Linking Genes to Node Development in Common Bean (*Phaseolus vulgaris*)

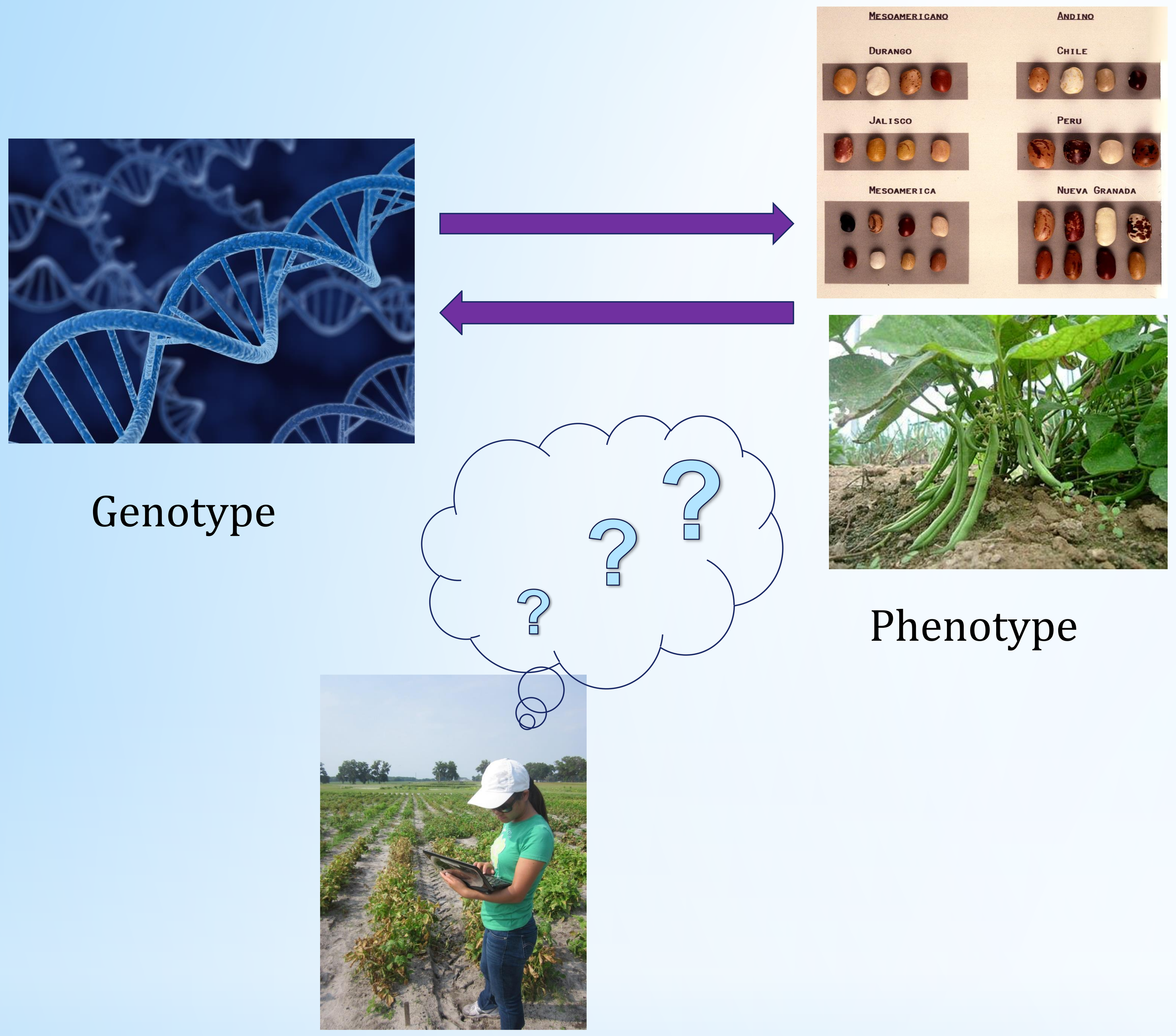
Li Zhang[1], Jim W. Jones[1], C. Eduardo Vallejos[2], Kenneth J. Boote[3], Senthold Asseng[1], Wei Hou[4], Jose A. Clavijo[3], Mehul Bhakta[2], Melanie Correll[1]

