

# Main change from the previous discussion

- Adding one table to store the unit and condition of living stock for phenotyping and genotyping measurement – we could call that table ‘OU’
- Storing the hierarchical ‘living stock’ data in Stock table instead of OU table
  - It’s almost like changing the name from OU to Stock, and add a new table called ‘OU’
  - OU is for both observation\_unit [a living stock, part of living stock (specimen), or a clone of a living stock] and environment (treatments, collection year, etc)
- Why? - To add another layer on phenotype (genotype) data
  - Project - plant (Fuji) – phenotype (sugar content)
  - Project - plant (Fuji) – property (planted in plot A, treated with fertilizer B, and collected in 2008) – phenotype (sugar content)
  - Those properties can’t be stored in the phenotype assay (sugar content assay)
- Examples:
  - [GDPDM](#)
  - [ICIS](#) (International Crop Information System)

\*\*Chado user could choose to store individual clone or specimen in ‘stock’ – then ‘ou’ table will be more to store environment/collection date..

# Example breeding data

PlantYear	Orchard	*Code	**Rep	spring '04	fall '04	fall '05	Flower	Fruit	Mean wt
				Diam(mm)	Diam(mm)	Diam(mm)	Number	Number	
2004	A	13	2	.	18	36	77	17	232
2004	A	6	2	.	16	34	32	2	241
2004	A	3	2	.	18	37	60	24	92
2004	A	8	2	.	16	25	36	6	164
2004	A	12	2	.	19	27	46	27	250
2004	A	5	2	.	19	25	22	6	210

\*genotype; \*\*plot number

- Users would like to search for all the stocks with Fruit number over 20
- Only the “fruit number” can be a phenotype, not orchard, plot number, plant year or any treatment

## 1. Stock

- Stock (germplasm) named by breeders, researchers or germplasm centers
- can be a population, an established cultivar/breeding line or an individual (eg. collected from wild or from a cross) but not likely to be clones that are propagated in breeding programs (presentation on [Germplasm Ontology](#))
- link to pedigree and passport data

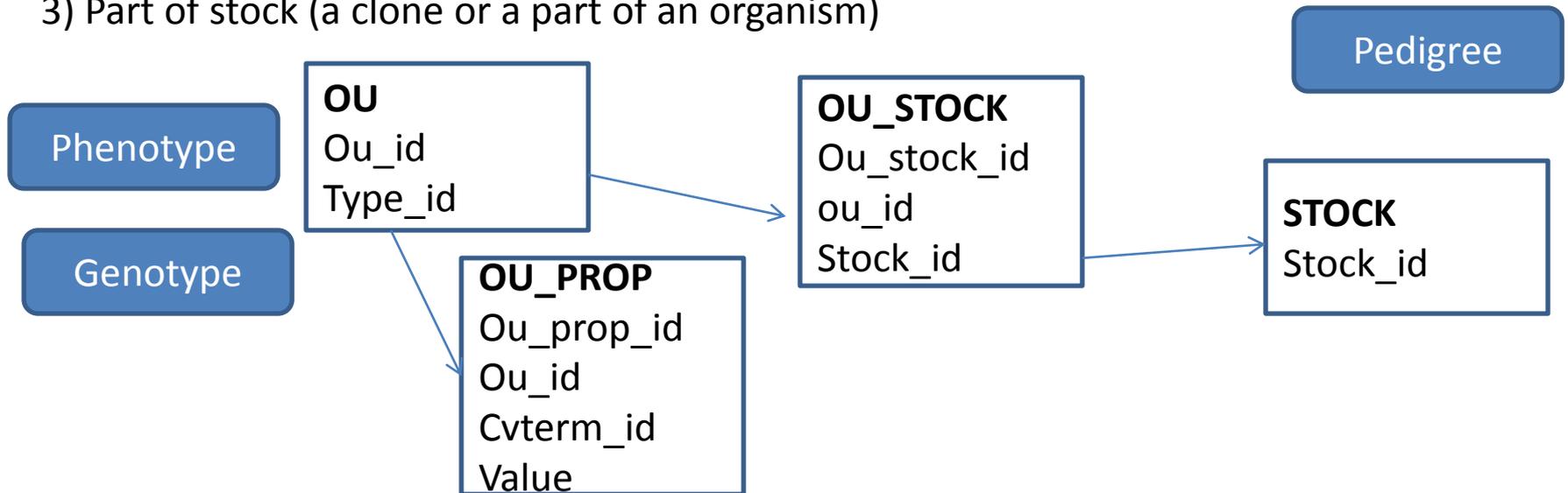
## 2. OU

- observational unit with specific properties (inventory/plot number, treatment, plant/harvest date, etc)
- links to a distinct phenotype and genotype value per assay
- can be any of the below with SPECIFIC FACTORS

