IPK GATERSLEBEN

Special Symposium to Honour the Life´s Work of Ulrich Wobus

Using Genomics to Dissect Seed Development

BOB GOLDBERG

6/27/08

Today's Headlines

The New York Times

A Global Need for Grain That Farms Can't Fill

Published: March 9, 2008

Los Angeles Times

Economist.com

High Rice Cost Creating Fears of Asia Unrest

By KEITH BRADSHER
Published: March 29, 2008

U.S.News & WORLD REPORT

CNN.com

THE FOOD CHAIN

A Drought in Australia, a Global Shortage of Rice

Across Globe, Empty Bellies Bring Rising Anger

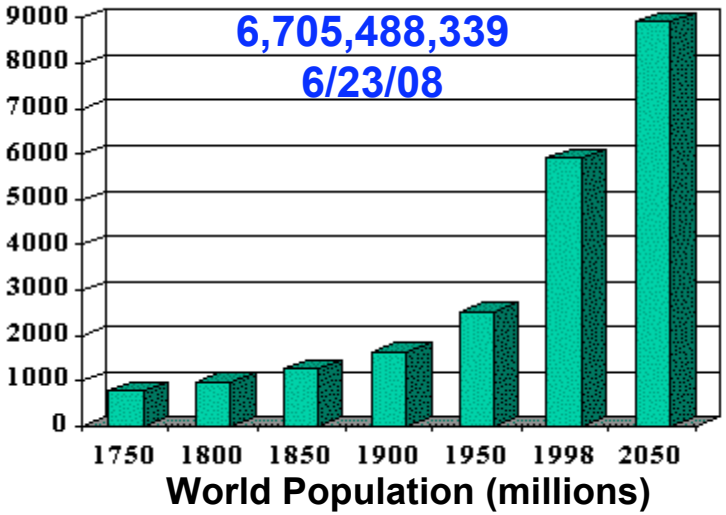
Newsweek

updated 10:42 p.m. EDT, Mon April 14, 2008

Riots, instability spread as food prices skyrocket

The Washington Post

We Face Major Challenges In Agriculture Even Greater Than Those in Today's Headlines

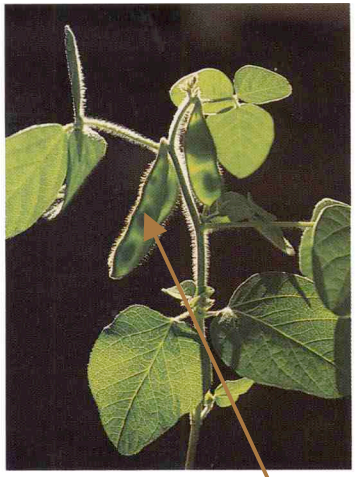


OVER 50 YEARS WE WILL NEED TO PRODUCE MORE FOOD THAN IN THE WHOLE OF HUMAN HISTORY -- AND DO IT ON THE SAME (or less) AMOUNT OF ARABLE LAND!!!!

Aerial Photograph of UCLA in 1929


Aerial Photograph of UCLA in 2008
Bel-Air
Beverly Hills
Former Farms
Sunset Blvd.
Hilgard Blvd.
Original Agricultural College
and Citrus/Avocado Orchard
Westwood Blvd.
Image © 2007 Sanborn
Technologies
5/21/08
Google

So….How Can Seed Yields Be Improved?
Use a Variety of Approaches To Identify Genes Critical For Fundamental Seed Processes (Yo!!-It's the Yield That Counts!)



Seeds Without Fertilization

Hybrids

Reduced Pod Shattering

Architecture Designed For Specific Growth Conditions

More Seeds

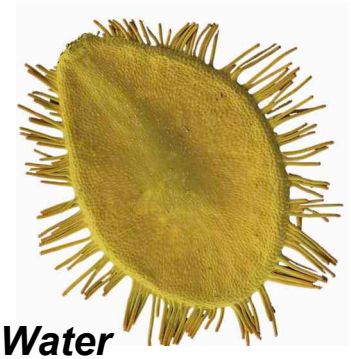
Bigger Seeds

Seeds Optimal For Human/Animal Health & Nutrition

…..and → **Re-Engineer Existing Crops** → *Increase Yields!*

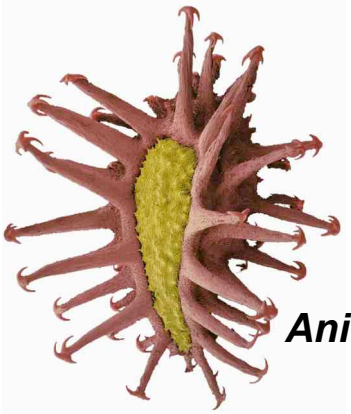
And……..There's Also A Problem With Using Land For Energy Production……

So……Why Seeds??

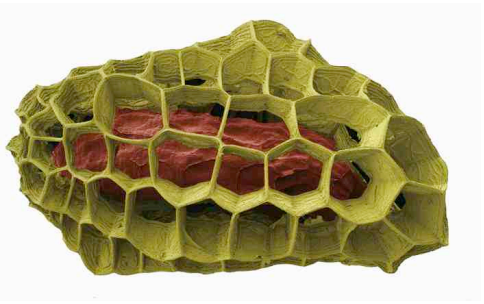


Water

Wind



Animals



Seeds Protect and Disperse Plant Embryos and Come in Many Shapes and Sizes!

Seeds Are Used in Many Ways as Food, Beverages, Spices, and Fuels!

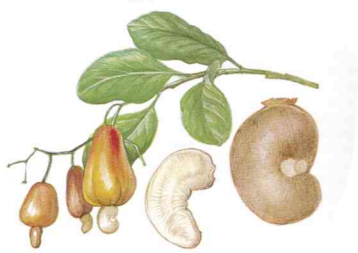
Beans

Peas

Wheat

Corn

Coconut



Cashew Nuts

Peanuts

Pecans

Cocoa Beans

Coffee Beans

Nutmeg

Mustard

Most Importantly….. Our Food is Derived From Fourteen Crops & <u>Over Half</u> Produce Seeds For Human and Animal Consumption

<u>Seed Crops</u>

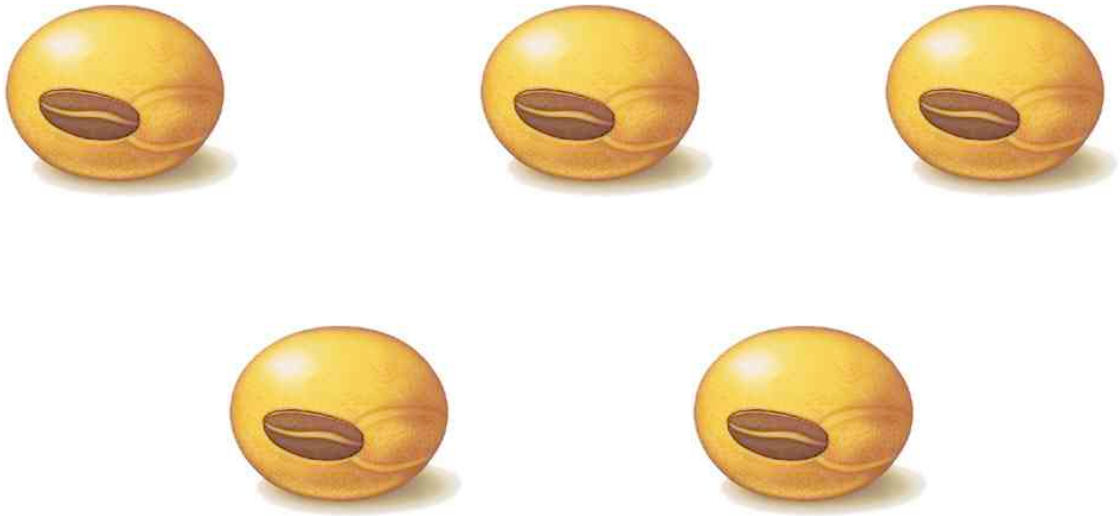
- ***Wheat***
- ***Rice***
- ***Corn***
- ***Barley***
- ***Sorghum***
- ***Soybean***
- ***Common Bean***
- ***Coconut***

<u>Non-Seed Crops</u>

- ***Potato***
- ***Sweet Potato***
- ***Cassava***
- ***Sugar Beet***
- ***Sugar Cane***
- ***Banana***

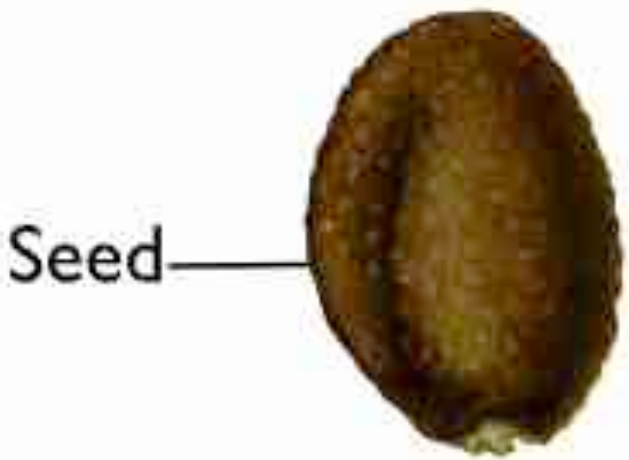
In Some World Populations 75% of Calories Are Derived From Seeds!

How Is a Seed Formed?