[1] Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL, USA [2] Department of Horticultural Sciences, University of Florida, Gainesville, FL, USA

[3] Department of Agronomy, University of Florida, Gainesville, FL, USA [4] Department of Preventive Medicine, Stony Brook University Medical center, Stony Brook, NY, USA

UF IFAS UNIVERSITY of FLORIDA

Background and Challenge



Goal

To identify Quantitative Trait Loci (QTL) controlling node development under a range of environments, and to incorporate this genetic information for node development.

Hypothesis and Model Description

Hypothesis : node development during the vegetative phase is only a function of temperature (Jones et al., 1999)

Model:

$$dN/dt = Nm \bullet f(T)$$

dN/dt, node addition/appearance rate (node per physiological day (**PD**));
Nm, maximum node appearance rate under optimum condition (node PD^{-1});
PD, Physiological Day, normalized form of thermal time accumulated per calendar day, *PD = f(T) × f(PP), f(PP)* is photoperiod function and equal to 1 in our model based on our hypothesis, therefore, daily *PD = f(T);*
f(T), cardinal temperature function used in most crop model (see **Results**):

T_H, hourly temperature;
T_{base}, base temperature below which no growth occurs;
T_{opt1}, 1st optimum temperature above which growth reaches maximum.

Reference

Jones, J., Kenig, A., & Vallejos, C. (1999). Reduced state-variable tomato growth model. Transactions of the ASAE-American Society of Agricultural Engineers, 42(1), 255-266.

Wang S., C.J. Basten, and Z.B. Zeng. 2012. Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm)

Wu, R. L., and Lin, M. (2006). Functional mapping-how to map and study the genetic architecture of dynamic complex traits. *Nat. Rev. Genet. 7*, 229-237.

Materials and Methods

Field Experiments

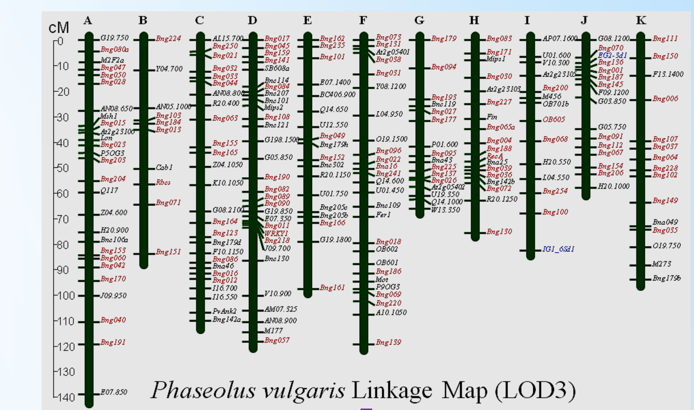
	Experiment 1	Experiment 2	Experiment 3	Experiment 4	Experiment 5
Location	Citra, Florida, USA (CT) †	Palmira, Colombia (PA) †	Isabella, Puerto Rico (PR) †	Popayan, Colombia (PO) †	Prosper, North Dakota, USA (ND) †
Latitude, Longitude	29° 39' N, 82° 06' W	03° 29' N, 76° 81' W	18° 28' N, 61° 02' W	02° 25' N, 76° 62' W	47° 00' N, 96° 47' W
Growing season	Mar, 2011-Jun, 2011	Nov, 2011-Jan, 2012	Feb, 2012-May, 2012	Mar, 2012-Jun, 2012	May, 2012-Aug, 2012
Previous culture	Fallow	Beans	Beans	Fallow	Wheat
Soil texture	Sand	Clay	Clayey Kaolinite	Medium Loam	Silt Clay Loam
Fertilization ($kg\ ha^{-1}$)	136 (N); 60 (P);112 (K)	40 (Urea)	NA	96 (P); 129 (N); 80.3 (K);	No fertilizer
Irrigation	Central Pivot Sprinkler System	Rain fed	Drip	Rain fed	Rain fed
Plant density ($plants/m^2$)	4.3	3.0	3.9	4.3	3.3
Row spacing (cm)	90	120	100	90	NA
# of replicates/harvest	3	3	3	3	3
Total # of genotypes‡	188	174	128	178	176
Frequency of Measurement	weekly	weekly	weekly	weekly	weekly

