

**Genetic Characterization and Estimation of  
Linkage Disequilibrium of Sunflower Oil and  
Confection Types using Genome-Wide  
Distributed SNP Markers**

**Venkatramana Pegadaraju & Quentin Schultz,  
Lili Qi & Brent Hulke**

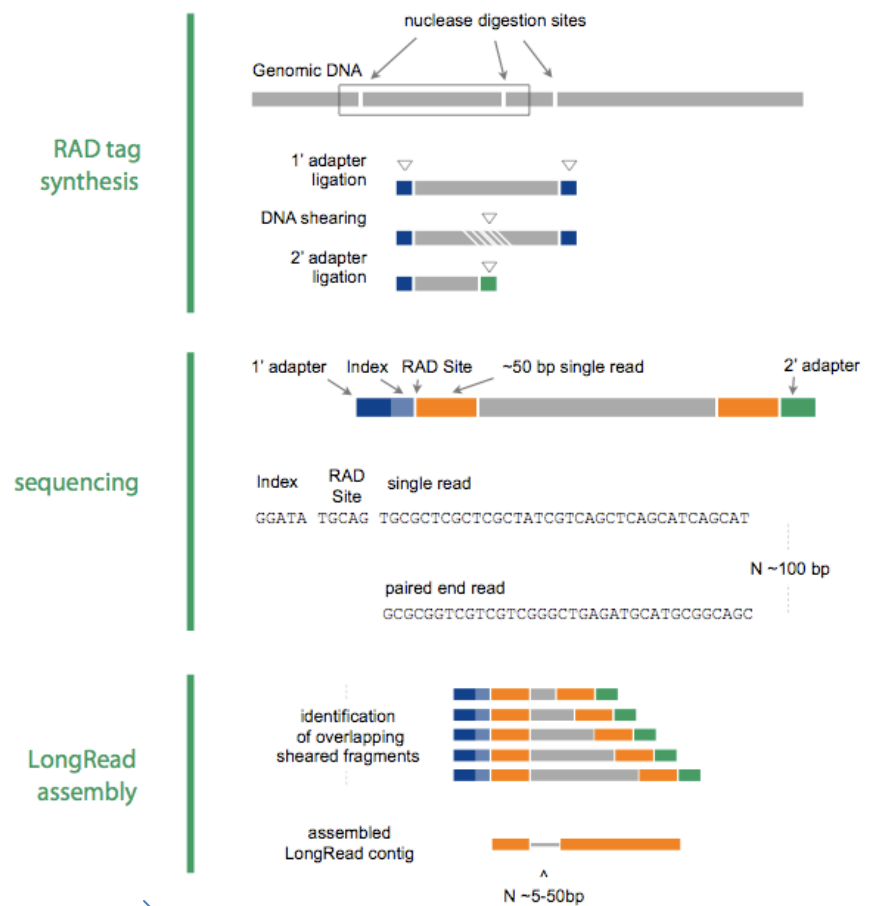
NSA Research Forum, 2012

# SNP Discovery in Sunflower

## Sequencing Panel

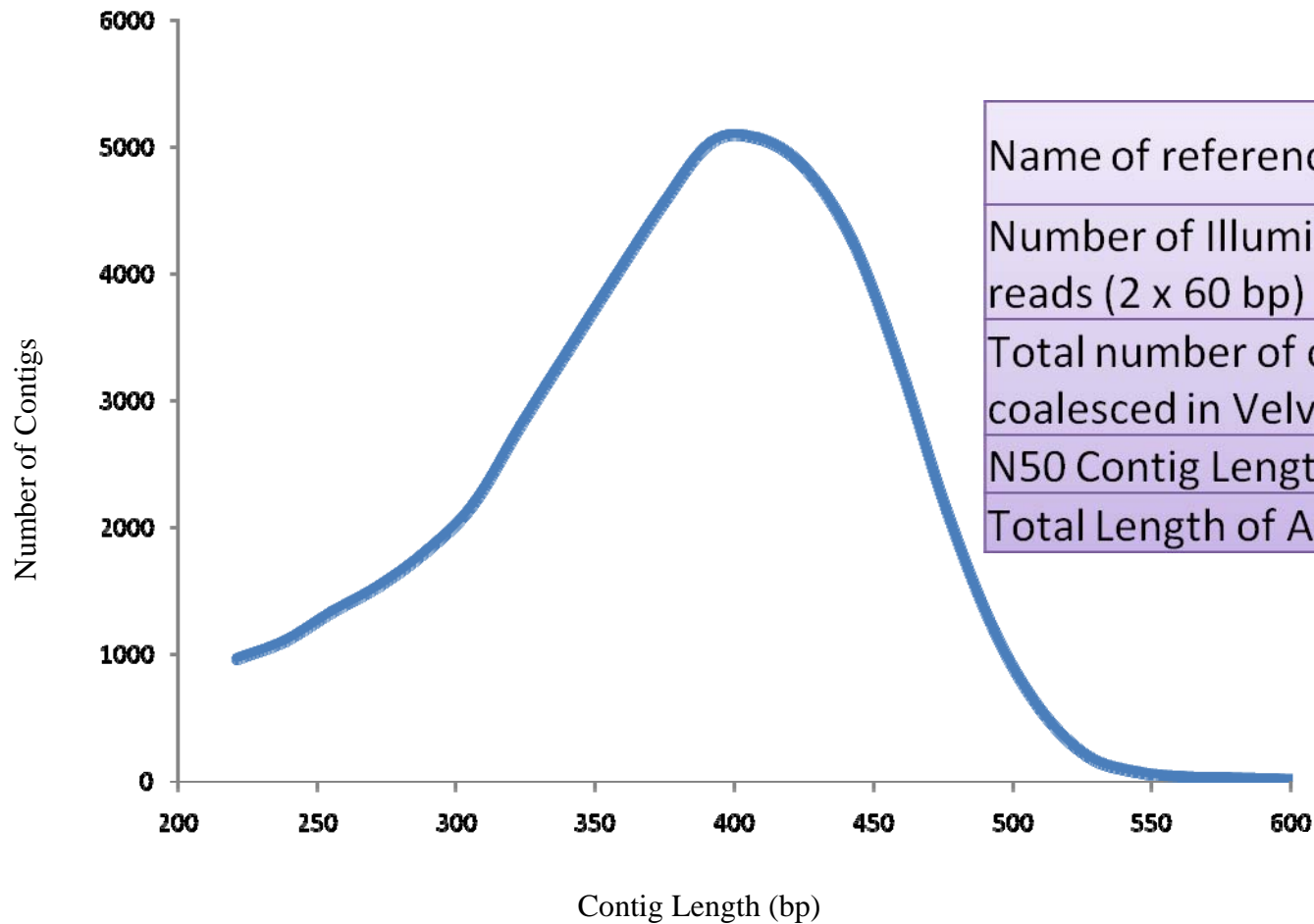
<b>Sunflower Line</b>
TX1612
CR29
