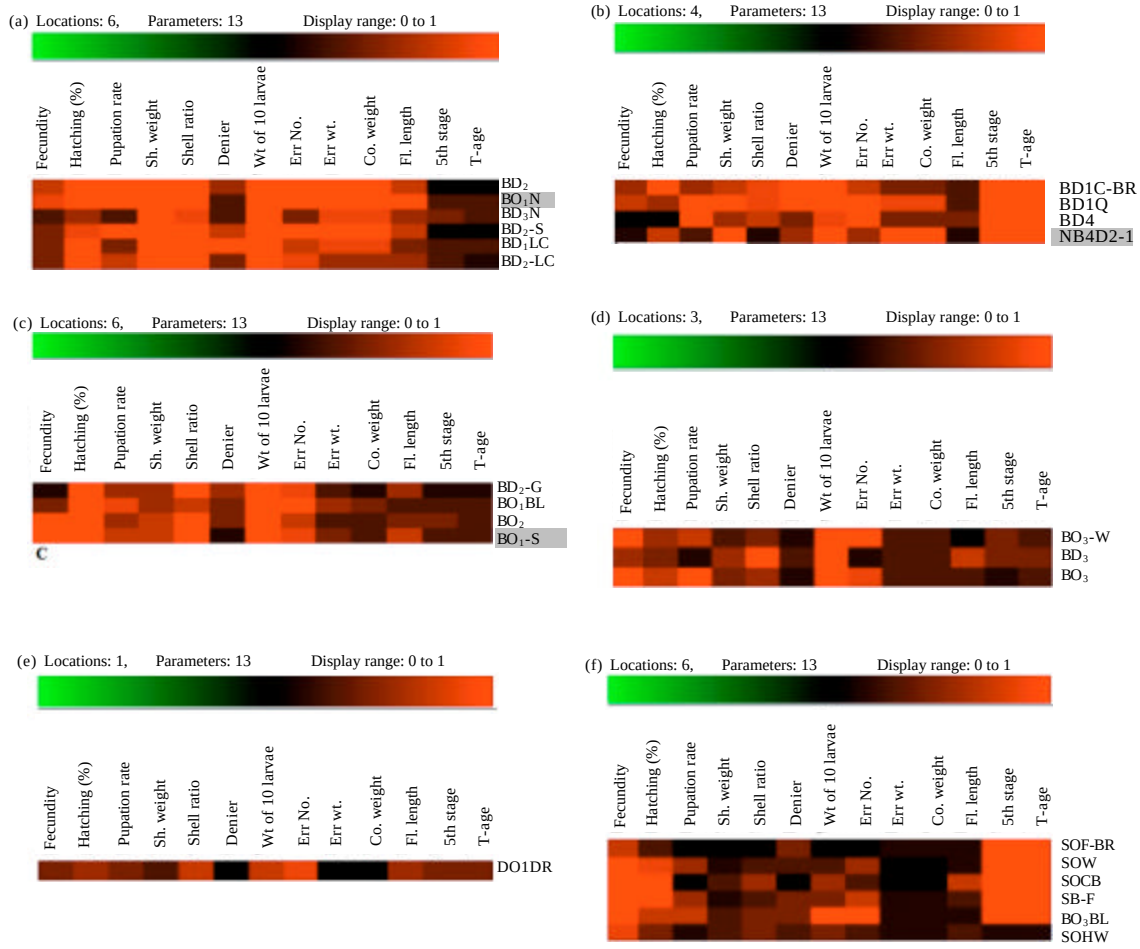

Fig. 4(a-f): Inter cluster divergences among bivoltine silkworm genotypes (a) SOM cluster 1.1 (b) SOM cluster 1.3 (c) SOM cluster 2.1 (d) SOM cluster 3.1 (e) SOM cluster 3.2 (f) SOM cluster 3.3

genotypes exhibited best performance in most of the qualitative and quantitative characters with high fecundity, hatching%, pupation rate%, shell weight, shell ratio, denier, cocoon weight and filament length. These breeds can be utilized for the improvement of specific desired character of a breed as well as can be utilized for overall improvement of races in breeding program.