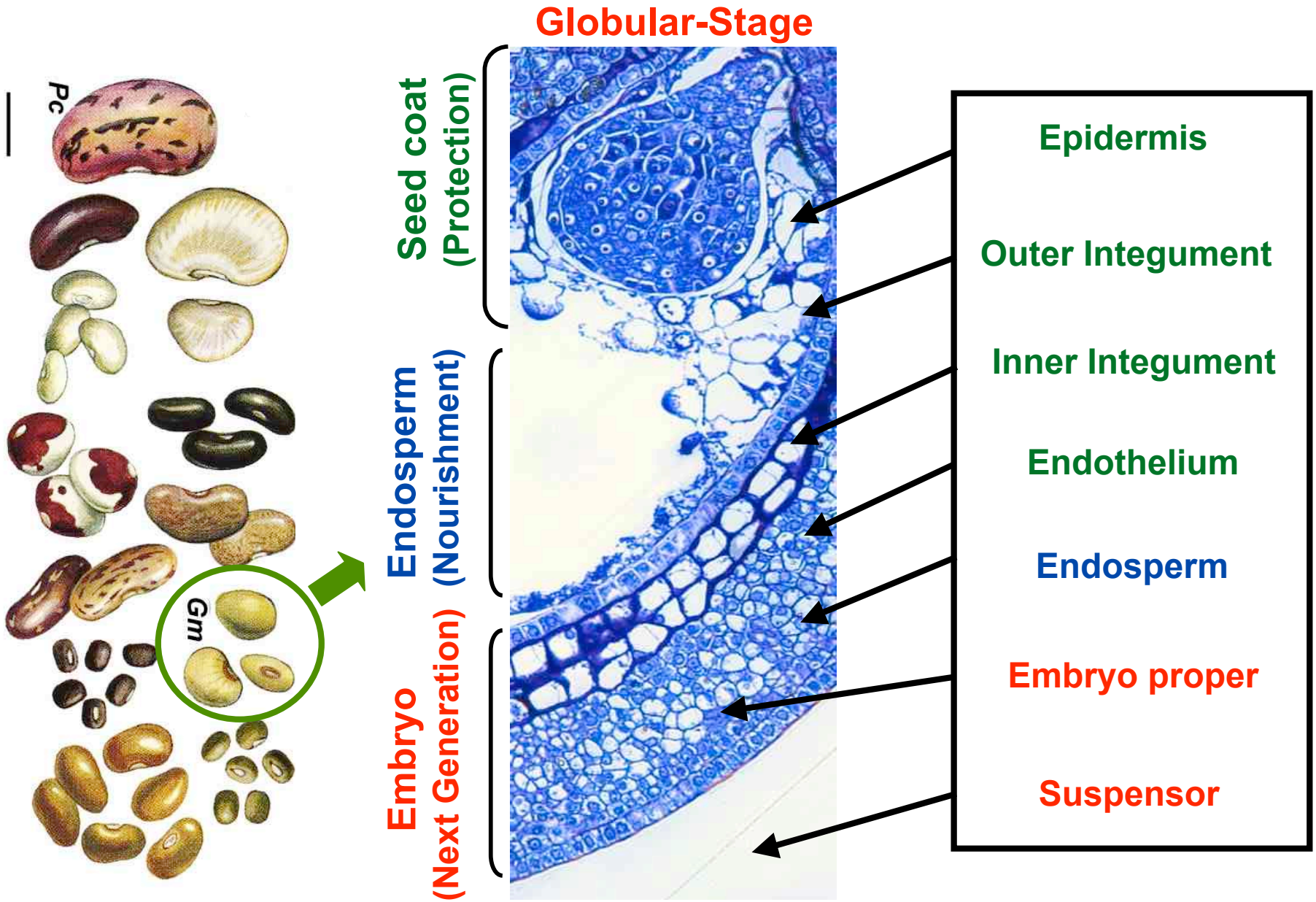


Remember……. a Seed Contains the Mature, Dormant Embryo
That is the Next Plant Generation

In the Beginning....

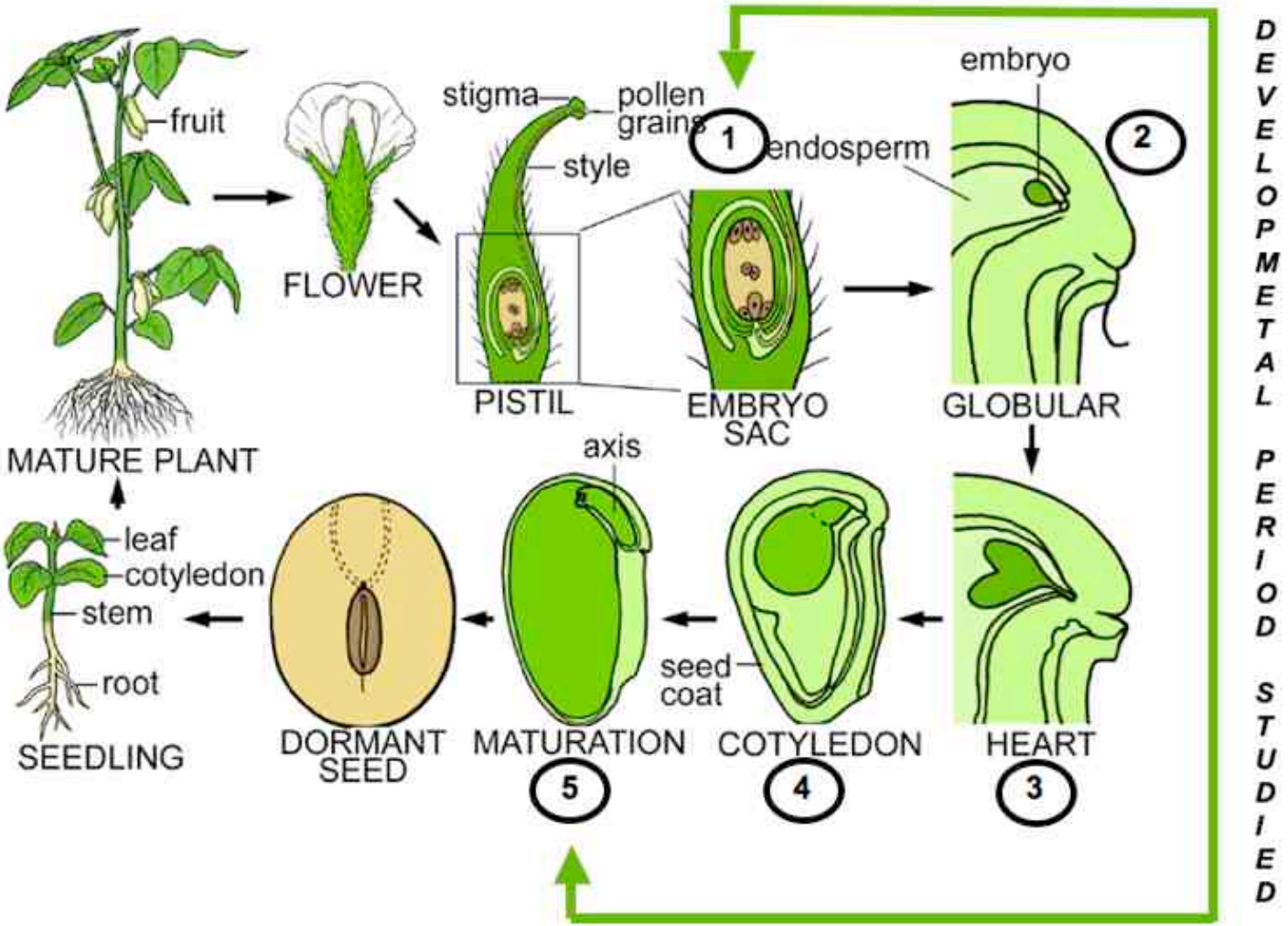


What Are the Genes Required to Make a Seed?

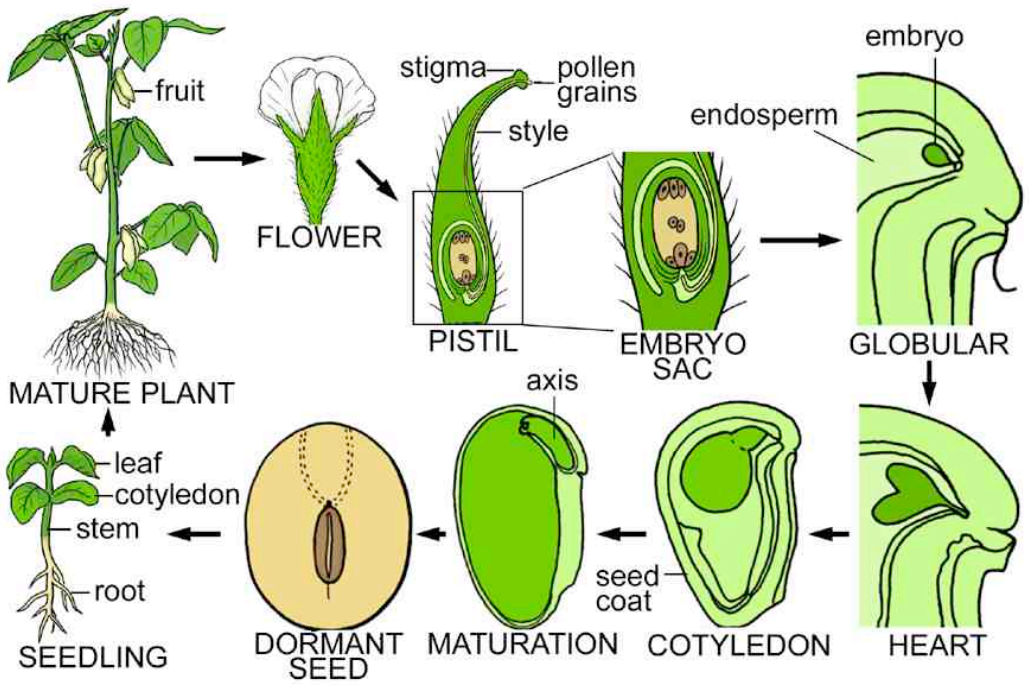


And How Are They Wired in a Plant Genome?

More Specifically……What Are The Genes Required to Program <u>Every</u> Compartment, Tissue, and Cell Type During Seed Development?

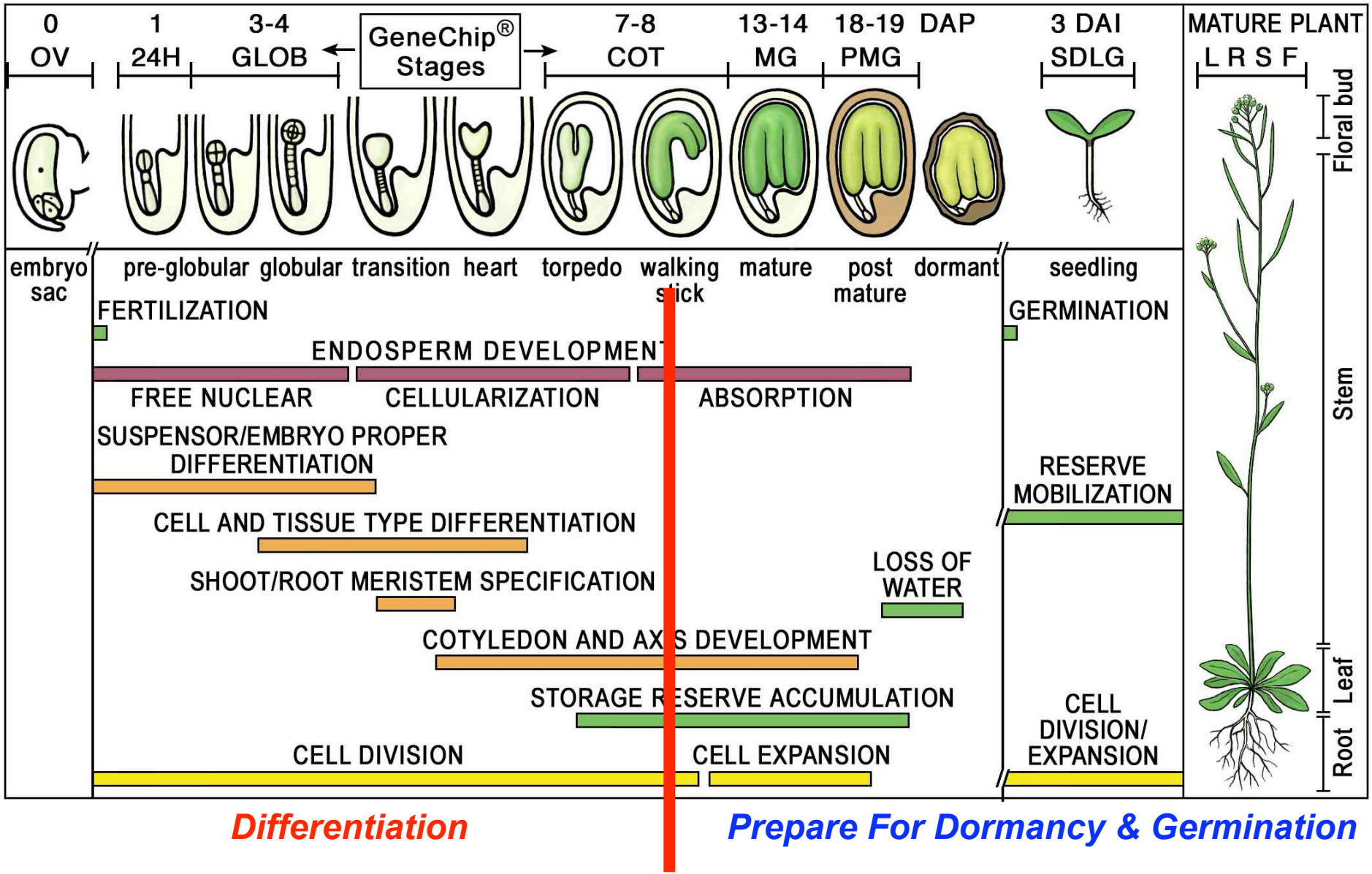


Major Seed Biology Questions Discussed Today



- ***What Is the Nature of Seed-Specific Regulators?***
- ***How Many Genes Required to Make an Entire Seed?***
- ***What DNA Sequences Are Required For Seed Region-Specific Transcription?***