### 1) Stock

### 2) Group of stocks (eg. F3 population)

### 3) Part of stock (a clone or a part of an organism)



# Breeding Data (apple)

Female x Male



Test 1<sup>st</sup> seedlings (1<sup>st</sup> Phase)



Pick 50 genotypes  
and propagate (15 trees per genotype)

Test in three different trial sites  
(5 trees per genotype in each site)  
Controls are included (2<sup>nd</sup> Phase)



Pick 15 genotypes  
and propagate (100 trees per genotype)

Test in four different sites  
(25 trees per genotype in each site)  
Controls are included (3<sup>rd</sup> Phase)

All seedlings are stored in  
Stock and OU (property  
include site, date, etc)

The 50 genotypes with  
specific properties (site,  
date, etc) is stored in OU  
The phenotype value will  
be the mean of 5 trees

The 5 individual tree with  
specific properties (position  
in a plot, site, date, etc) will  
be stored in OU

# Breeding Data (apple)

Female x Male



Test 1<sup>st</sup> seedlings (one of them WSU123)



Pick 50 genotypes  
and propagate (15 trees per genotype)

Test in three different trial sites  
(5 trees per genotype in each site)  
Controls are included



Pick 15 genotypes  
and propagate (100 trees per genotype)

Test in four different sites  
(25 trees per genotype in each site)  
Controls are included

OU entries

1. Wsu123 @ nursery @tested on 1/1/2008 @1<sup>st</sup> phase
2. Wsu123 @ site 1 @tested on 1/1/2009@2<sup>nd</sup> phase
3. Wsu123 @ site 2 @plot 1 @tested on 1/1/2009@2<sup>nd</sup> phase



Stock entry  
WSU123

1 and 2 is 'stock' and 3 is 'part of stock'

# QTL (tomato data from SGN)

Sausage (F) x LA1589 (M)



F1



107 F2 lines



Phenotypes were evaluated

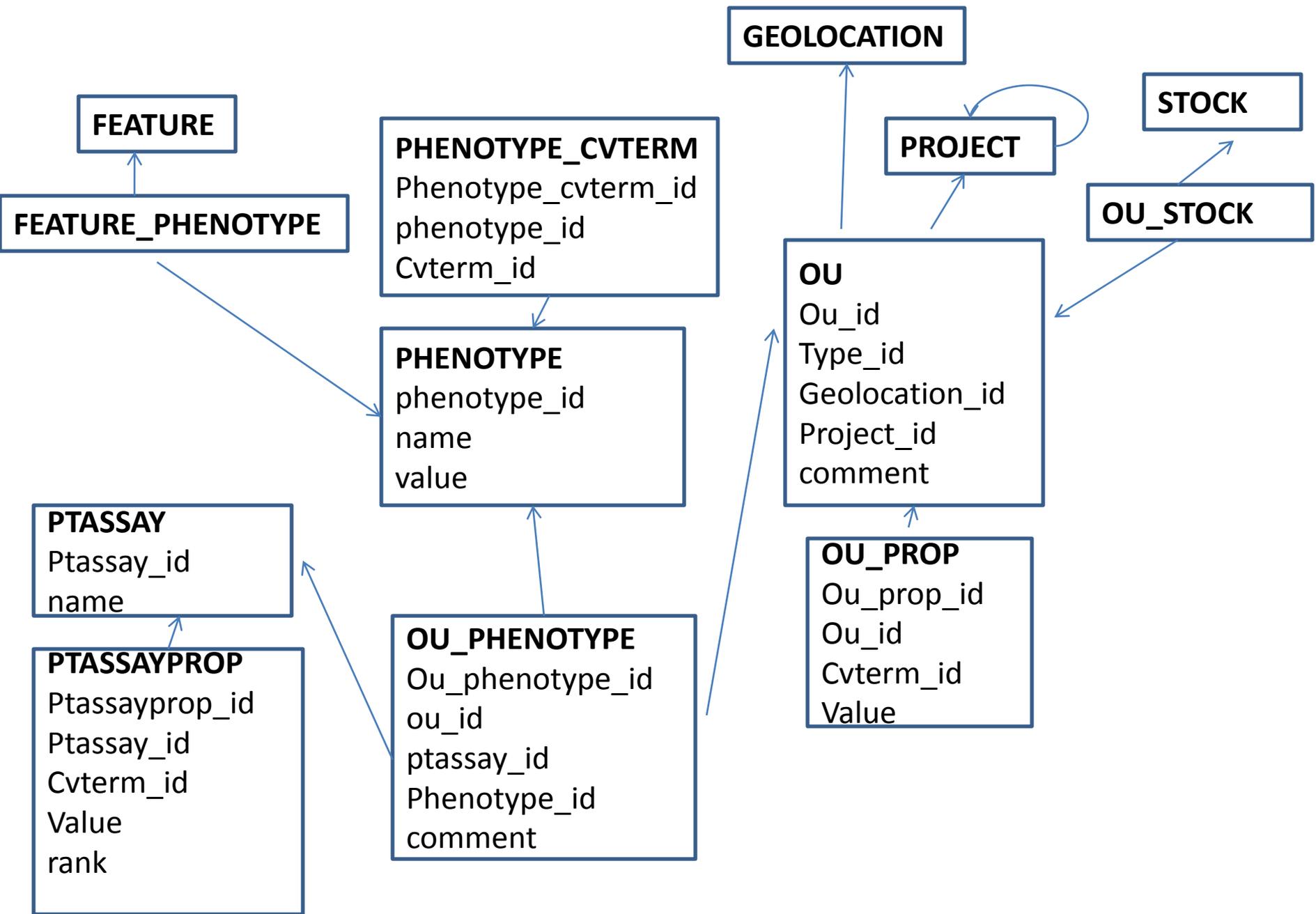
General info on QTL experiment is stored in 'project' table