Cluster 1.3 consisting of four genotypes BD_1C-BR, BD_1O, BD_4, NB_4D_2-1. These genotypes exhibit high hatching percentage, larval weight, shell ratio and filament length. However, rest of the characters is comparatively less than the races of the cluster 1.1. The reference NB_4D_2-1 showed high in most of qualitative and quantitative characters but grouped in this cluster may be due to high pupal weight, very less filament length and low shell ratio percentage similar to other breeds of the cluster.

Cluster 2.1 comprises four genotypes BD_2-G, BO_1BL, BO_2 and BO_1S. These races are performing better in shell weight, shell ratio, fine denier however with very less pupation rate and ERR. Hence these races can be utilized

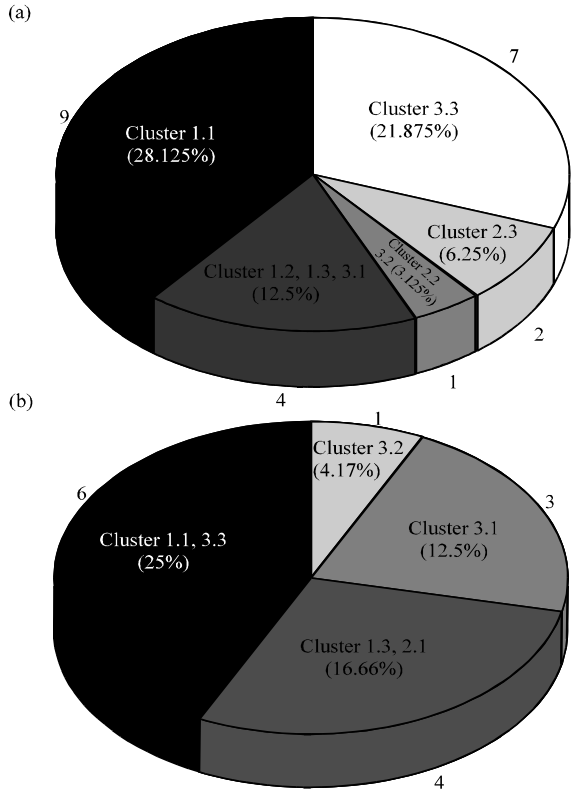

Fig. 5(a-b): SOM clusters details on percentage contribution of genetic traits in silkworm genotypes (a) Polyvoltine (b) Bivoltine

for induction of qualitative characters like shell weight and shell ratio for improving a particular race in breeding programme.

Cluster 3.1 comprises of three genotypes BO_3W, BD_3 and BO_3. These genotypes are promising with many important qualitative characters like higher fecundity, high hatching percentage, shell weight, shell ratio and fine denier. But quantitative characters are not much better than other breeds of cluster 1.1 and 2.1.

Cluster 3.2 consists only one genotype i.e., BO_1BR. This genotype differs from all the other genotypes due to higher hatching percentage, higher shell ratio (%) and moderate in all other qualitative and quantitative characters. Among all the characters the important of this breed is of with fine denier. Hence among all bivoltine breeds of the germplasm, this breed can be utilized in breeding program for induction of these important characters in new breeds or hybrid formation.

Cluster 3.3 comprised of six genotypes SOF-BR, SOW, SOC-B, SB-F, BO_3BL and SOHW. These races are with moderate fecundity, hatching percentage, shell weight, denier and uniform cocoon weight. But are with low pupation rate, low quality of silk reflected higher denier and low filament length and longer 5th stage and total larval duration.

More contribution for the genetic diversity in weight of 10 larvae (Wt of 10 larvae), followed by hatching%, yield by number (ERR NO), fecundity and less in filament length (Fl. Length) and high percentage (each 25%) intra cluster difference attributed in clusters 1.1, 3.3 followed by clusters 1.3, 2.1 (each of 16.66%), clusters 3.1 (12.5%) and low percentage intra cluster differences attributed in clusters 3.2 (4.17%) was envisaged in the process of SOM genetic divergence among the bivoltine genotypes (Fig. 2, 4 and 5b).