Genome-Wide Profiling of mRNAs During the Entire Arabidopsis Life Cycle



Gene Activity Before, During, And After Arabidopsis Seed Development

SEED DEVELOPMENT

	OV	24H	GLOB	COT	MG	PMG	SDLG
Total mRNAs	12,591	12,421	13,722	13,103	10,875	8,779	13,185
TF mRNAs	999	995	1,089	1,051	851	699	1,016
Unique mRNAs	22	16	100	50	26	31	505
Unique TFs	4	0	17	9	4	6	57
Shared mRNAs (TFs)	6,937 (477)	→					

http://estdb.biology.ucla.edu/genechip

Gene Activity Before, During, And After Arabidopsis Seed Development

SEED DEVELOPMENT

	OV	24H	GLOB	COT	MG	PMG	SDLG
Total mRNAs	12,591	12,421	13,722	13,103	10,875	8,779	13,185
TF mRNAs	999	995	1,089	1,051	851	699	1,016
Unique mRNAs	**22**	**16**	**100**	**50**	**26**	**31**	**505**
Unique TFs	**4**	**0**	**17**	**9**	**4**	**6**	**57**
Shared mRNAs (TFs)	6,937 (477)	→					

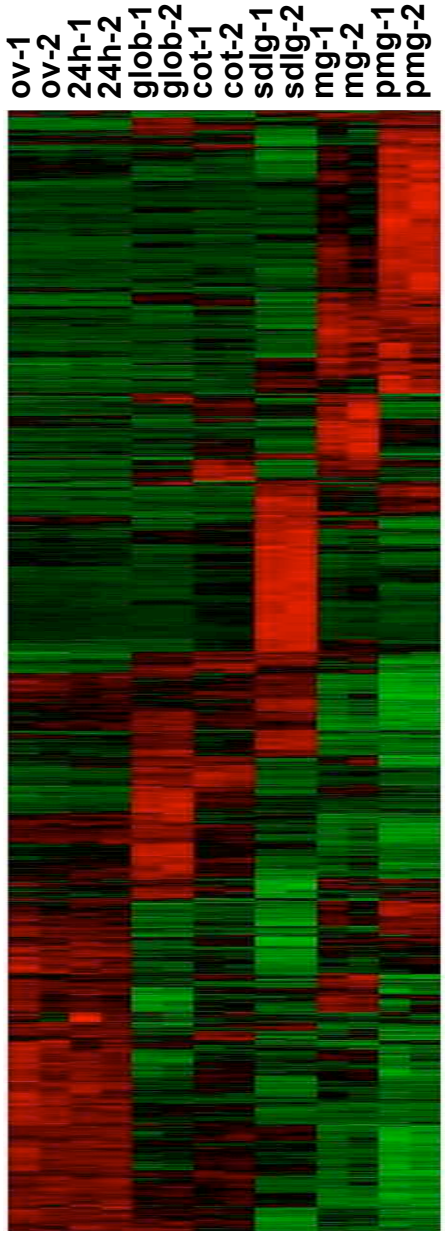
http://estdb.biology.ucla.edu/genechip

Gene Activity Before, During, And After Arabidopsis Seed Development

SEED DEVELOPMENT

	OV	24H	GLOB	COT	MG	PMG	SDLG
Total mRNAs	12,591	12,421	13,722	13,103	10,875	8,779	13,185
TF mRNAs	999	995	1,089	1,051	851	699	1,016
Unique mRNAs	22	16	100	50	26	31	505
Unique TFs	4	0	17	9	4	6	57
Shared mRNAs (TFs)	**6,937 (477)**	→					

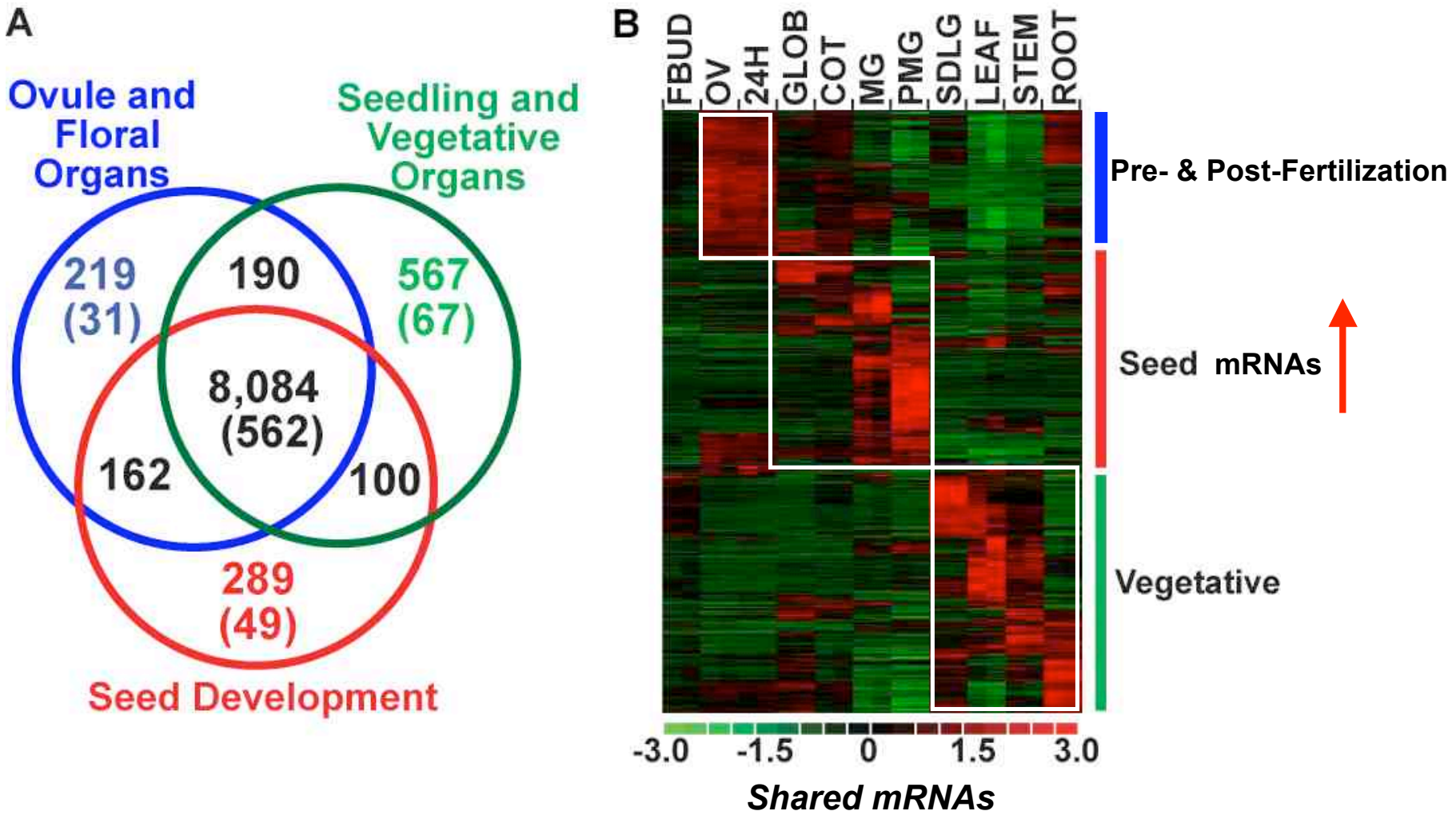
http://estdb.biology.ucla.edu/genechip



Shared mRNAs


Are There Seed-Specific Genes That May Play a Critical Role in Programming Seed Development?