Seeds 2000 Confection B Line
HA467
RHA468 699-10
RHA464 09 098-4

## Sequencing Approach



# Reference Genome Assembly

---

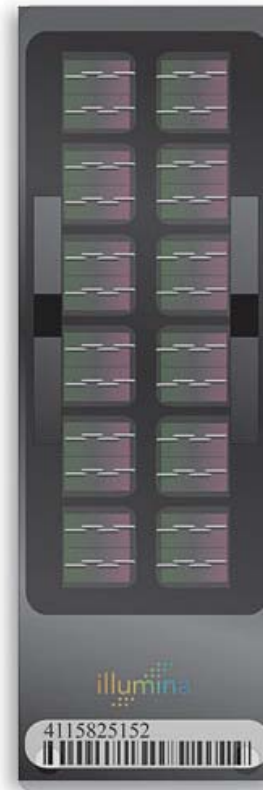
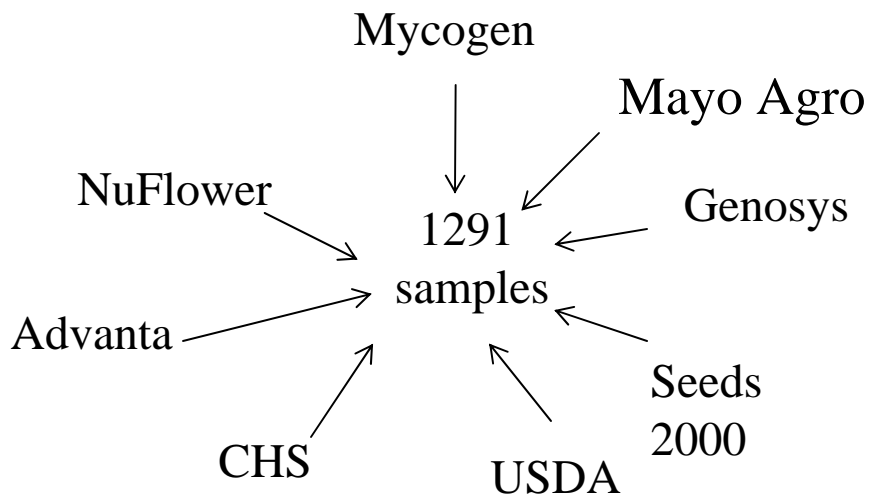


Name of reference cultivar	BDI_Sunflower_06
Number of Illumina paired end reads (2 x 60 bp) to obtain ref	9,016,941
Total number of contigs coalesced in Velvet assembly	50,726
N50 Contig Length	379 bp
Total Length of Assembly	18.87 Mbp

