



Routine Computation of Coefficient of Parentage Using the International Crop Information System

Graham McLaren, Ian DeLacy and Jose Crossa
Crop Research Informatics Laboratory

The Coefficient of Parentage

- The COP measures the genetic relationship between strains of germplasm according to the proportion of identical alleles that they share by descent through their pedigrees.
- Wright first proposed the measure he called the Coefficient of Relationship in 1922 in order to compute the Coefficient of Inbreeding in livestock.
- They are extensively used in livestock breeding where replication of genotypes is difficult so information on genetic effects must come from relatives

Use of COP in Plant Breeding

- Used in diversity studies for soybean (Cox 1985 and Sneller 1994), winter wheat (Murphy, 1986 and Cox et al., 1986), in rice (Dilday, 1990) and spring wheat (Souza et al., 1994)
- Used to study the relationship between genealogical distance and breeding behavior in rapeseed (Lefort-Buson et al., 1986), oats (Cowen and Frey, 1987) and maize (Bernardo, 1994).
- Bernardo (2002) demonstrates the use of COP values for cross prediction in hybrid breeding programs

Constraints to use of COP in Plant Breeding

- Lack of pedigree records and data management systems to record them
- Lack of software to compute COP values and take account of self fertilization and relationships between sister lines
- Lack of easy to use mixed model software to take advantage of relationships between relatives in analysis of evaluation data

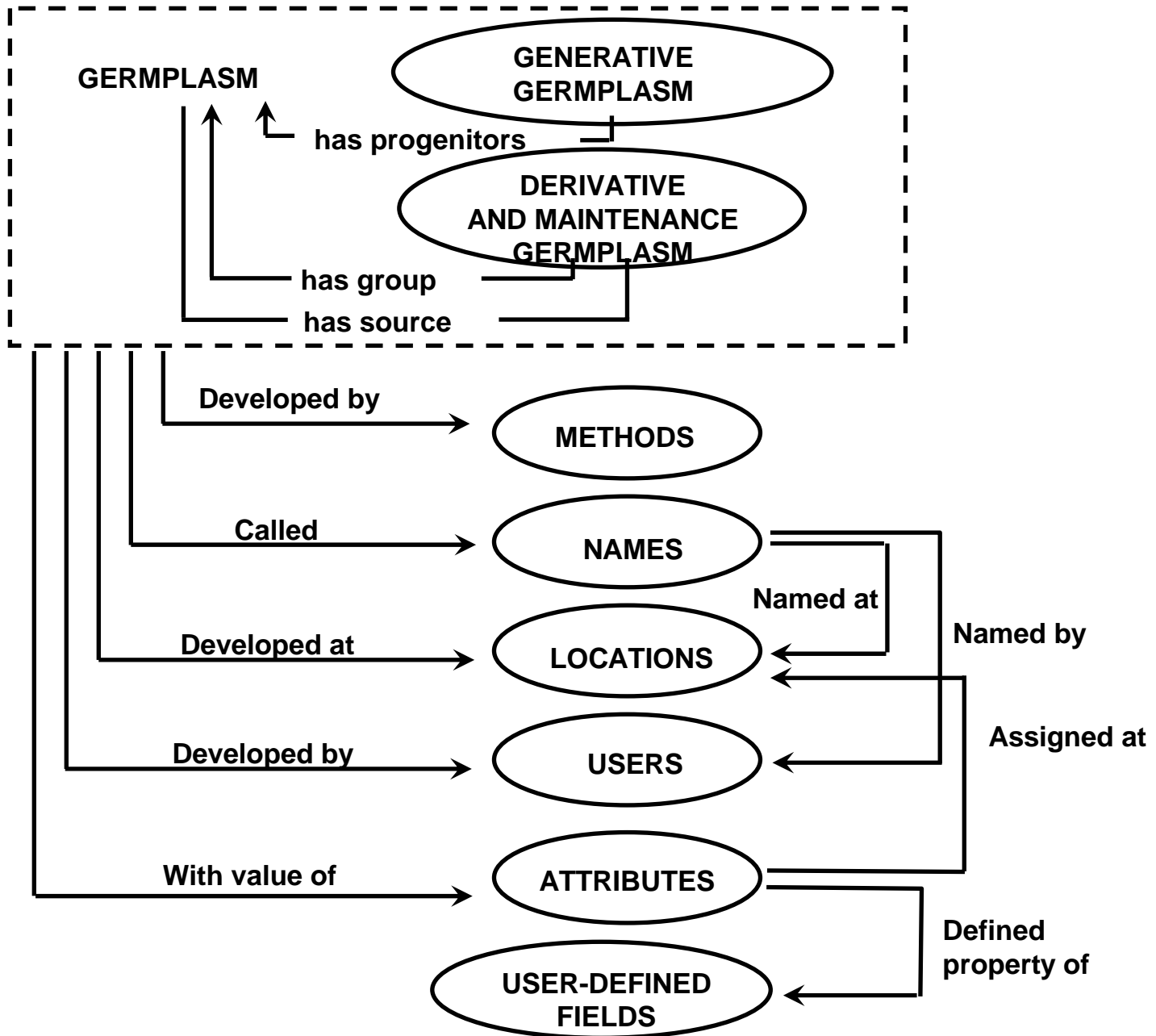
Basic relationships for Calculating the COP