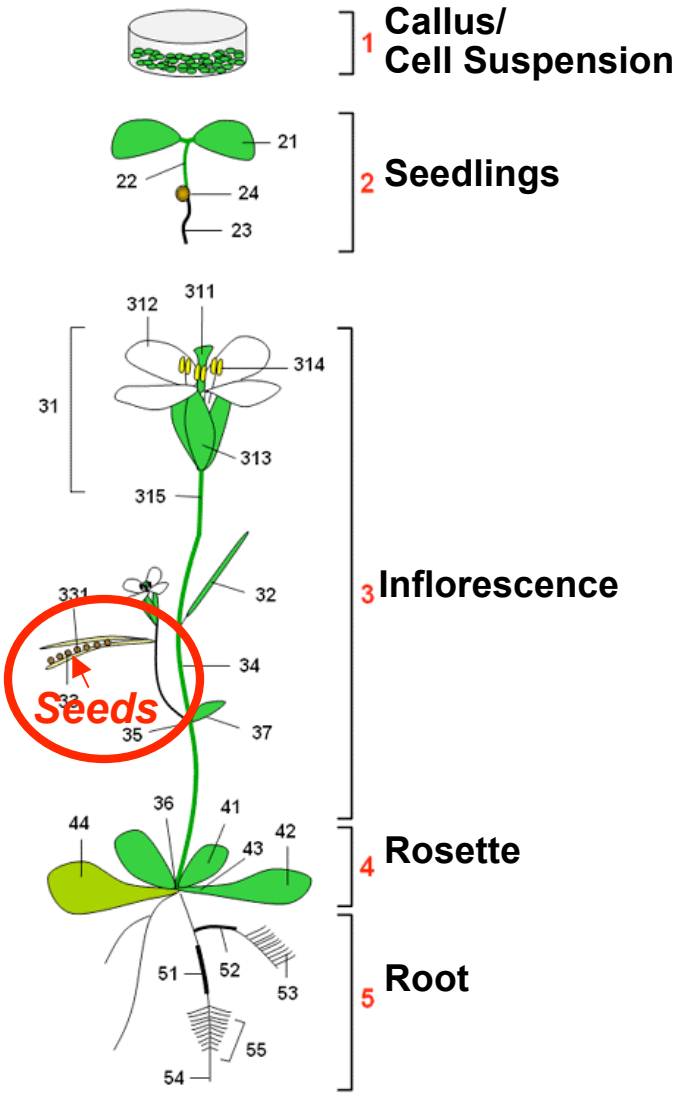

Identification of Seed-Specific mRNAs in the Arabidopsis Life Cycle

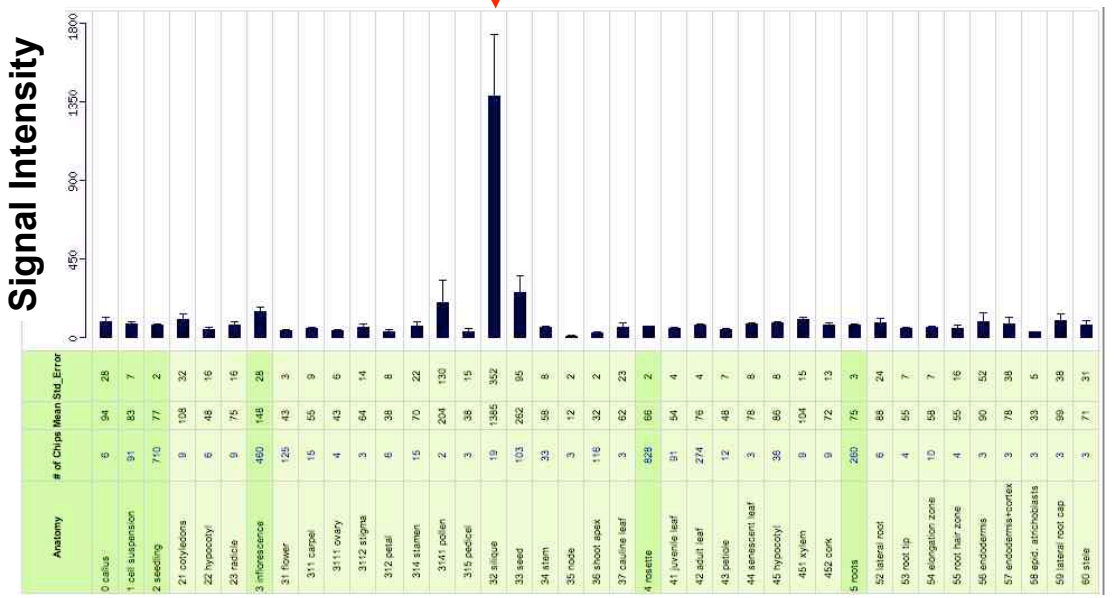


…at the GeneChip Level!!

( ) Indicates number of transcription factor mRNAs

Validation of "Seed-Specific" Genes Using Public Microarray Data





>2,500 Sets of Public Microarray Data

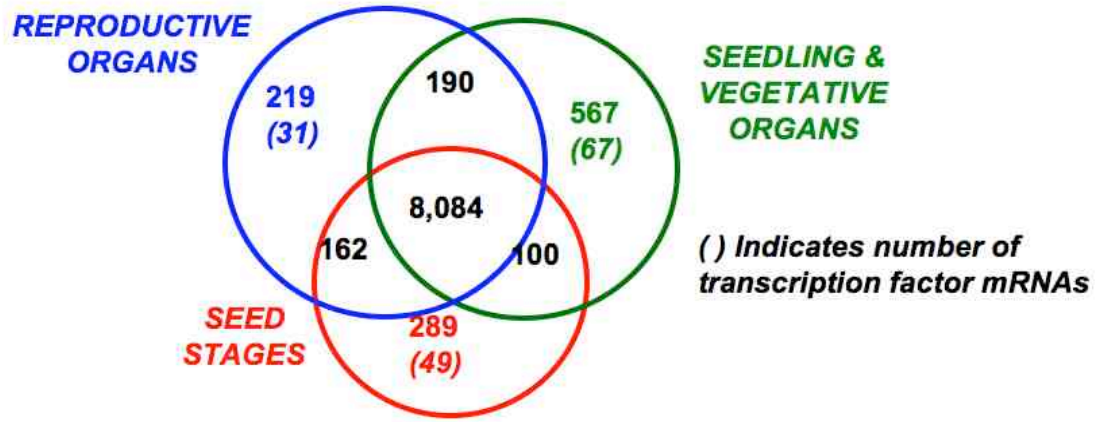
http://www.genevestigator.ethz.ch/at

Majority of Seed-Specific Genes are Seed-Specific Compared to All Available Public Microarray Data Set!!

Identification of "Seed-Specific" Transcription Factor mRNAs

		ABI3-VP1
(G,C)	*	LEC2
(G,C,M,PM)	*	FUS3
(C,M)		B3 Family TF
		AP2/EREBP
(PM)	*	AP2 Domain Protein
(C,M,PM)		Aintegumenta-Like 7
		ARF
(G)	*	ARF21
		ARR-B
(G,C)		ARR19
(C)	*	ARR21 / ARR13
(24H,G,C,M,PM)	*	ARR22
		AS2
(C,M)		LOB Domain Protein 18
(G)		LOB Domain Protein 35
		AUX-IAA
(24H,G,C,M)		IAA31
		Basic Leucine Zipper
(C,M,PM)	*	AtbZIP67; DPBF2
(C)	*	AtbZIP72
(G,C)		AtbZIP15
		CCAAT-Box
(G,C)	*	LEC1
(G,C,M)	*	L1L
(G)		HAP5A-Like
		Heat-Shock
(G,C)	*	HSF1-Like
(PM)	*	Heat Shock TF

** Mutation in these genes affects or disrupts embryo/seed development (15)*



		Homeobox
(G,C,M)		Homeodomain Protein
(G,C)	*	ATML1-Like
(G,C)	*	Homeodomain Protein
		MADS-Box
(M,PM)	*	MADS-Box TF
(M,PM)	*	MADS-Box TF
(G,C)		AGL33
(G)		AGL35
(G)		AGL36
(G)		AGL45
(G)		AGL57
(G)		AGL91
		MYB
(M,PM)		AtMYB107
(C)	*	Myb-Related Protein
(PM)	*	AtMYB67; AtY53
(G,C)	*	Myb-Related Protein

** Mutation in these genes affects or disrupts embryo/seed development (7)*

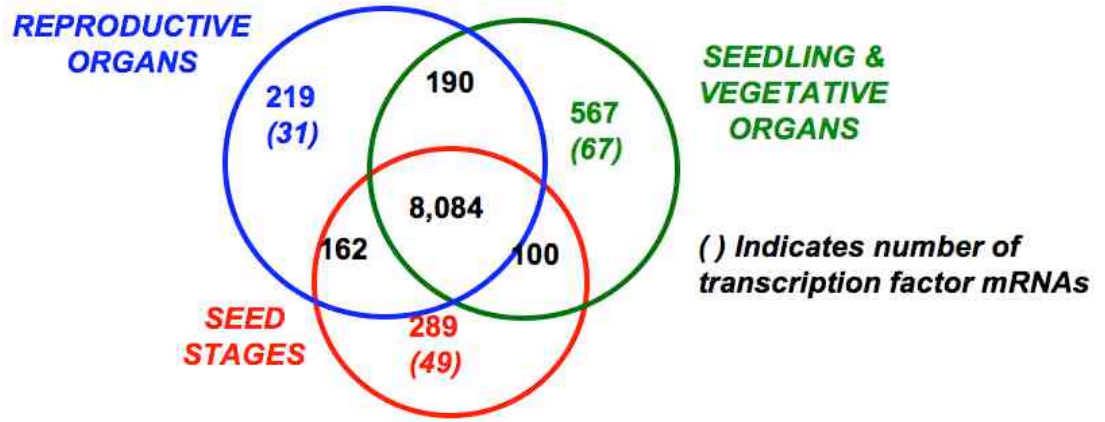
(G,C)		Myb-Related Protein
		NAC Domain
(C)		No Apical Meristem (NAM)
		Polycomb Group
(M)	*	MEDEA
(PM)		CCHC-Type Family Protein
		WRKY
(G)	*	WRKY10; MINI3
		Zinc Finger
(C,M,PM)	*	PEI1
(G)		C2H2-Type Zinc Finger
(G,C)		C2H2-Type Zinc Finger
(G,C)		C3HC4-Type RING Finger
(G)		C3HC4-Type RING Finger
(G,C)		Dof-Type Zinc Finger
(24H,G)		RABBIT EARS (RBE)
		Unclassified
(C)		SAP; Sterile Apetala

Seed Stages: 24H, 24Hr Post-Pollination; G, Globular; C, Cotyledon; M, Mature Green; PM, Post-mature Green

Identification of "Seed-Specific" Transcription Factor mRNAs

		ABI3-VP1
(G,C)	*	LEC2
(G,C,M,PM)	*	FUS3
(C,M)		B3 Family TF
		AP2/EREBP
(PM)	*	AP2 Domain Protein
(C,M,PM)		Aintegumenta-Like 7
		ARF
(G)	*	ARF21
		ARR-B
(G,C)		ARR19
(C)	*	ARR21 / ARR13
(24H,G,C,M,PM)	*	ARR22
		AS2
(C,M)		LOB Domain Protein 18
(G)		LOB Domain Protein 35
		AUX-IAA
(24H,G,C,M)		IAA31
		Basic Leucine Zipper
(C,M,PM)	*	AtbZIP67; DPBF2
(C)	*	AtbZIP72
(G,C)		AtbZIP15
		CCAAT-Box
(G,C)	*	**LEC1**
(G,C,M)	*	**L1L**
(G)		HAP5A-Like
		Heat-Shock
(G,C)	*	HSF1-Like
(PM)	*	Heat Shock TF