Parameter Estimation

- According to the hypothesis that temperature is the only factor affecting node addition, node appearance rate for a specific genotype should be the same or similar across locations if calendar days are adjusted to physiological day by choosing proper T_{base} and T_{opt1}.
- The default cardinal temperature function (T_{base}= 5° C, T_{opt1}=27° C) during vegetative phase in CROPGRO-Dry Bean model embedded in DSSAT software package were tested.
- A new set of T_{base}, T_{opt1}, and *Nm* were estimated based on the model using Levenberg-Marquardt nonlinear least square (*nls.lm*) function from minpack.lm package in R (version 2.14.2).
- Levenberg-Marqurdt optimization algorithm searches for the best fits of T_{base} and T_{opt1}, and *Nm* across locations for a specific genotype by minimizing the sum square of the estimates with the observations from each location.

Genotypic Analysis

High Density Maker Map (2013)



Phenotypic data: Nodes on Main Stem

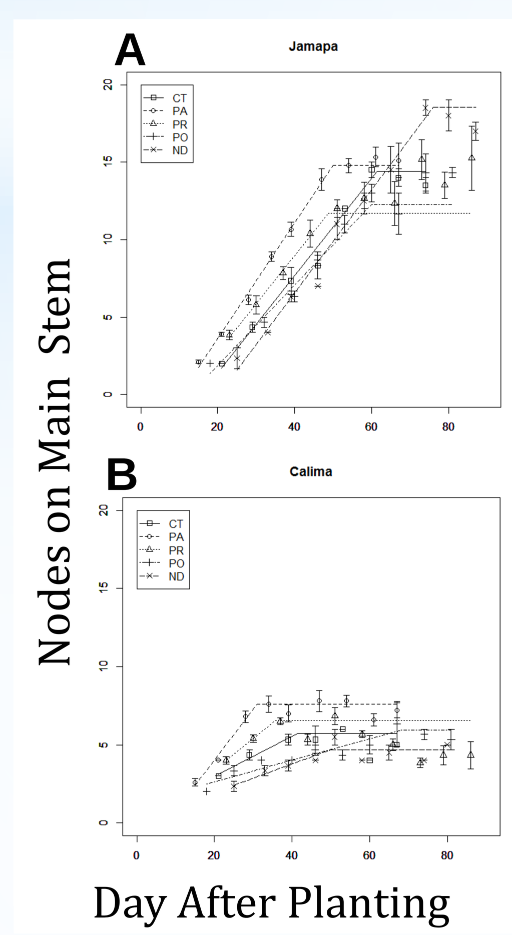
Traditional QTL Mapping

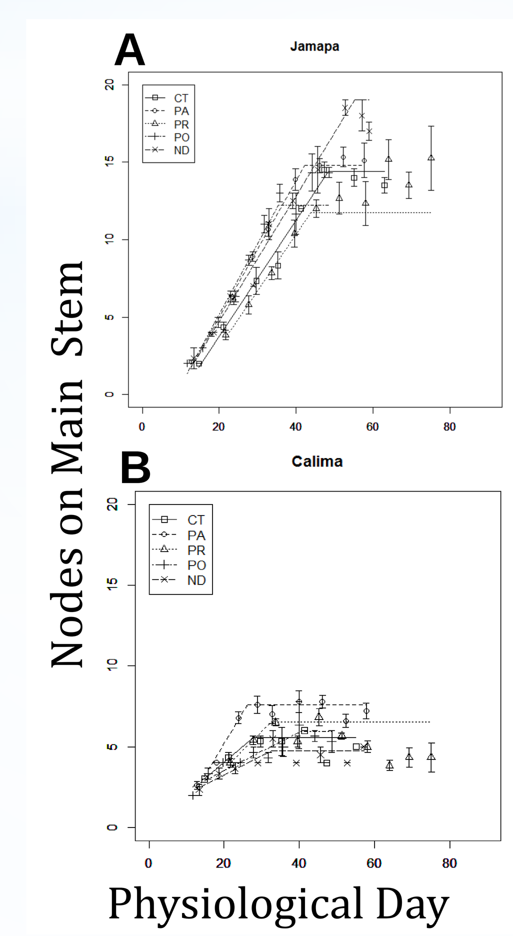
- Composite Interval Mapping (CIM);
- Multiple Interval Mapping (MIM);
- Deal with end-point traits