# Sunflower Genotyping Panel

---

## Diversity Panel



**8723**  
**SNP markers**

## Mapping Panel

HA89 x RHA464



F1



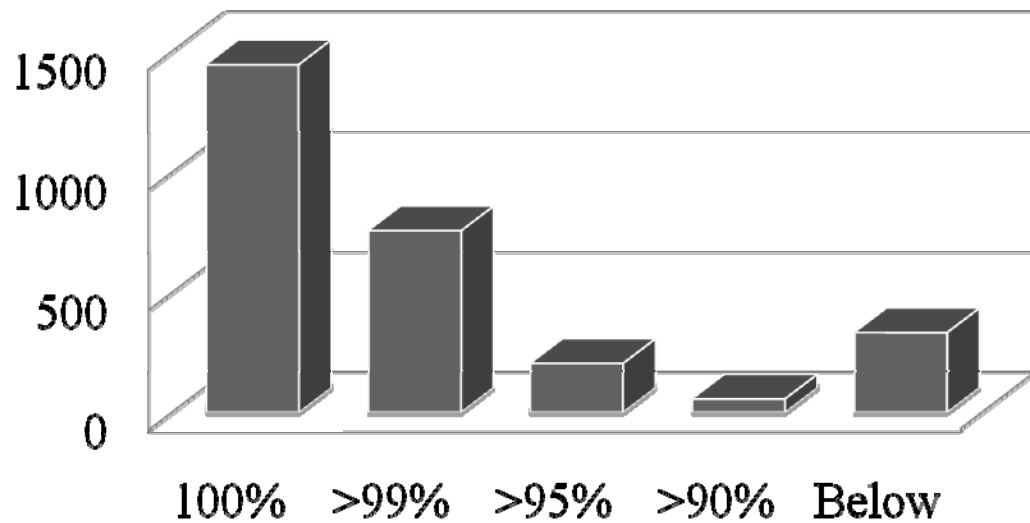
Self

141, F2 lines  
genotyped

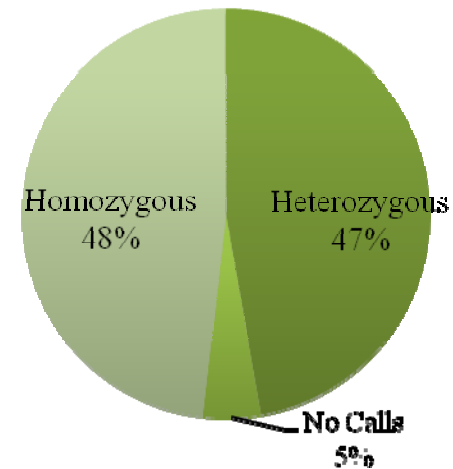
# HA89xRHA464 Mapping Population Summary Stats