- The COP between strains P and Q , f_{PQ} , can be expanded as $f_{PQ} = (f_{PC} + f_{PD})/2$ if Q resulted from a cross between C and D.
- If Z results from n generations of selfing after a cross between C and D the the coefficient of inbreeding of Z is $F_Z = 1 - (1 - f_{CD})/2^n$
- The COP, f_{ZZ} , for a strain Z with itself is given by $f_{ZZ} = (1 + F_Z)/2$

Relationship of COP to Genetic Variance

- The breeding value of a strain is the effect of its alleles on the phenotype of its progeny
- The breeding value of a strain determines its potential as a parent
- The variance of breeding values for a trait over all individuals in a population is the additive genetic variance for that trait in that population
- Breeding values of related strains in a population are correlated due to IBD alleles – the additive genetic correlation between strains is twice their COP

Data model for the ICIS Genealogy Management System

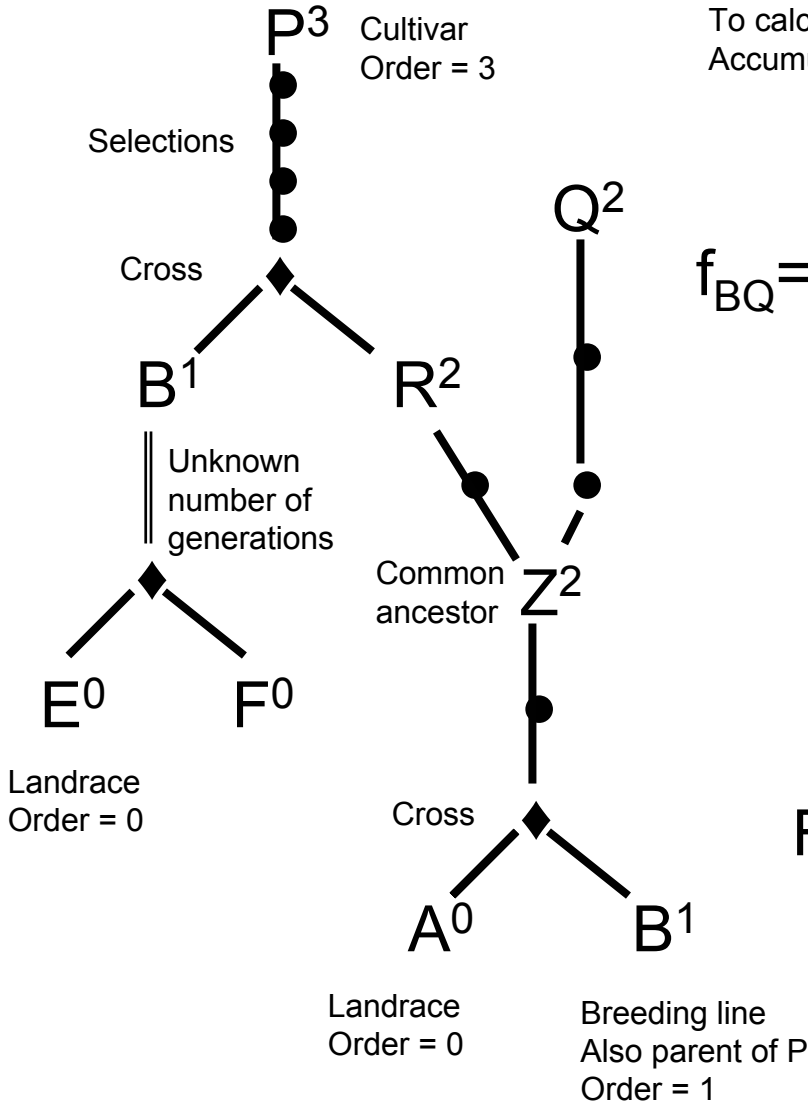


Computation of COPs with BROWSE

1. Extract the full pedigree of both lines from the database.
2. Determine the order of the ancestors – the largest number of generative steps from an ancestor to a terminal ancestor.
3. Use the recursive equation $f_{PQ} = (f_{PC} + f_{PD})/2$ to accumulate COP values from terminal ancestors.
4. Use the inbreeding relationships $F_Z = 1 - (1 - f_{CD})/2^n$ and $f_{ZZ} = (1 + F_Z)/2$ to account for inbreeding and sister lines.

Example of COP calculations for strains P and Q

To calculate the COP expand the formula from the top down and Accumulate the computations from the bottom up