** Mutation in these genes affects or disrupts embryo/seed development (15)*



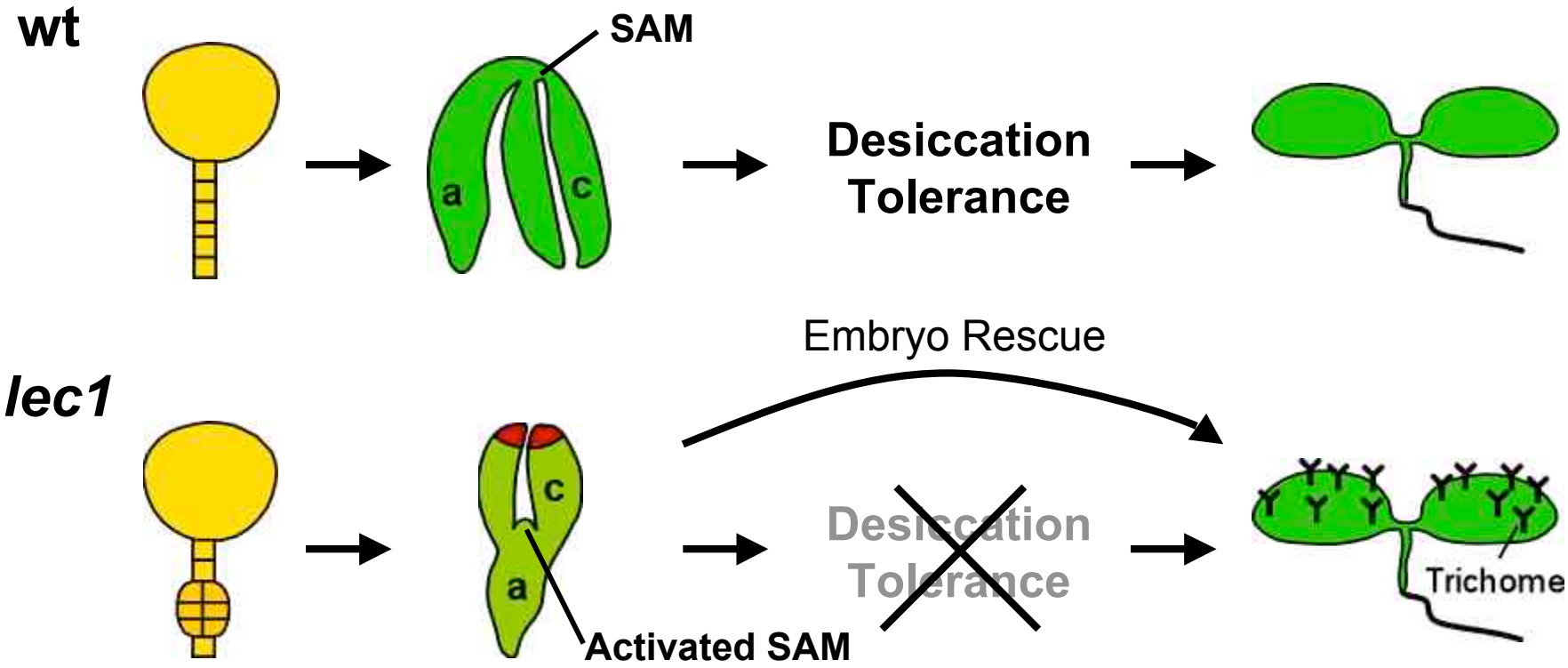
		Homeobox
(G,C,M)		Homeodomain Protein
(G,C)	*	ATML1-Like
(G,C)	*	Homeodomain Protein
		MADS-Box
(M,PM)	*	MADS-Box TF
(M,PM)	*	MADS-Box TF
(G,C)		AGL33
(G)		AGL35
(G)		AGL36
(G)		AGL45
(G)		AGL57
(G)		AGL91
		MYB
(M,PM)		AtMYB107
(C)	*	Myb-Related Protein
(PM)	*	AtMYB67; AtY53
(G,C)	*	Myb-Related Protein

** Mutation in these genes affects or disrupts embryo/seed development (7)*

(G,C)		Myb-Related Protein
		NAC Domain
(C)		No Apical Meristem (NAM)
		Polycomb Group
(M)	*	MEDEA
(PM)		CCHC-Type Family Protein
		WRKY
(G)	*	WRKY10; MINI3
		Zinc Finger
(C,M,PM)	*	PEI1
(G)		C2H2-Type Zinc Finger
(G,C)		C2H2-Type Zinc Finger
(G,C)		C3HC4-Type RING Finger
(G)		C3HC4-Type RING Finger
(G,C)		Dof-Type Zinc Finger
(24H,G)		RABBIT EARS (RBE)
		Unclassified
(C)		SAP; Sterile Apetala

Seed Stages: 24H, 24Hr Post-Pollination; G, Globular; C, Cotyledon; M, Mature Green; PM, Post-mature Green

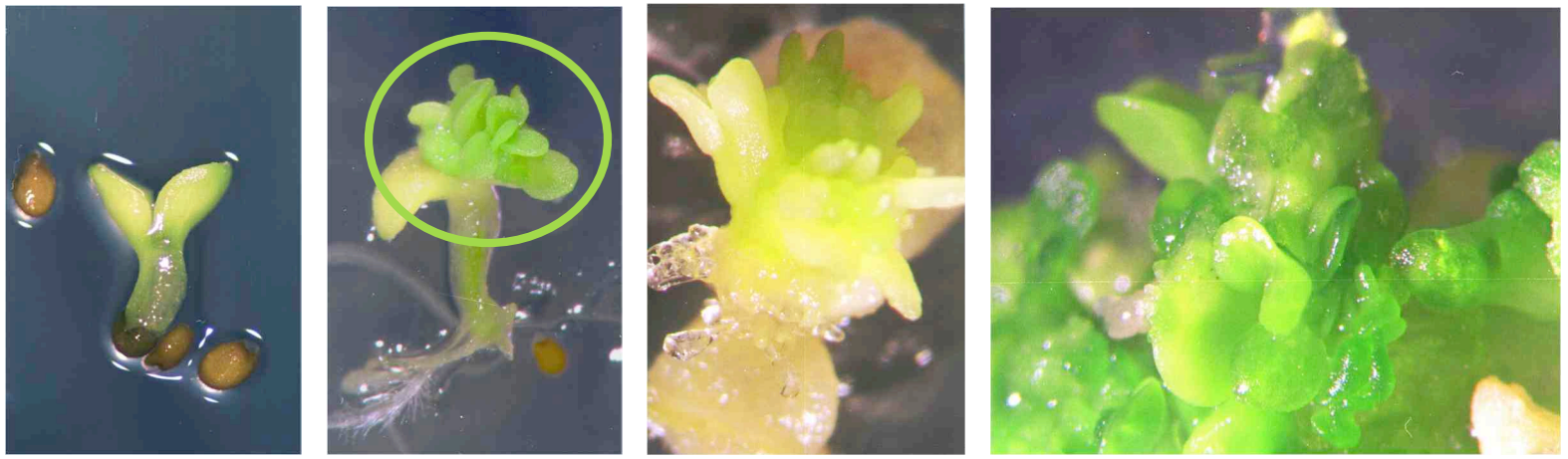
leafy cotyledon1 (lec1) Mutants Disrupt Seed Development

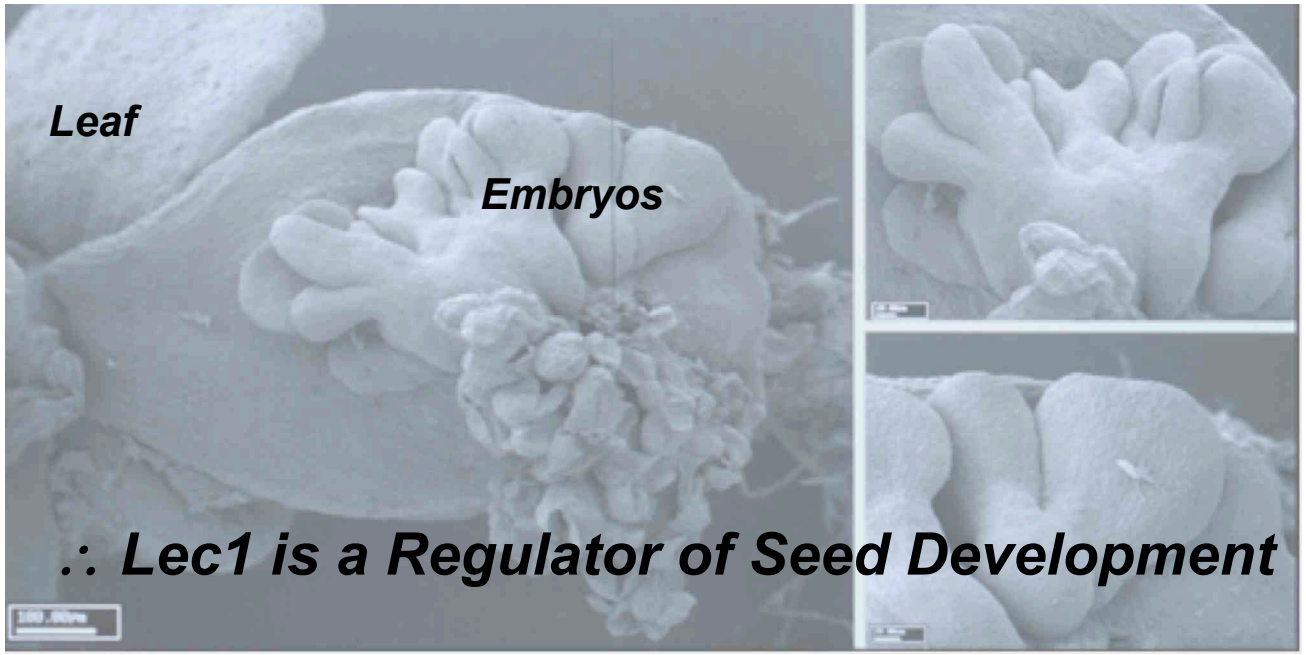


- *Suppression of Suspensor Embryonic Potential*
- *Development of Cotyledon Identity*
- *Initiation and Maintenance of Seed Maturation*
- *Inhibition Germination*

Lotan et al., Cell, 1998

Lec1 Induces Embryo Development on Engineered Leaves!!