---

**Marker Call Rates  
(polymorphic)**



**Mapping Population  
Composition**

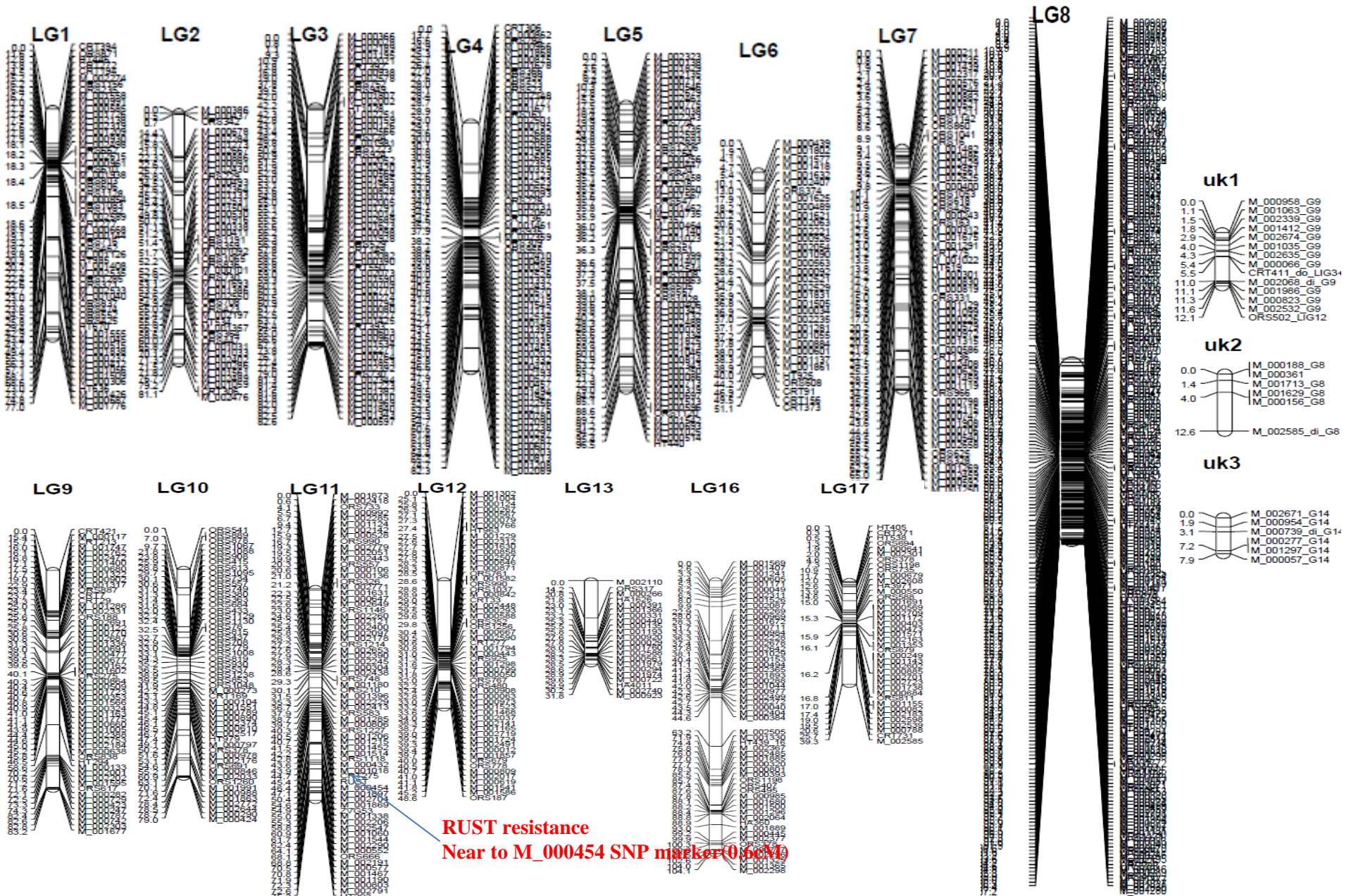


# Markers interrogated	8723
# Successfully scored	8027
# Polymorphic markers	2797

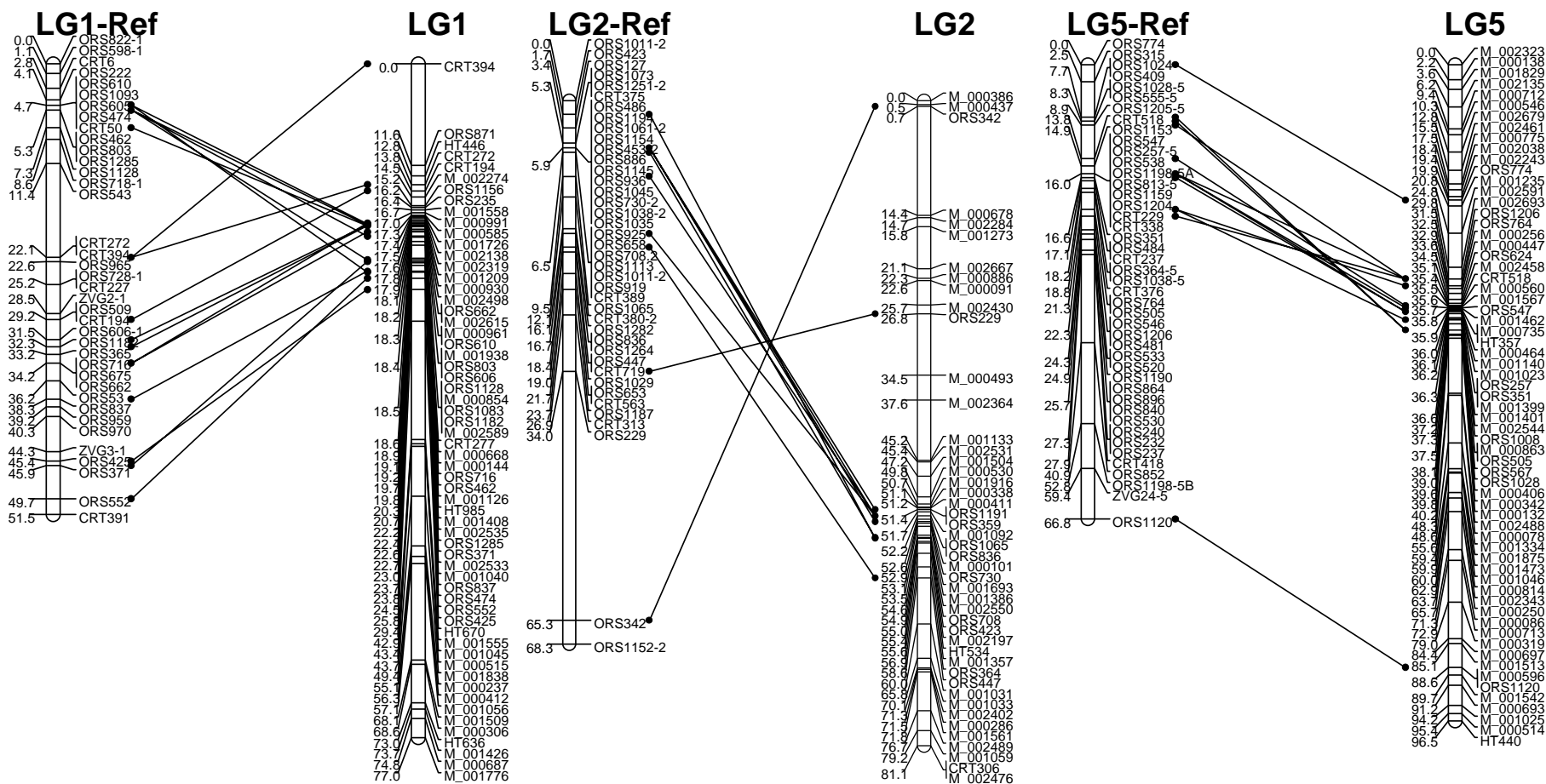
# Mapping SNP markers to Sunflower Linkage Groups using HA89xRHA464 (F2-population)