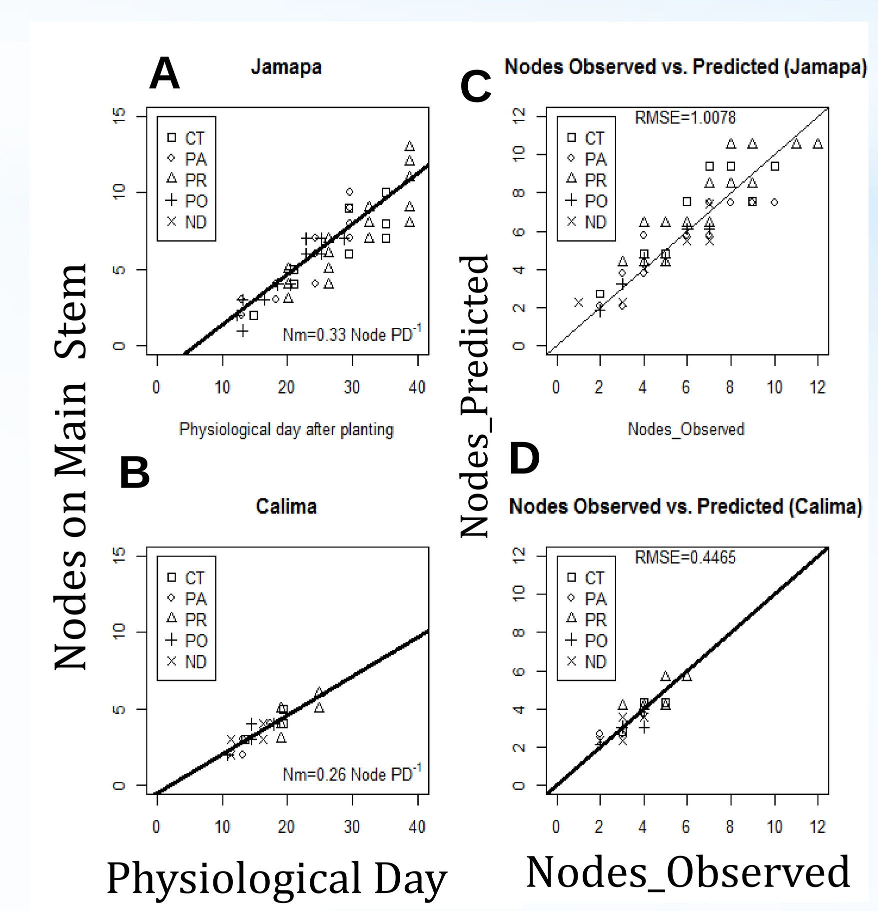
Functional Mapping

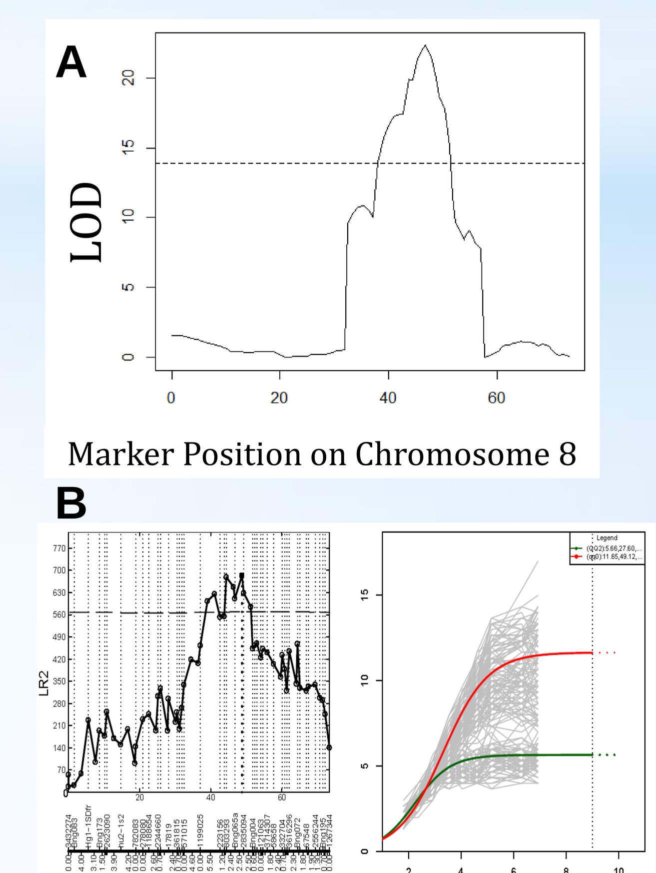
- Statistical mapping framework;
- Deal with dynamic traits

Results









Gene-based model

$$dN/dt = [Nm\_Det + (Nm\_Indet - Nm\_Det) * fin] \bullet f(T)$$

$$fin = \begin{cases} 0, & \text{if a given genotype is determinate;} \\ 1, & \text{if a given genotype is indeterminate;} \end{cases}$$

$$f(T) = \begin{cases} 0, & \text{if } T_H < Tbase \\ \sum_{H=1}^{24} \frac{T_H - Tbase}{24 * (Topt1 - Tbase)}, & \text{if } Tbase <= T_H < Topt1 \\ 1/24, & \text{if } T_H >= Topt1 \end{cases}$$

T_{base} = 9.3 ° C
T_{opt1} = 22.2 ° C

Step 1:

- Nodes on Main Stem vs. **Calendar day**;
- Significant difference among some locations($P < 0.05$) for most of genotypes ;
- Temperature plays an important role in node development.

Step 2:

- Nodes on Main Stem vs. **Physiological Day** (PD) using default T_{base} **(5° C)** and T_{opt1} **(27° C)** from CROPGRO-Bean;
- Significant difference among some locations for most of genotypes;
- T_{base} and T_{opt1} need to be estimated for our population.

Step 3:

- **Parameter estimation** using nls.lm function;
- Using newly estimated T_{base} and T_{opt1} for each genotype, nodes on main stem vs. **PD** from different locations can be lined-up with the same node appearance rate (Nm);
- The average T_{base} and T_{opt1} are **9.3** and **22.2° C** respectively for our population.

Step 4:

- Traditional QTL mapping (**CIM & MIM**) peak (above) for Nm;
- Functional Mapping (**FM**) peak (below) for nodes on main stem (N_m & N_{max});
- The same QTL region on **chromosome 8** where ***fin*** gene is occupied was identified.
- Gene ***fin*** is responsible for **node development.**

Step 5:

- ***Nm_Det*** and ***Nm_Indet*** represent maximum node appearance rate (node PD^{-1}) under optimum condition for indeterminate and determinate genotypes respectively;
- The average *Nm_Det* and *Nm_Indet* are **0.32** and **0.25** node PD^{-1} in our population.

NSF # 93903