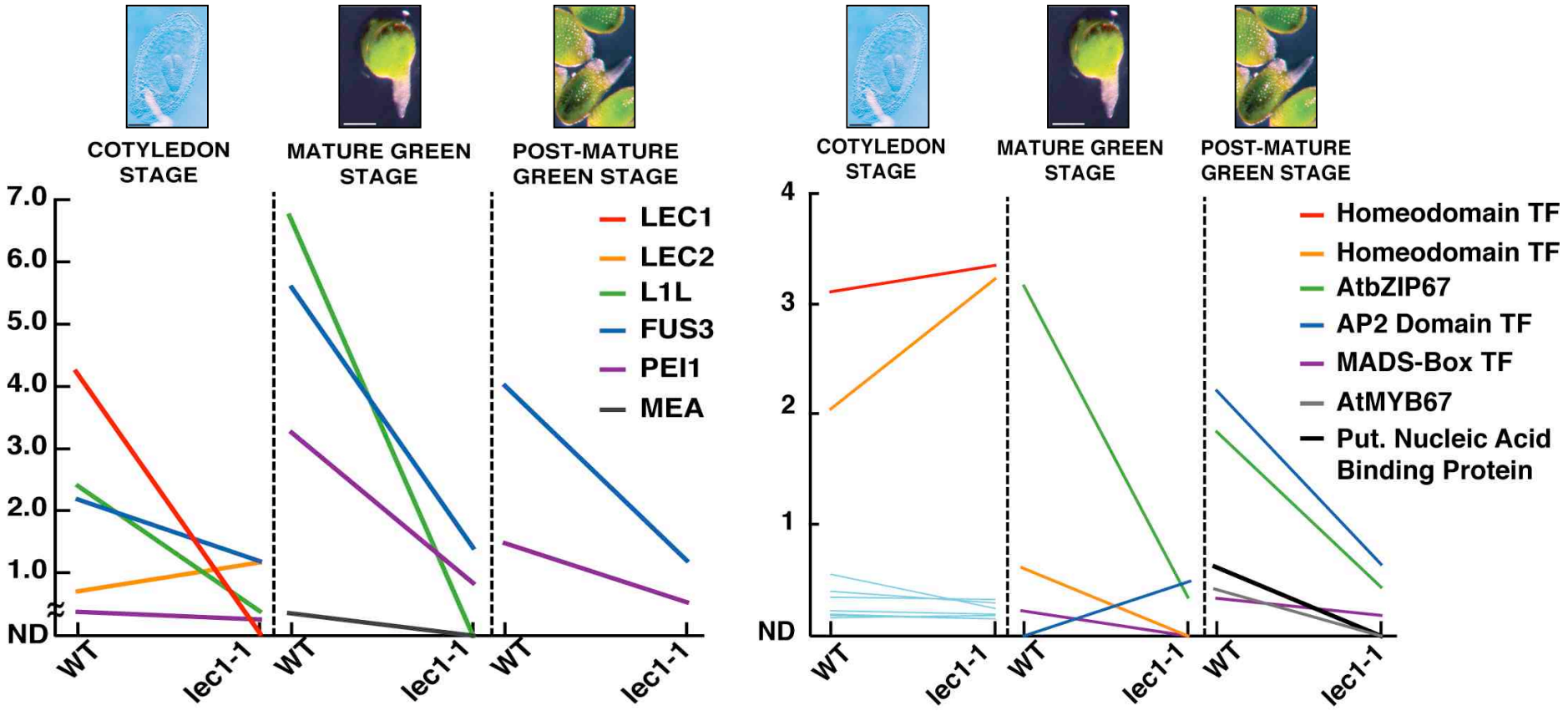




∴ Lec1 is a Regulator of Seed Development

Lotan et al., Cell, 1998; Lee et al., PNAS, 2003; Kwong et al., Plant Cell, 2003

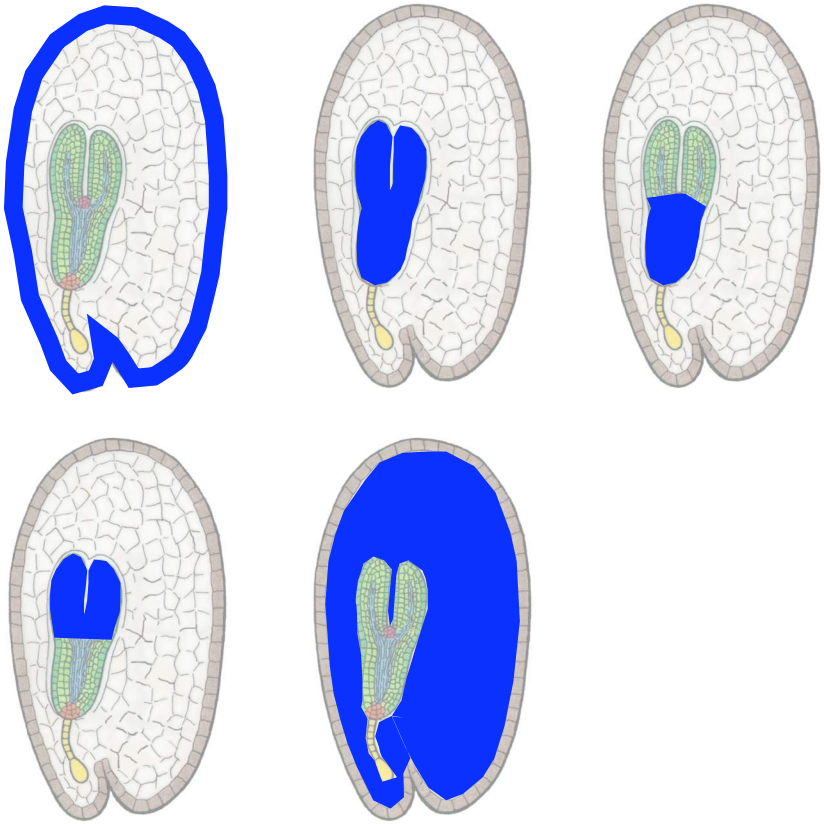
The lec1 Mutation Disrupts the Activity of "Seed-Specific" Genes



Where Are Seed-Specific Genes Active Within the Seed?

Chimeric Reporter Gene

Transform Arabidopsis

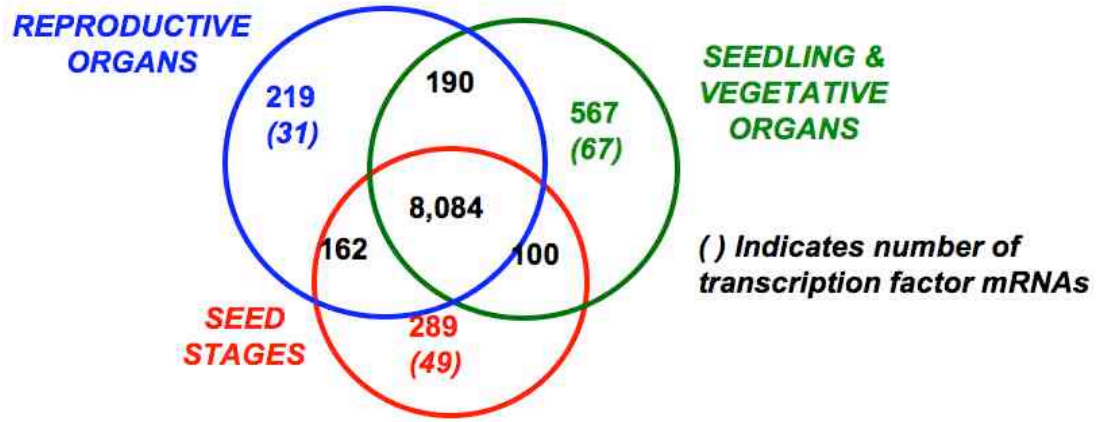


Observe Promoter Activity

Identification of "Seed-Specific" Transcription Factor mRNAs

		ABI3-VP1
(G,C)	*	LEC2
(G,C,M,PM)	*	FUS3
(C,M)		B3 Family TF
		AP2/EREBP
(PM)	*	**AP2 Domain Protein**
(C,M,PM)		Aintegumenta-Like 7
		ARF
(G)	*	ARF21
		ARR-B
(G,C)		ARR19
(C)	*	ARR21 / ARR13
(24H,G,C,M,PM)	*	ARR22
		AS2
(C,M)		LOB Domain Protein 18
(G)		LOB Domain Protein 35
		AUX-IAA
(24H,G,C,M)		IAA31
		Basic Leucine Zipper
(C,M,PM)	*	**AtbZIP67; DPBF2**
(C)	*	**AtbZIP72**
(G,C)		**AtbZIP15**
		CCAAT-Box
(G,C)	*	LEC1
(G,C,M)	*	**L1L**
(G)		HAP5A-Like
		Heat-Shock
(G,C)	*	**HSF1-Like**
(PM)	*	**Heat Shock TF**

** Mutation in these genes affects or disrupts embryo/seed development (15)*



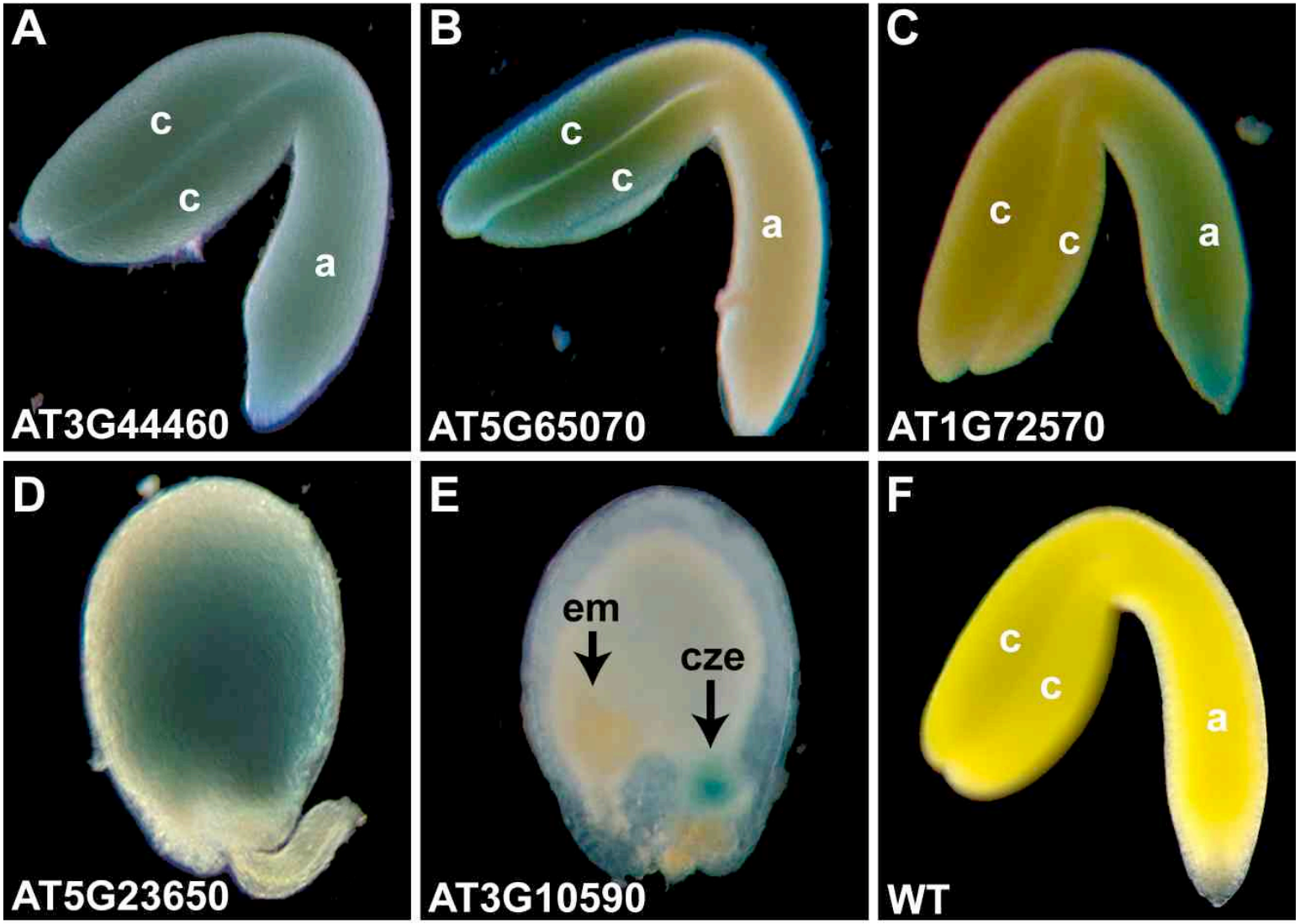
		Homeobox
(G,C,M)		**Homeodomain Protein**
(G,C)	*	**ATML1-Like**
(G,C)	*	Homeodomain Protein
		MADS-Box
(M,PM)	*	**MADS-Box TF**
(M,PM)	*	MADS-Box TF
(G,C)		**AGL33**
(G)		**AGL35**
(G)		**AGL36**
(G)		**AGL45**
(G)		AGL57
(G)		AGL91
		MYB
(M,PM)		**AtMYB107**
(C)	*	Myb-Related Protein
(PM)	*	AtMYB67; AtY53
(G,C)	*	**Myb-Related Protein**