SSR LG	Markers Mapped (Phase I)		Markers Mapped (Phase II)				
	SNPs Loci	Total SNP markers (with redundant markers)	Mapped SSRs	Total SSR/SNP markers (SSR markers intergated in the SNP map without the redundant SNPs)	Total (SSR/SNP) loci	Total SSR/SNP markers (with redundant markers)	Lengh (cM) Kosambi
LG 1	76	328	29	105	80	357	76.96
LG 2	33	98	13	46	46	111	81.10
LG 3	87	139	12	99	98	151	82.56
LG 4	65	159	12	77	75	171	82.32
LG 5	51	199	15	66	65	214	96.53
LG 6	36	60	6	42	37	66	51.11
LG 7	65	117	15	80	67	132	66.40
LG 8	212	756	60	272	260	816	107.21
LG 9	73	113	10	83	69	123	83.18
LG 10	20	47	35	55	54	82	79.04
LG 11	58	160	16	74	72	176	75.30
LG 12	86	152	14	100	84	166	48.578
LG 13	22	59	1	23	22	60	37.87
LG-D	5	6	0	5	14	6	12.64
LG-A	12	18	2	14	5	20	12.09
LG-B	3	8	2	5	5	10	0.48
LG-E	35	54	0	35	5	54	45.52
LG-C	6	6	0	6	34	6	7.89
LG 16	23	39	2	25	25	41	55.67
LG 17	13	44	9	22	21	53	22.65
	981.00	2562.00	253.00	1234.00	1138.00	2815.00	<b>1125.09</b>

# Sunflower SNP Linkage Map



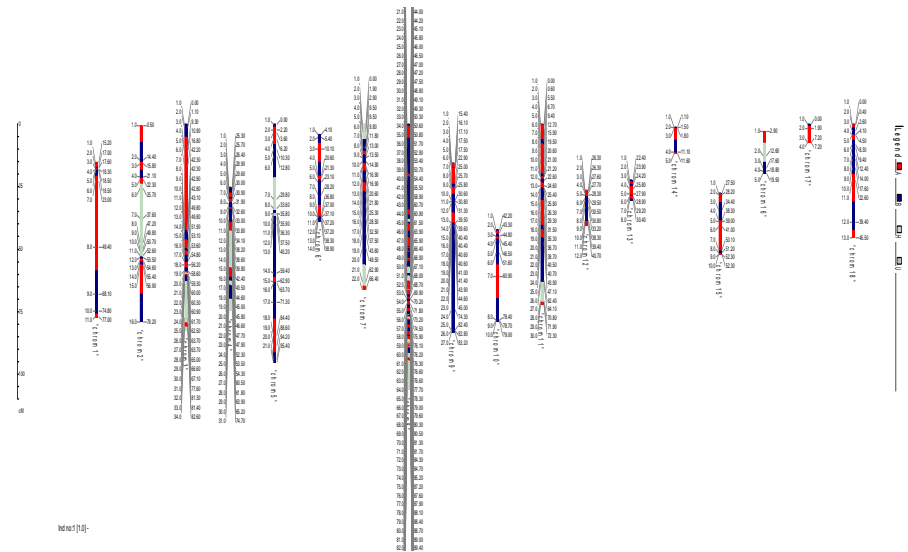
# Comparison of NSA SNP Linkage Map to SSR Reference Map