DISCUSSION

Since last few decades, the biometrical analysis concept for generalized distance D^2 analysis developed by Mahalanobis (1936) is one of the potent techniques extensively used in measuring genetic divergence between closely related parents in silkworm by multivariate analysis and varieties grouped according to Tocher's method and image processing methods by various silkworm breeders (Ramamohana Rao and Nakada, 1998; Kumaresan *et al*., 2007) as the success of breeding is mainly dependent on the selection of suitable parental resource from germplasm, experimental design, intensity of selection, mating system and major contribution traits to the overall cocoon yield. It was revealed that variability on gene interaction between commercial traits will be the source of breeding process (Frankel and Brown, 1984; Dalton, 1987). Analysis of variance indicated significant variation among the silkworm genotypes is the source for its genetic exploitation in sericulture. Genetic variance between genotypes may be calculated by various statistical measures based on obtained data on genetic traits (Mohammadi and Prasanna, 2003; Weir, 1996; Beaumont *et al*., 1998).

The genetic variability included phenotypic, genotypic and environmental variation will be the resource for the improvement in commercial hybrids. Thus, visualized clustering method of SOM also indicates the possibility for recombining low and high-yielding genotype from genetically distant clusters in hybrid selection as sericulture depends on crossbreed layings for cocoon production commercially, need of appropriate genotype in order to increase the production of hybrid eggs in the tropical countries like India (Ramesha *et al*., 2009). Reports from Arunachalam *et al*. (1984) emphasized that genetic diversity of the parents will certainly facilitate to heterotic hybrid combinations for commercial purpose.

The results of the present study revealed that the clusters 2.2 and 3.2 of polyvoltines genotypes has isolated with zero value of inter cluster distance, which has also shown minimum percentage contribution of 3.125 in SOM cluster analysis. Among bivoltine genotypes, an intra cluster distance of 4.17 was observed in cluster 3.2 which will be considered as minimum percentage contribution when compared to remaining clusters. But, the clusters 1.1 and 3.3 showed maximum percentage contribution of 25.00 and 28.125% intra cluster genetic distance in polyvoltine and bivoltine respectively clearly indicated the genetically diversification.

This study also revealed the inclusion of genotypes of the same origin in different clusters clearly indicate the presence of considerable genetic diversity among the population used in this study and results of analysis are compatible and complementary with the previous study based on D^2 analysis (Kumaresan *et al*., 2007; Nezhad *et al*., 2010). The overall genetic profile of polyvoltine genotypes, 8 clusters supports the popular genotypes of Pure Mysore (PM), NT-M, NT-P included in the same cluster with 21.875% percentage contribution of intra cluster differences indicates its longer adaptation with long history. In this study, a similar situation was also observed as a great diversity of polyvoltine genotype must exist globally as revealed by Anonymous (1997). This diversity is considered to be the result of adaptation during long generation (Murakami, 1994). It was proven that in case of PM genotype has one of the major local native genotype of Southern India, where sericulture is being practiced largely in a traditional manner. This was also used in the breeding for more than three decades as one of the female counterpart in the evolution of new polyvoltine as well as poly/bivoltine hybrids for commercial exploitation.

Similarly, popular NB_4D_2-1 included with 16.66% intra cluster differences along with other evolved genotypes viz., BD_1C-BR, BD_1O, BD_4 indicates possibility of exploitation among bivoltine genotype. Further, it could be indicated from the above results that the genotypes included in cluster 3.3 of polyvoltine and cluster 1.3 of bivoltine enhancements may be improved for the local region. Similarly, the optimum genetic distances obtained between clusters of polyvoltine and bivoltine genotype along with higher values on cocoon traits emphasized the utilization of these genotypes in the conventional silkworm breeding programs for sustainable silk production by exploitation of hybrid vigour in crossbreed in the South East Asian region particularly in India. The cluster analysis provides scope for adopting a recombination breeding program using distant cluster genotype.