$$f_{PQ} = (f_{BQ} + f_{RQ}) / 2 = 87/128$$

$$f_{BQ} = (f_{BA} + f_{BB}) / 2 = 31/64$$

$$f_{RQ} = f_{ZZ} = (1 + F_Z) / 2 = 7/8$$

$$F_Z = 1 - (1 - f_{AB}) / 2^n = 1 - 1/2^2 = 3/4$$

$$f_{BB} = (1 + F_B) / 2 = 31/32$$

$$F_B = 1 - (1 - f_{EF}) / 2^{n+1} = 1 - 1/2^4 = 15/16$$

$$f_{BA} = (f_{EA} + f_{FA}) / 2 = 0$$

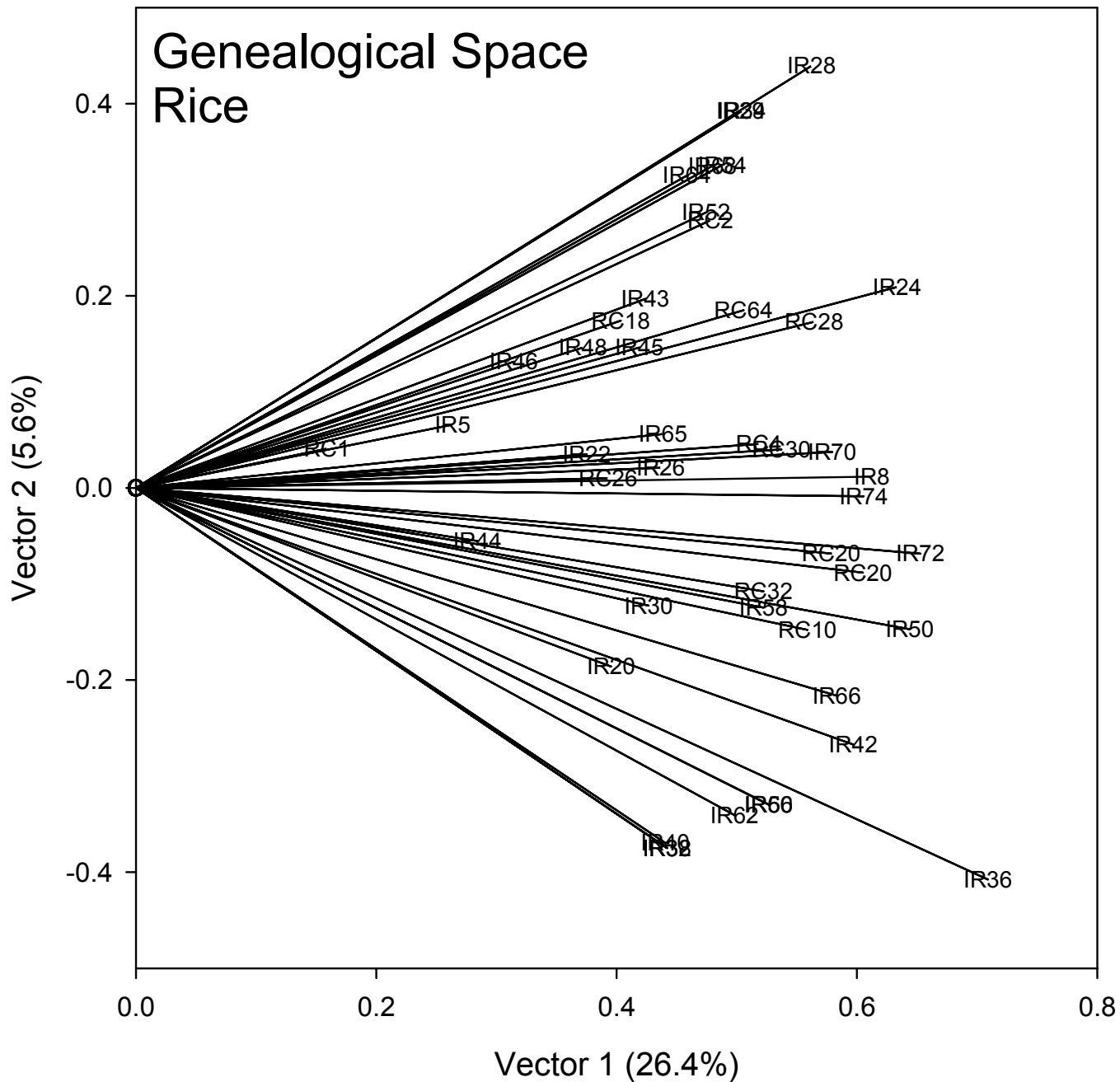
Generalizations and assumptions

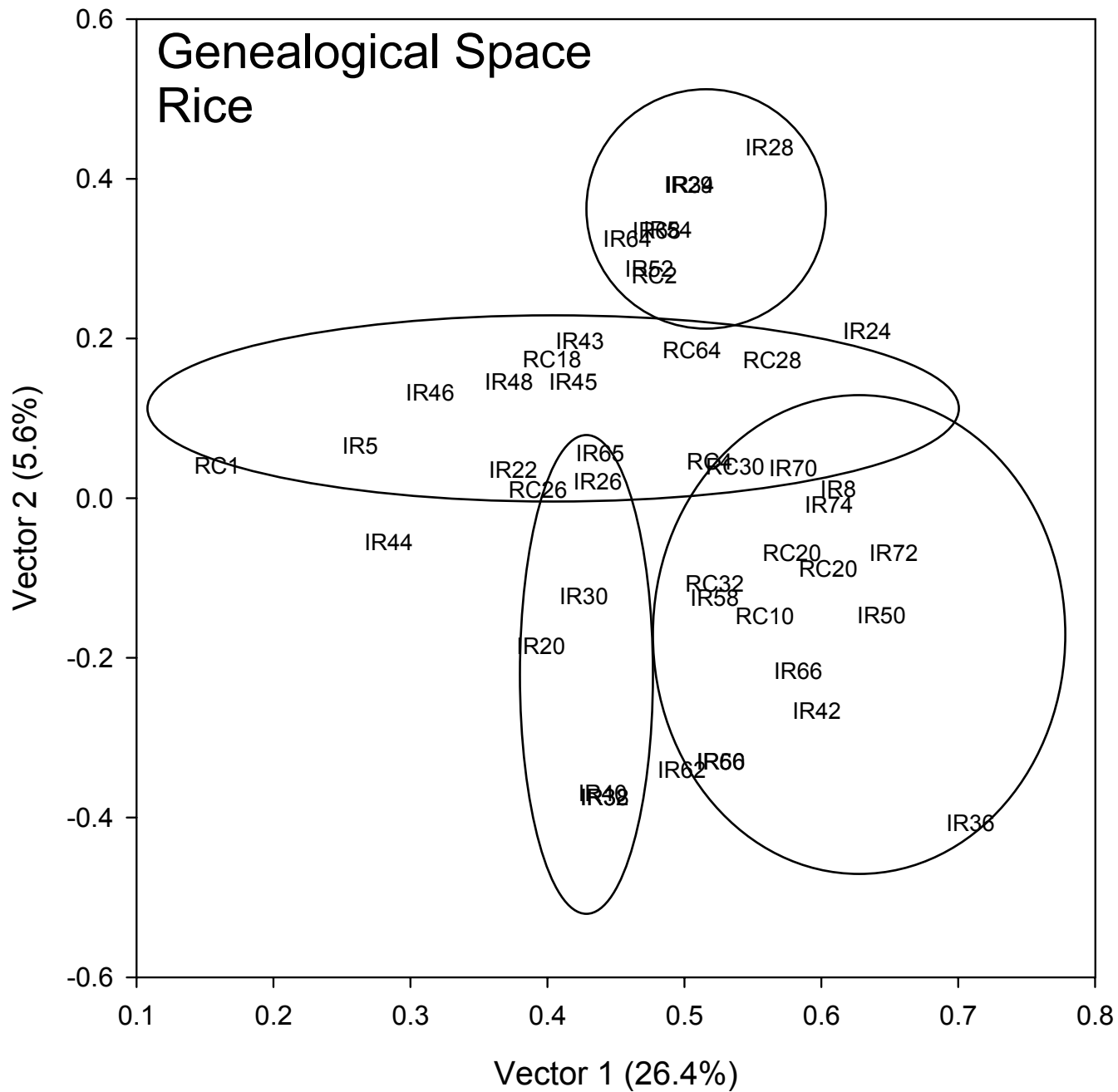
- If Q has progenitors $Q_1, Q_2 \dots Q_m$ then $f_{PQ} = (1/m) \sum f_{PQi}$
- If P and Q are derived from different crosses the COP between P and Q is unaffected by inbreeding.
- If the progenitors of strain Z are unknown then set F_Z to 1 in self fertilizing crops and to 0 in outcrossing crops.
- If the number of selfing generations is unknown for a breeding line set F_Z to 15/16, equivalent to F4 generation.
- If P and Q are sister lines then their COP is affected by selfing up to their most recent common ancestor Z and $f_{PQ} = (1+F_Z)/2$.

Calculating matrices of COPs

- Use an ICIS germplasm list to identify all the strains
- Extract all ancestors into a single pedigree table
- Store intermediate COP values in a sparse matrix
- The COP matrix is symmetric so we only output the lower triangular part with its inverse