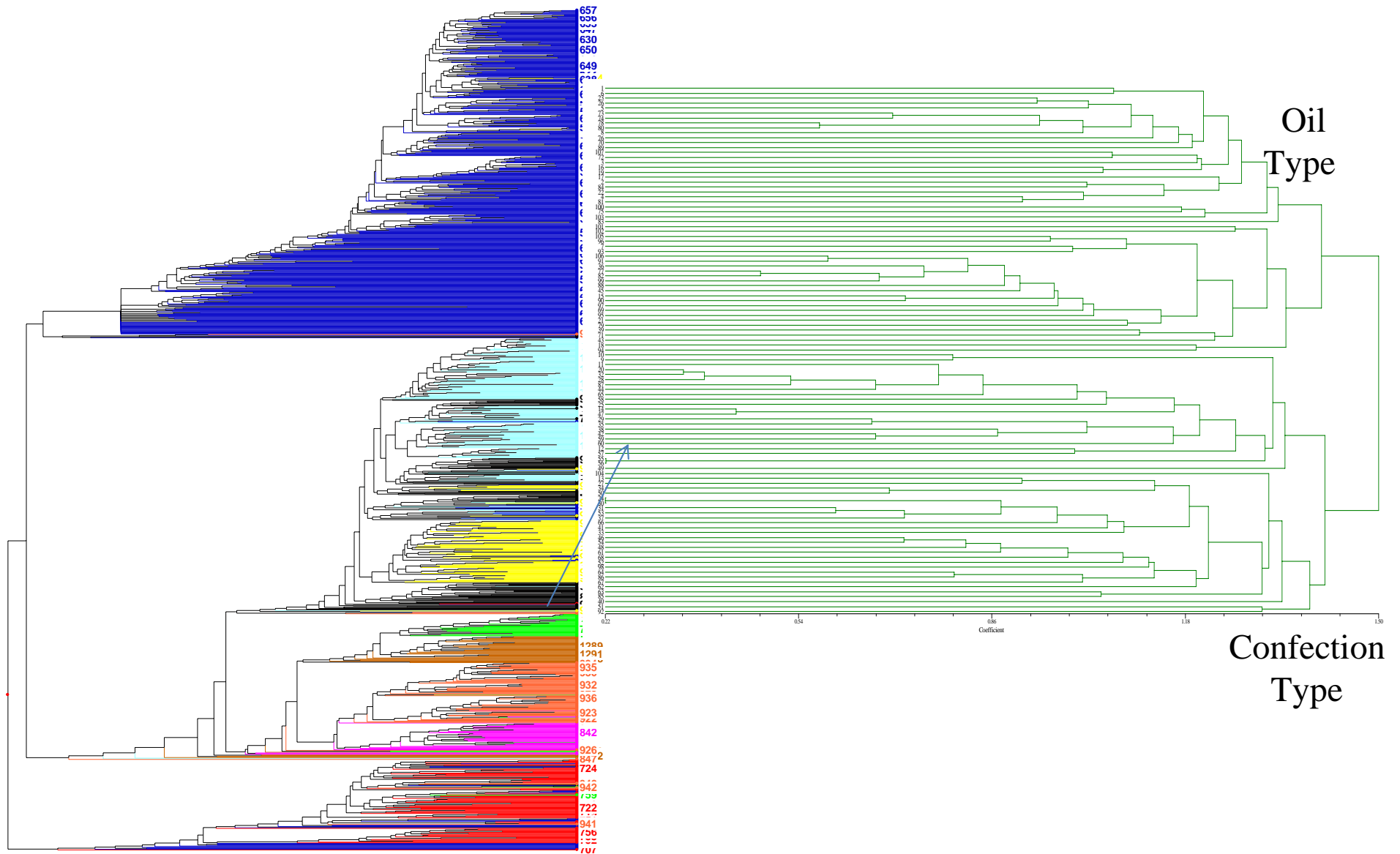


# Selection of Genome Wide Distributed 384 SNP Panel

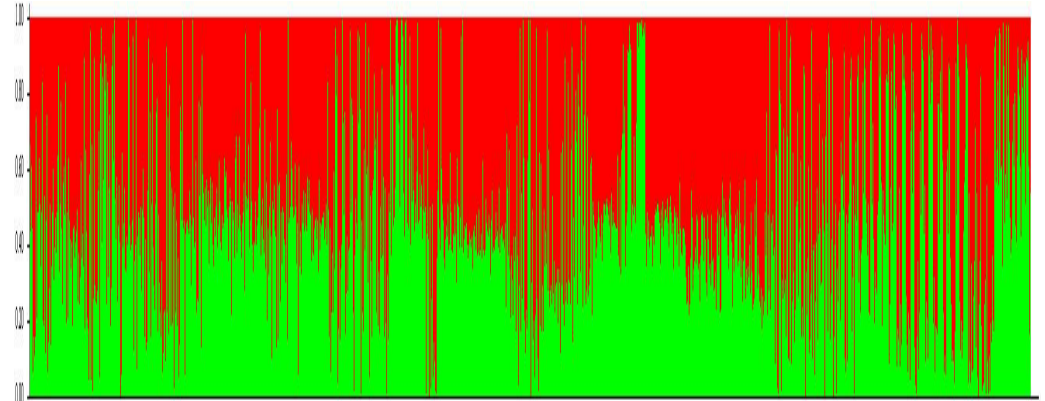
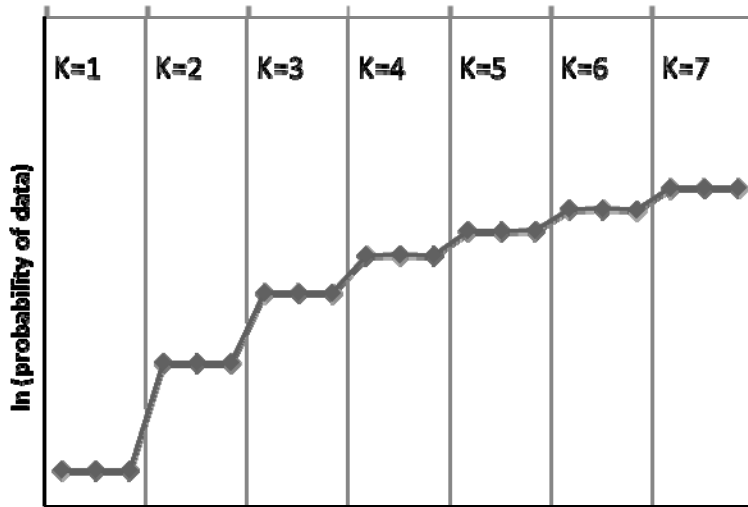
Chromo #	SNP markers	Length (cM)	PIC range
1	11	76.9	0.35-0.375
2	16	79	0.31-0.375
3	34	82.56	0.31-0.375
4	31	74.739	0.3-0.375
5	21	95.375	0.33-0.375
6	14	38.9	0.31-0.375
7	22	66.4	0.32-0.375
8	104	106.38	0.30-0.375
9	17	83.18	0.31-0.375