** Mutation in these genes affects or disrupts embryo/seed development (7)*

(G,C)		**Myb-Related Protein**
		NAC Domain
(C)		No Apical Meristem (NAM)
		Polycomb Group
(M)	*	MEDEA
(PM)		**CCHC-Type Family Protein**
		WRKY
(G)	*	WRKY10; MINI3
		Zinc Finger
(C,M,PM)	*	PEI1
(G)		C2H2-Type Zinc Finger
(G,C)		C2H2-Type Zinc Finger
(G,C)		C3HC4-Type RING Finger
(G)		C3HC4-Type RING Finger
(G,C)		**Dof-Type Zinc Finger**
(24H,G)		RABBIT EARS (RBE)
		Unclassified
(C)		SAP; Sterile Apetala

Seed Stages: 24H, 24Hr Post-Pollination; G, Globular; C, Cotyledon; M, Mature Green; PM, Post-mature Green

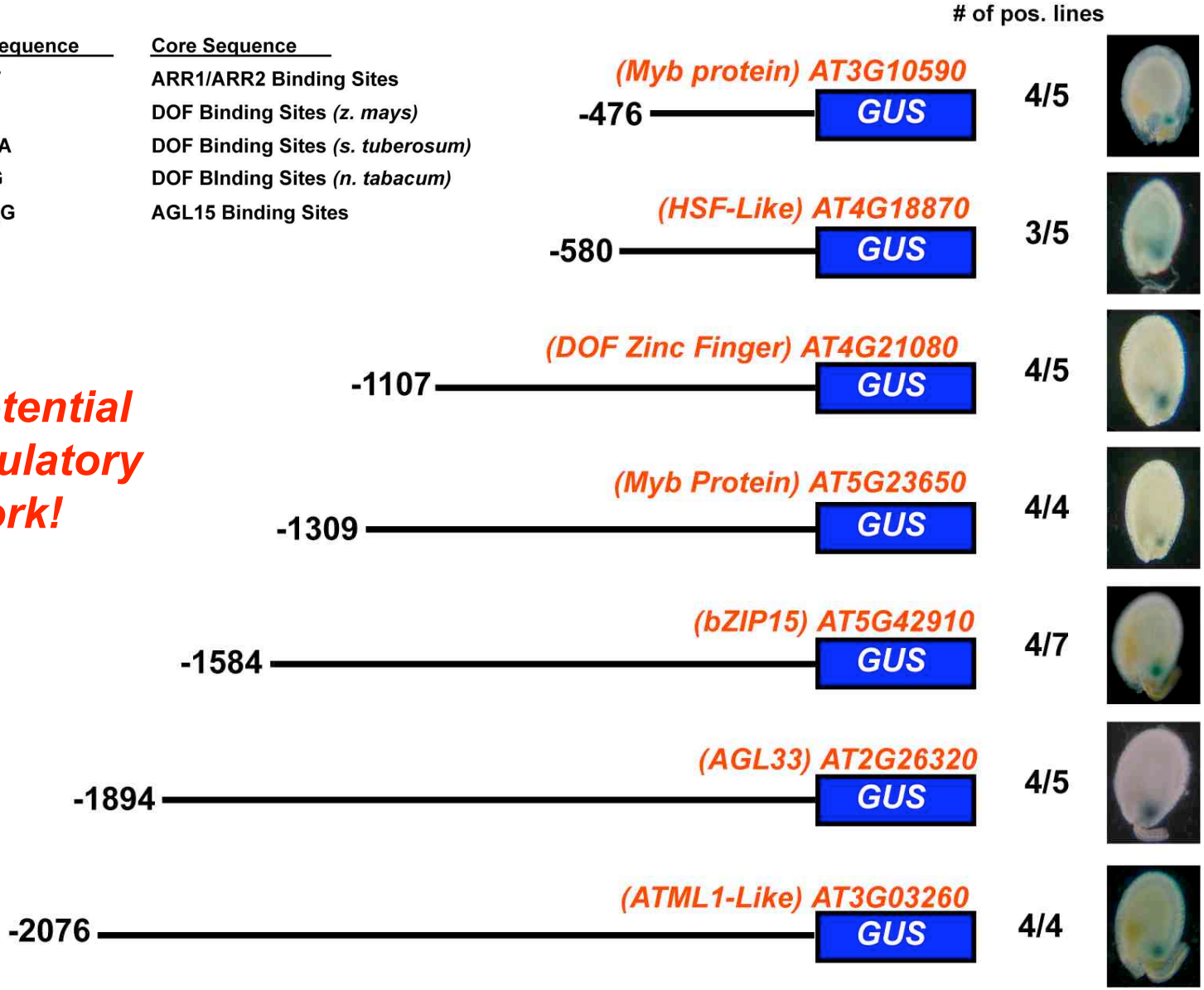
Transcriptional Patterns of Seed-Specific Transcription Factor Upstream Regulatory Region



What Transcription Factor Genes Are Active in the Chalazal Endosperm?

Motif Name	Core Sequence	Core Sequence
ARR1AT	NGATT	ARR1/ARR2 Binding Sites
DOFCOREZM	AAAG	DOF Binding Sites *(z. mays)*
NTBBF1ARROLB	ACTTTA	DOF Binding Sites *(s. tuberosum)*
TAAAGSTKST1	TAAAG	DOF BInding Sites *(n. tabacum)*
CARGCW8AT	$C(A/T)_8G$	AGL15 Binding Sites

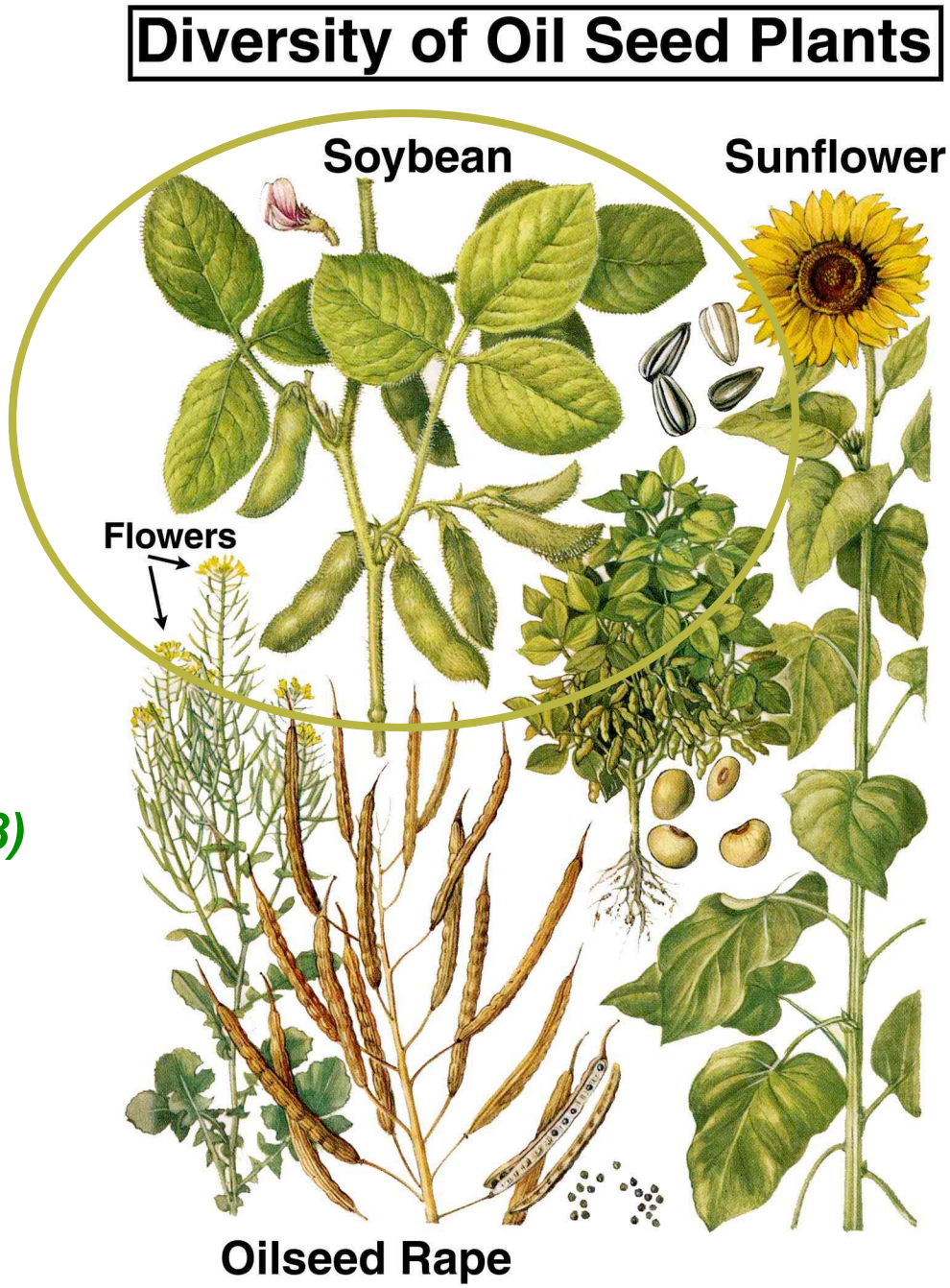
This a Potential Seed Regulatory Network!



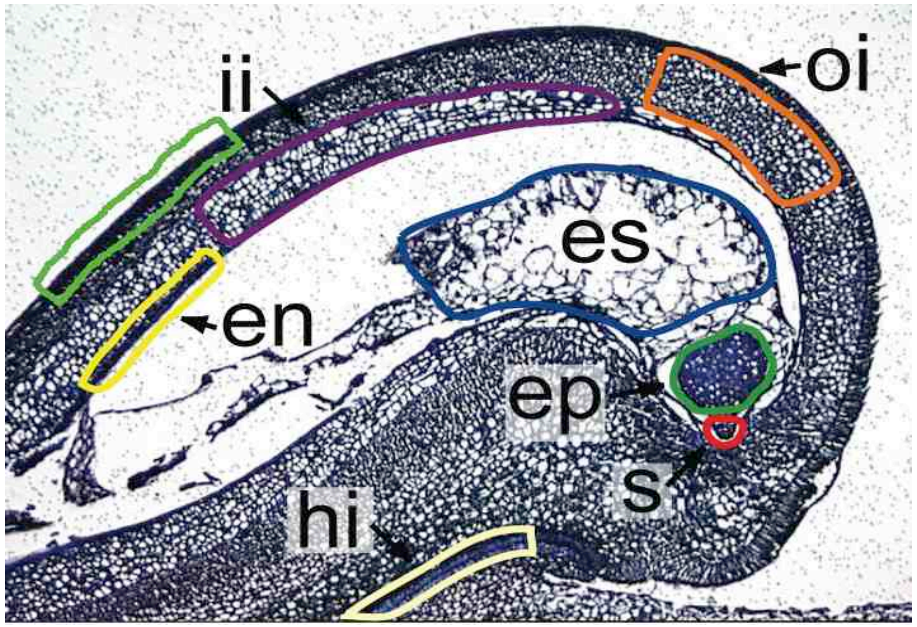
“Making A Globular Stage Soybean Seed”