Genealogical Space Rice

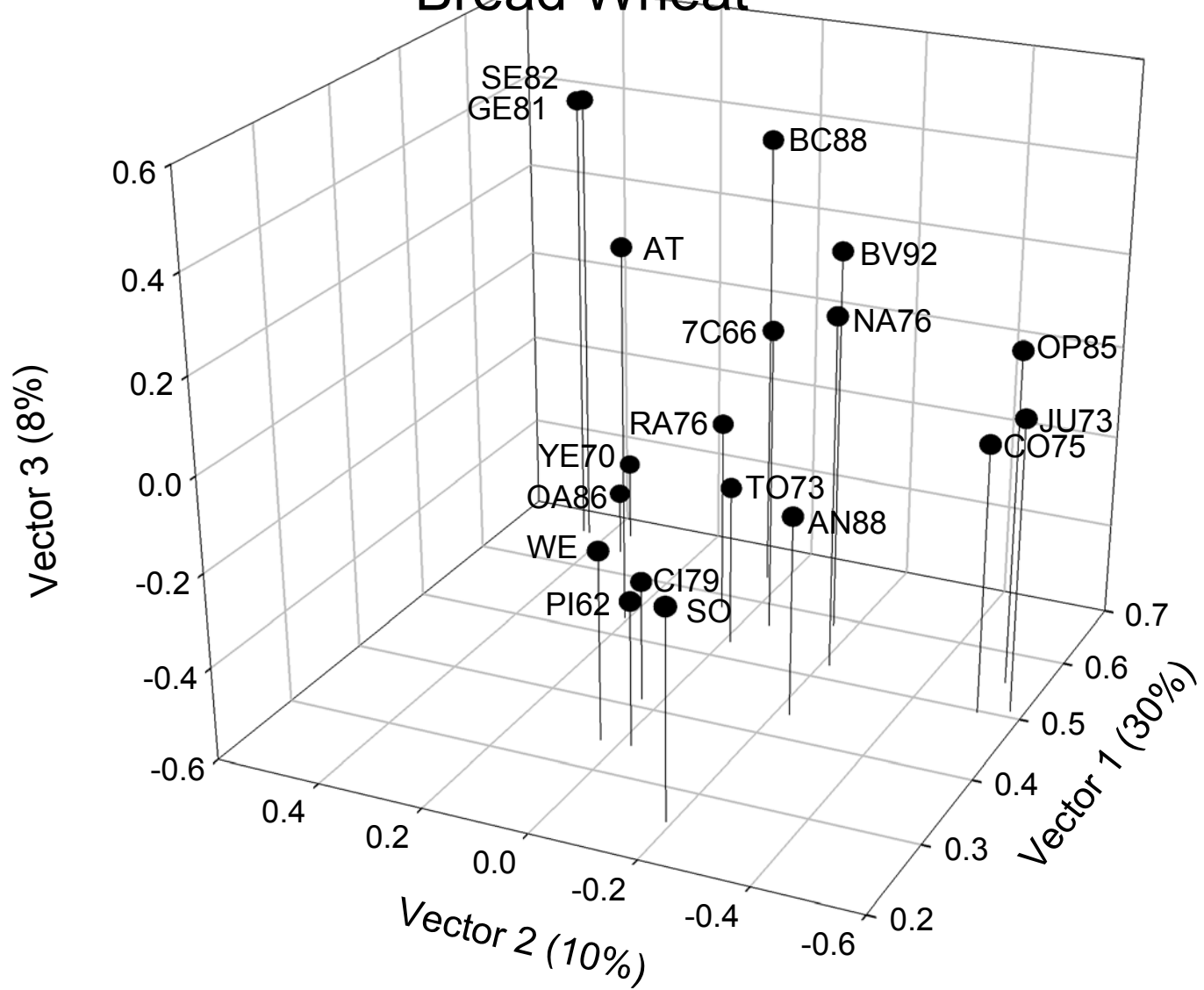




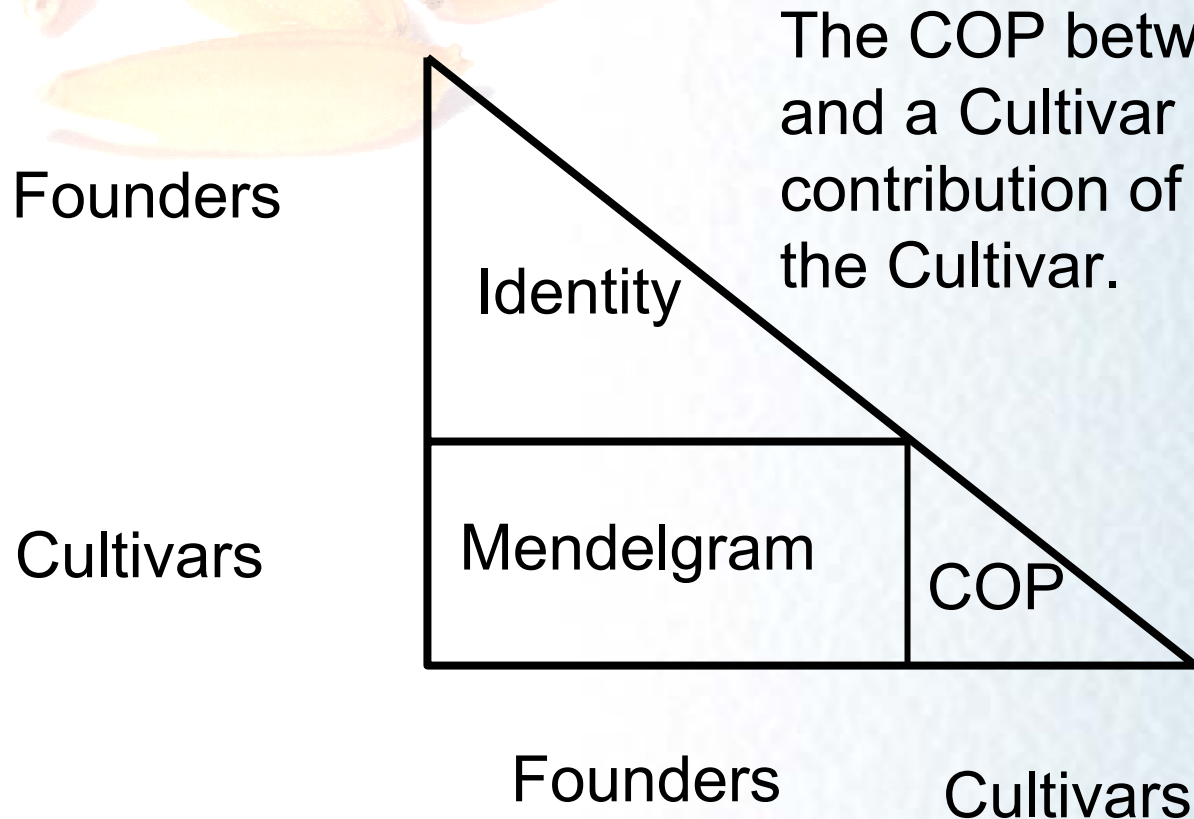
Wheat cultivars released in Mexico derived from CIMMYT breeding lines

ENTRYID	ENTRYCD	SOURCE	DESIGNATION
1	PI62	MEXICO	PITIC 62
2	7C66	MEXICO	SIETE CERROS T 66
3	SO	MEXICO	SONALIKA
4	YE70	MEXICO	YECORA F 70
5	JU73	MEXICO	JUPATECO F 73
6	TO73	MEXICO	TORIM F 73
7	CO75	MEXICO	COCORAQUE F 75
8	NA76		NACUZARI F 76
9	PA76	MEXICO	PAVON F 76
10	CI79	MEXICO	CIANO T 79
11	GE81	MEXICO	GENARO T 81
12	SE82	MEXICO	SERI M 82
13	OP85	MEXICO	OPATA M 85
14	OA86	MEXICO	OASIS F 86
15	AN88	MEXICO	ANGOSTURA F 88
16	AT	MEXICO	ATTILA
17	BC88	MEXICO	BACANORA T 88
18	BV92	MEXICO	BAVIACORA M 92
19	WE		WEAVER

Genealogical Space Bread Wheat



Viewing a Mendelgram



The COP between a Founder and a Cultivar is the proportional contribution of the Founder to the Cultivar.

BIPLOT OF FIRST TWO ORDINATION SCORES FOR MENDELGRAM
 OF CIMMYT CULTIVARS (MODEL FIT: 88.7% OF GXE SS
 PLOT LABELS ARE CULTIVAR CODES AND FOUNDER NUMBERS)