The genotypes of polyvoltine of the cluster 1.1 belongs to high yielding varieties with many qualitative characters like good hatching percentage (96.19-97.19%), pupation rate (90.28-91.78%), high shell weight (0.204-0.224 g), high SR% (15.61-17.24%) and low denier (1.7-1.9). In qualitative characters most of the characters are almost similar nature. Hence, these genotypes are considered as better parents and can be utilized for improvement of desired characters in subsequent breeding programs. In contrast to cluster 1.1 the genotypes of cluster 1.2 belongs to breeds those are with uniform and good hatching percentage (97. 21-97.39%), shell weight (1.190-1.197 g), shell ratio (16.06-16.35%), low denier (1.6-1.9) and shorter total larval duration (526-527 h). Hence, these genotypes can be used for introduction of these characters in high yielding varieties in the breeding program. The races of cluster 1.3 are with very poor in both qualitative and quantitative characters however, there is higher percentage of hatching of eggs when compared to other races. Hence, these races can be utilized to improve the hatching percentage of eggs in development of a new breed. In cluster 2.2, only one race CYP is noticed which is with moderate in both qualitative and quantitative characteristics. Such type of races cannot be used for the breeding program. While analyzing these two races of cluster 2.3, almost all the characters are similar each other and can be considered as subline of one of the genotype. However, due to high hatching percentage and less total larval duration and lower denier of silk thread, these two races are in the same cluster. These breeds can be utilized for induction of character of denier as well as hatching percentage. In cluster 3.1, overall performance is said to be poor however, the most important characters of silk such as shell weight and filament length are very high compared to other races. All the genotypes of cluster 3.3 are considered as very poor quality races as, all the characters are significantly lower compared to other qualitative and quantitative races. Hence, these races cannot be utilized for any of the character improvement in a new breed and could not be part in the normal breeding program. Similarly, the breeds of bivoltine races are with many important characters in cluster 1.1 and considered to be the best breeds of the germplasm. However, other clusters were shown to have some specific characters and can be utilized in the breeding program for induction of specific characters in new breed or hybrids.

The genotypes of bivoltine cluster 1.1 can be chosen as better parental races with their better qualitative characters with maximum hatching percentage (96.65- 97.92%), high pupation rate (89.84-92.64%), high shell weight (0.307-0.307 g) and high shell ratio

(20.04-21.04%). In qualitative characters maximum ERR (15.35) was noticed in this cluster. Therefore, these genotypes are considered as better parents and can be utilized for improvement of desired characters in subsequent breeding programs. Breeds in cluster 1.3 are with good hatching percentage ranging from (94.58-97.34) larva weight ranging from (40.2-41.3 g) shell ratio ranging from (18.29-20.00) and shell weight ranging from (0.277-0.311 g) are grouped under this cluster. In this cluster good hatching percentage and pupation percentage was observed on par with previous cluster. In cluster 2.1 maximum fecundity was observed in two races i.e., 618 and 619. Good hatching percentage ranging from (96.96-97.50) was observed in the races belonging to this cluster and above 90% of pupation was observed in few races (90.63-91.22%). Comparatively less shell weight ranging from 0.278-0.289 g and Shell ratio ranging from 19.42-20.49% was observed in this cluster. In this cluster races with maximum fecundity of more than 600 eggs per each laying was observed. To increase character like fecundity this particular cluster races can be utilized in the breeding programme. In cluster 3.1, high fecundity was observed i.e., 602 and 616. Hatching percentage ranging from 95.81-96.51% was observed which can be considered as good qualitative characters to improve the breed. Pupation rate can be compared with races of previous clusters. Shell weight ranging from 0.264-0.281 g was observed. Shell ratio of 18.91-20.14% was observed in the cluster. When coming to denier races having fine quality silk are included under this particular cluster. Cluster 3.2 includes only one race with moderate hatching% i.e., 95.98%. Cluster 3.3 includes 6 breeds of which fecundity of minimum 570 and maximum 611 was observed in this clusters. Pupation and shell weight observed to be less compared to other cluster in most of the breeds when compared to cluster 1.1 and 1.3. Most of the breeds exhibited less shell ratio when compared to remaining clusters. Hence, qualitative traits when compared to other clusters were poor, thus the breeds of cluster 3.3 may not be recommended for utilization of breeding programme. Thus, SOM has the main advantage of visualized method over other analysis methods on the study of silkworm genetic diversity and map unit clustering makes it easy to observe similarities and dissimilarities on the data of silkworm genotypes for the first time efficiently.