1. All F2s are stored in Stock and OU

2. The F2 population of 102 is stored in OU. The phenotype value will be the mean of 107 lines

The OU entries for 1 is 'stock' type and the OU entry for 2 is 'group of stocks/population'

# Genotype/phenotype diversity data (Data from GRIN)

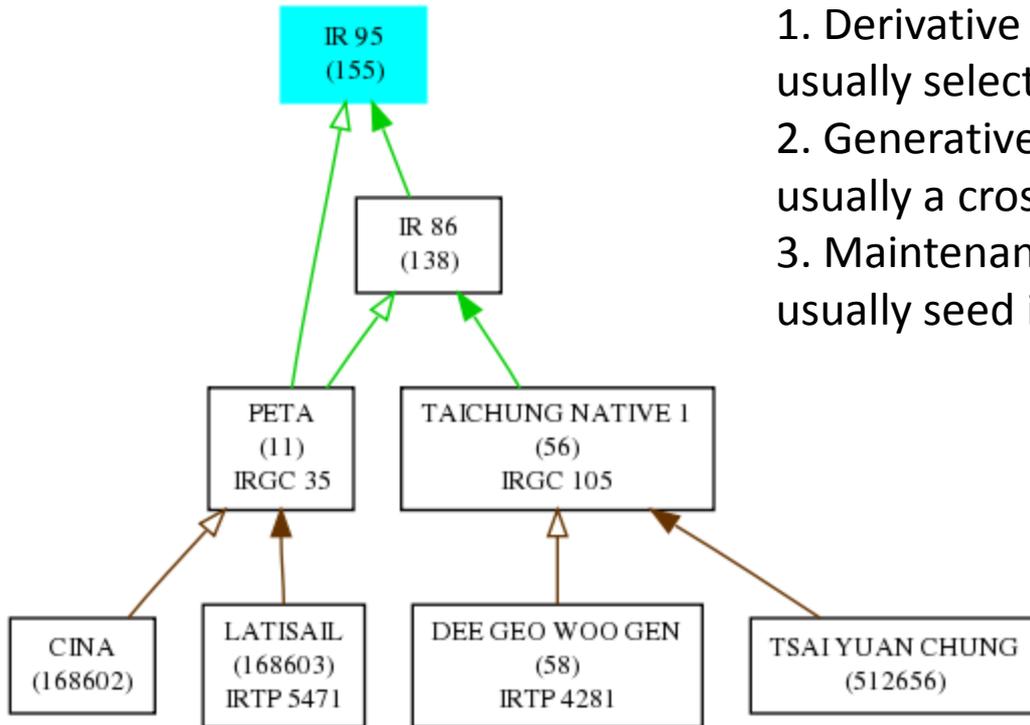


# Pedigree Data

(rice data generated from [ICIS pedigree viewer](#))

Germplasm is created by methods below  
([ppt](#), [wiki](#))

1. Derivative (to decrease genetic variation - usually selection)
2. Generative (to increase genetic variation - usually a cross)
3. Maintenance (to maintain genetic variation - usually seed increase or import)



- We need to differentiate method 2 from 1 & 3 to display pedigree
- We need to track down the original parent

This pedigree shows only parentage, empty arrow is female parent and filled arrow is male parent. Green is for generative methods and brown is for maintenance. PETA went through maintenance step(s) and CINA and LATISAIL is the parents of the cross before those maintenance step(s).

## In Germplasm Ontology

Group\_ID: Identification code for last instance of germplasm produced by a generative process (cross) from which the current instance was derived.

Source\_ID: Identification code for the immediate source of a derived or preserved germplasm. (Not valid for generated (by cross) germplasm.)

eg. When a Gala tree (Gala), generated by a cross, is imported to a WSU breeder, it will get a new germplasm\_id (Gala\_wsu), and the id of the original gala (Gala) would be group\_id. When WSU gala is imported to a MSU breeder, the source\_id would be Gala\_wsu, and group\_id would be Gala.