10	10	79.04	0.33-0.375
11	30	72.27	0.32-0.374
12	12	40	0.31-0.375
13	8	30.423	0.33-0.375
16	15	52.309	0.31-0.375
17	5	19.481	0.36-0.375
LG-A	5	11.568	0.33-0.372
LG-B	1	0.48	0.36
LG-C	4	7.224	0.33-0.372
LG-D	1	0	0.35
LG-E	13	45.23	0.32-0.372
Total	384	1061.459	0.36



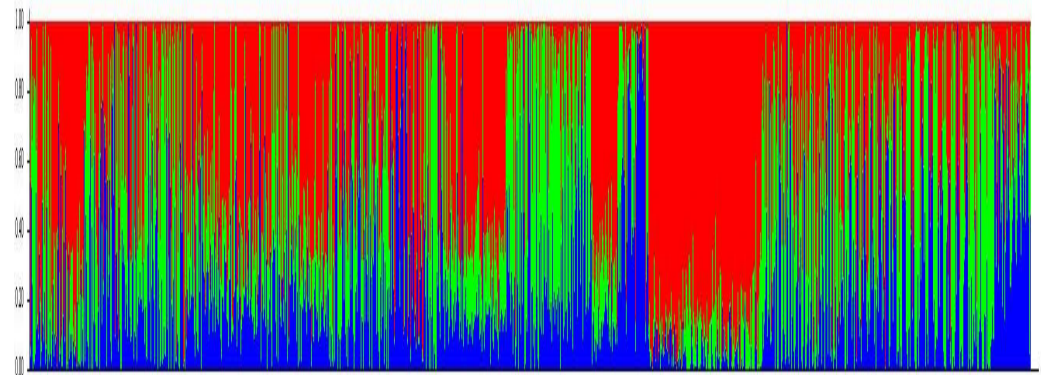
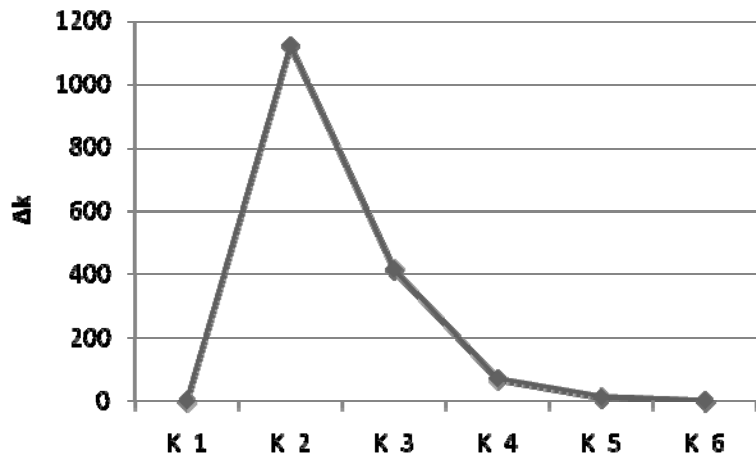
# Genetic Characterization of Sunflower Diversity Panel