CONCLUSION

This study provides insight into the genetic diversity among the 56 silkworm genotypes maintained at CSIR-Indian Institute of Chemical Technology, Hyderabad, India were distinctly differentiated into 14 clusters by SOM analysis. First, the overall genetic profile of polyvoltine genotypes with 8 clusters supports the popular genotypes of Pure Mysore (PM), Nistari (NT-M, NT-P) included in the same cluster with 21.875% intra cluster differences indicates its longer adaptation with long history. Second, bivoltine genotypes have established 6 clusters and popular NB_4D_2-1 included with 16.66% intra cluster differences along with other evolved genotypes viz., BD_1C-BR, BD_1O and BD4 indicates possibility of commercial exploitation. Third, these results indicated better ascertaining on the genetic diversity and genetic background on these silkworm genotypes for the first time by SOM as it is the most efficient visualized method. Finally, this study provides ample scope to utilize this genetic diversity to achieve the desired objectives of highly heterotic silkworm hybrids to increase silk productivity in South East Asian region.

ACKNOWLEDGMENTS

Authors wish to thank the Department of Science and Technology, Government of India, New Delhi for financial support in the form of Women Scientist Fellowship SR/ WOS-A/LS- 31/2004 to S. Sugnana Kumari. Authors are grateful to the Director, CSIR-Indian Institute of Chemical Technology, Hyderabad, India for his useful encouragement for this investigation.

REFERENCES

Anonymous, 1997. Principles and Techniques of Silkworm Breeding. ESCAP, Oxford and IBH Publishing Co. Pvt. Ltd., United Nations, New York, Pages: 114.

Arunachalam, V., A. Bandyopadhyay, S.N. Nigam and R.W. Gibbons, 1984. Heterosis in relation to genetic divergence and specific combining ability in groundnut (*Arachis hypogaea* L.) Euphytica, 33: 33-39.

Arunachalam, V., 2005. Phenotypic traits of genetic twins in coconut palm *Cocos nucifera* L. and their utility in seed production. Seed Technol., 27: 104-107.

Beaumont, M.A., K.M. Ibrahim, P. Boursot and M.W. Bruqord, 1998. Measuring Genetic Distance. In: Molecular Tools for Screening Biodiversity: Plants and Animals, Karp, A., P.G. Isaac and D.S. Ingram (Eds.). Chapman and Hall, London, UK., ISBN: 9780412638305, pp: 315-325.

Behbahani, S. and A.M. Nasrabadi, 2009. Application of SOM neural network in clustering. J. Biomed. Sci. Eng., 2: 637-643.

Ceccarelli, S., S. Grando and J.A.G. Van Leur, 1987. Genetic diversity in barley landraces from Syria and Jordan. Euphytica, 36: 389-405.

Chauhan, K.K., P.K. Rout, P.K. Singh, A. Mandal, S.K. Singh and R. Roy, 2003. Genetic resistance of Barbari and Jamunapari kids to natural infection with gastrointestinal nematodes. Trop. Anim. Health Prod., 35: 397-408.

Dalton, D.C., 1987. An Introduction to Practical Animal Breeding. 2nd Edn., English Language Book Society/Collins, New York, pp: 182.

Frankel, O.H. and A.H.D. Brown, 1984. Current Plant Genetic Resources a Critical Appraisal. In: Genetics: New Frontiers, Chopra, V.L., B.C. Joshi, R.P. Sharma and H.C. Bansal (Eds.). Oxford and IBH, New Delhi, India, pp: 1-11.

Kohonen, T., 1995. Self-Organizing Maps. 1st Edn., Springer, New York, USA.

Kohonen, T., 2001. Self-organizing maps of massive databases. Int. J. Eng. Intell. Syst. Electr. Eng. Commun., 9: 179-185.

Kohonen, T., 2002. Overture. In: Self-Organizing Neural Networks: Recent Advances and Applications, Seiffert, U and L.C. Jain (Eds.). Chapter 1, Physica-Verlag, Heidelberg, Germany, ISBN-13: 9783790814170, pp: 1-12.

Krishnaswami, S., 1978. New Technology of Silkworm Rearing. Central Sericultural Research and Training Institute, Mysore, India, pp: 23.

Kumaresan, P., R.K. Sinha and S.R. Urs, 2007. An analysis of genetic variation and divergence in Indian tropical polyvoltine silkworm (*Bombyx mori* L.) genotypes. Caspian J. Environ. Sci., 5: 11-17.

Mahalanobis, P.C., 1936. On the generalized distance in statistics. Proc. Natl. Inst. Sci., 2: 49-55.

Mohammadi, S.A. and B.M. Prasanna, 2003. Analysis of genetic diversity in crop plants-salient statistical tools and considerations. Crop Sci., 43: 1235-1248.

Murakami, A., 1994. Growth phenomena in *Bombyx mori* L. with a special reference to genetic factors responsible for growth acceleration and moultinism. Indian J. Sericulture, 33: 12-14.

Murty, U.S.N., M. Rao and S. Misra, 2008. Prioritization of malaria endemic zones using self-organizing maps in the Manipur state of India. Inform. Health Social Care, 33: 170-178.

Murty, U.S., M. Rao, K. Sriram and K. Rao, 2011. Applications of Self-Organising Map (SOM) for prioritisation of endemic zones of filariasis in Andhra Pradesh, India. Int. J. Data Mining Bioinform., 5: 417-427.

Nagaraju, J., 2002. Application of genetic principles for improving silk production. Curr. Sci., 83: 409-414.

Nezhad, M.S., S.Z. Mirhosseini, S. Gharahveysi, M. Mavvajpour, A.R. Seidavi and M. Naserani, 2010. Genetic diversity and classification of 51 strains of silkworm *Bombyx mori* (Lepidoptera: Bombycidae) germplasm based on larval phenotypic data using Ward's and UPGMA methods. Afr. J. Biotechnol., 9: 6594-6600.

Ramamohana Rao, P. and T. Nakada, 1998. Clustering of polyvoltine strains of the silkworm *B. mori* by image processing method: Significance of cocoon size and weight variables. Indian J. Seric., 37: 33-39.

Ramesha, C., S.V. Seshagiri and C.G.P. Rao, 2009. Evaluation and identification of superior polyvoltine crossbreeds of mulberry silkworm, *Bombyx mori* L. J. Entomol., 6: 188-197.

Ramesha, C., C.M. Anuradha, H. Lakshmi, S.S. Kumari, S.V. Seshagiri, A.K. Goel and C.S. Kumar, 2010. Nutrigenetic traits analysis for the identification of nutritionally efficient silkworm germplasm breeds. Biotechnology, 9: 131-140.

Roy, J. and X. Lin, 2000. Latent variable models for longitudinal data with multiple continuous outcomes. Biometric, 56: 1047-1054.

Tazima, Y., 1964. The Genetics of the Silkworm. Logos Press, London, UK., pp: 174-195.

Thangavelu, K., R.K. Sinha and B. Mohan, 2003. Silkworm germplasm and their potential use. Proceedings of The Mulberry Silkworm Breeders Summit, July 18-19, 2003, Hindupur, India, pp: 14-23.

Vargas, F.J., M.A. Romero and I. Vargas, 2002. Flowering precocity in pistachio progenies. Acta Hort., 591: 297-303.

Weir, B.S., 1996. Genetic Data Analysis II: Methods for Discrete Population Genetic Data. 2nd Edn., Sinauer Association, Sunderland, Massachusetts. USA.