# Estimating Population Structure



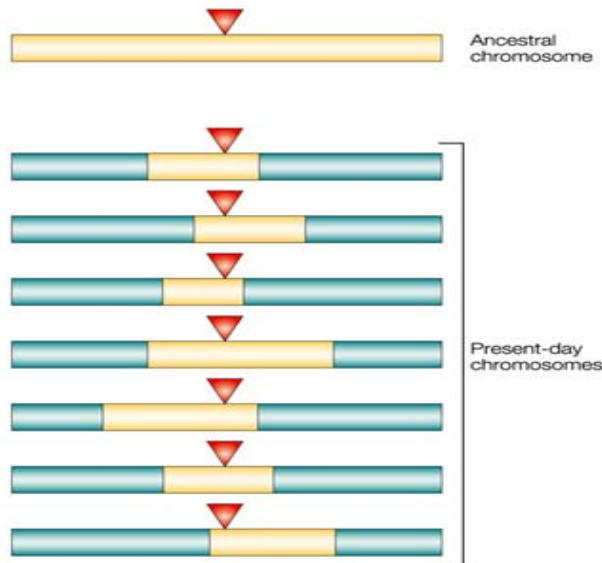
K value=2



K value=3

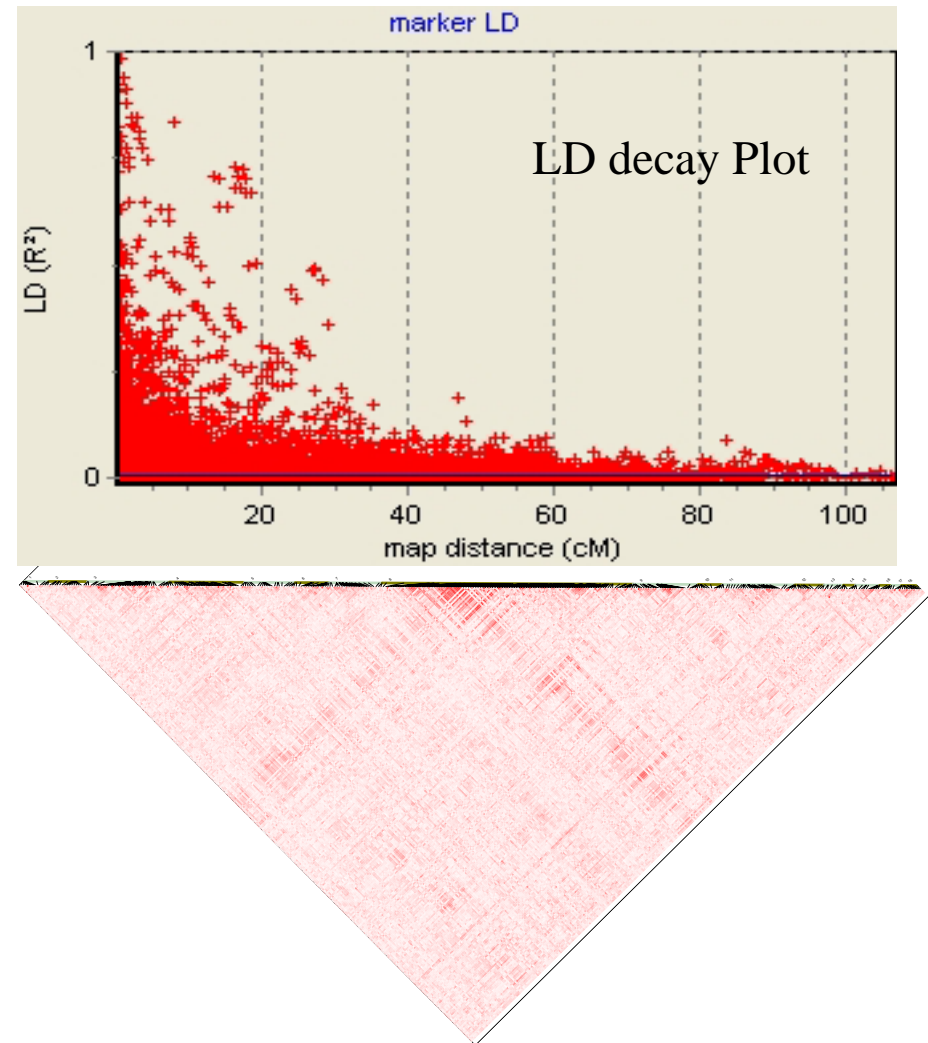
# Whole genome LD Analysis on Sunflower Diversity Panel

LD refers to nonrandom association of alleles at different loci.



Factors influencing LD :

- Recombination rate
- Mutation rate
- Population size
- Natural selection



# Future Directions

---

- Integrate additional SNP markers into NSA Linkage map ver. 1 & develop a high density SNP-based genetic map for sunflower (June-2012)
  - Nuseed R-line X RHA468 (144 F2 lines)
  - Nuseed B-line X RHA464(144 F2 lines)
- Develop trait linked marker sets for key agronomic traits through validation studies and start offering them as service for routine MAS projects to breeders(USDA collaboration)
- In collaboration with USDA(Dr. Brent Hulke )develop tools and framework models to support advance MAS programs.