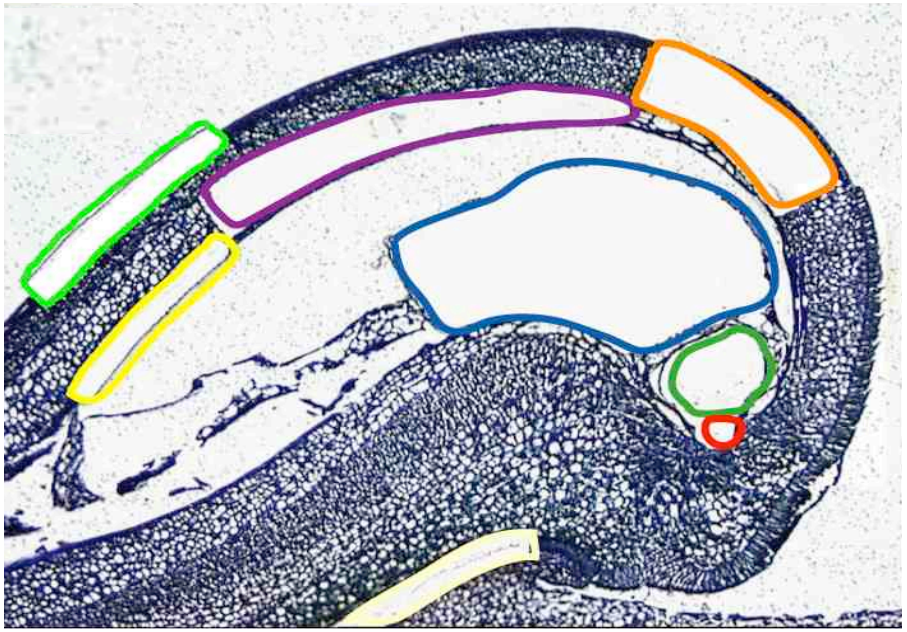
Why Soybean?

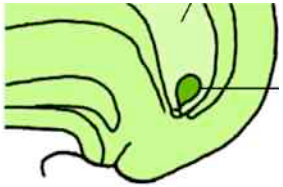
- *Second Major US Crop*
- *Major Food Source*
- *Major Biofuel Source*
- *Excellent Model Plant*
- *Genome Sequenced (2008)*
- *Major Funding Source*



How Can We Profile Gene Activity in All Seed Compartments, Regions, & Tissues?

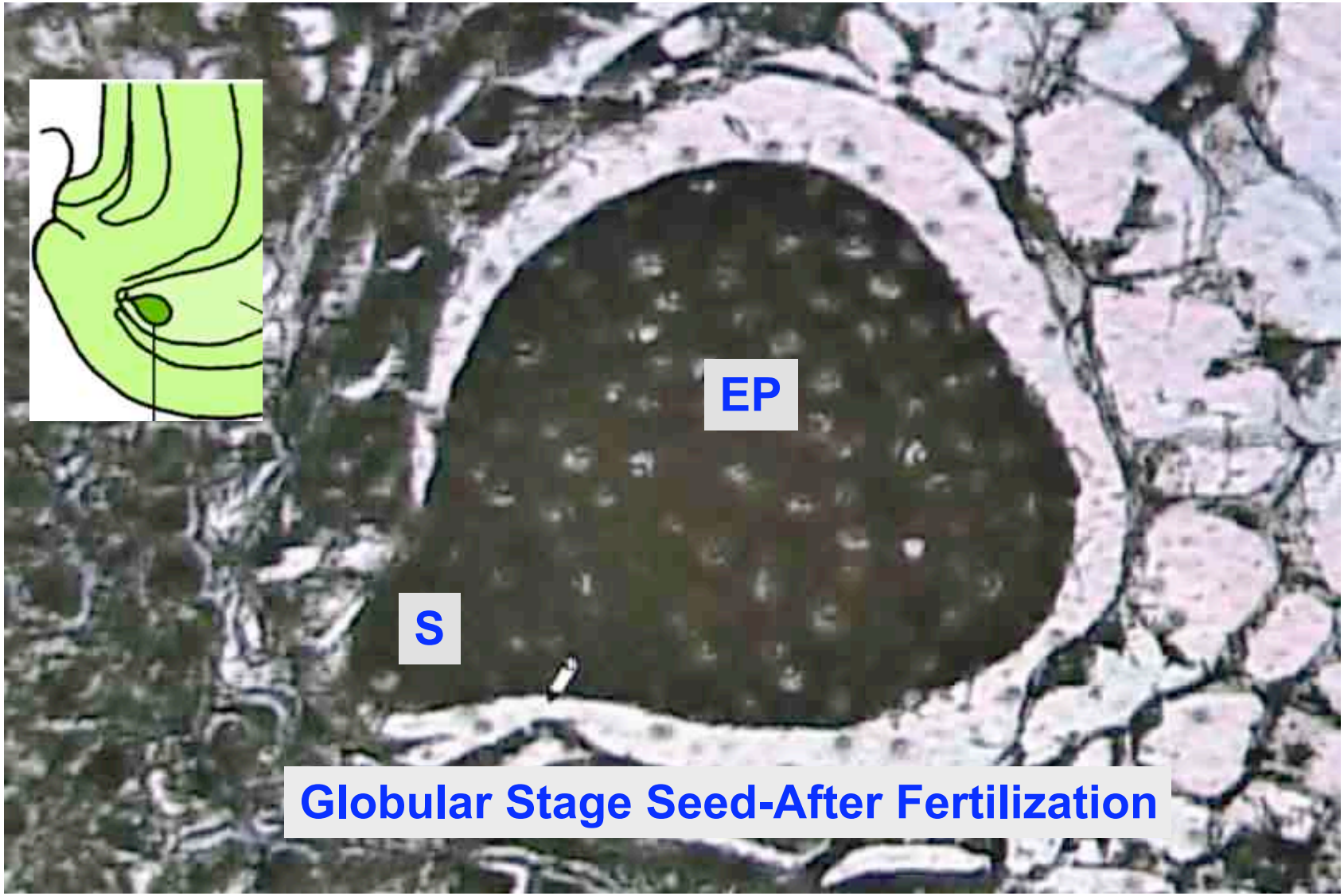






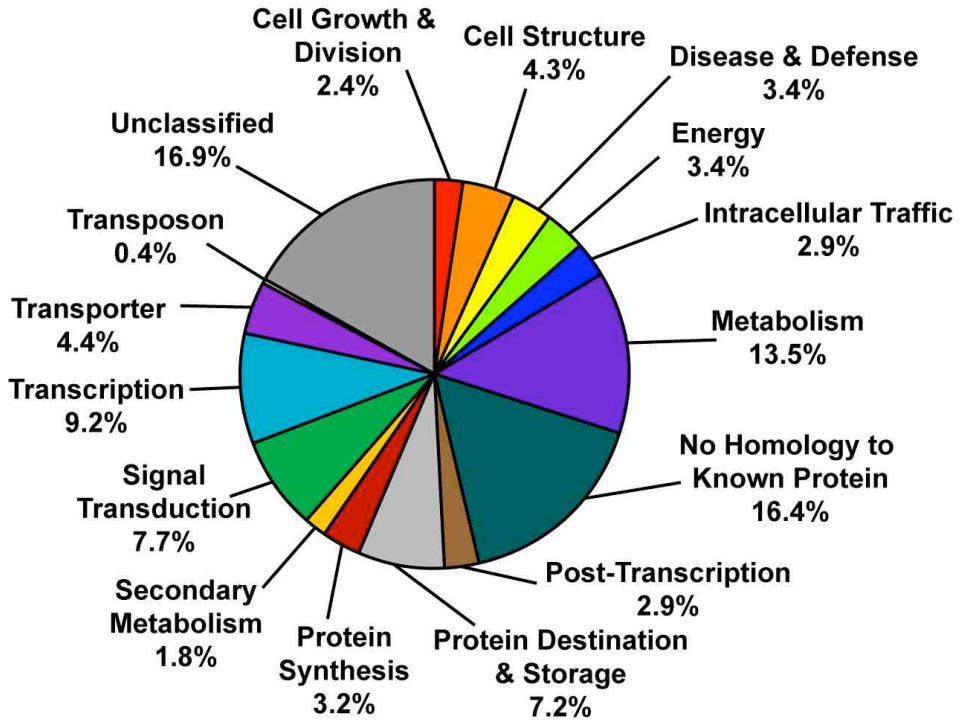
Combine Laser Capture Microdissection (LCM) Technologies With Genomics Approaches.

Using Laser Capture Microdissection (LCM) & Soybean GeneChips to Investigate Gene Activity In Seeds

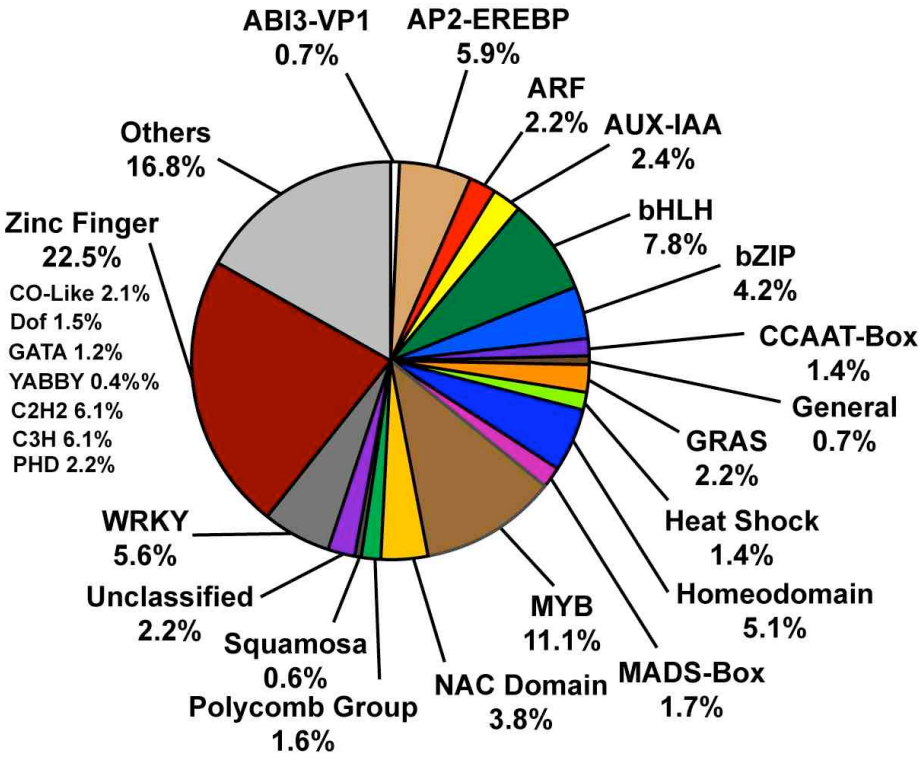


Spectrum of Gene Sequences Represented on the Soybean Affymetrix EST GeneChip (2007)

Functional Categories



Transcription Factors



~2,800 TF Transcripts

Contains Probe Sets Representing 38,000 Soybean Transcripts (~30,000 Clusters/~23,000 Predicted cDNAs) Derived From ~85 cDNA Libraries From Plant Regions and Multiple Developmental Stages (Not a Whole Genome Chip)

http://estdb.biology.ucla.edu/seed

Predicted Gene Model

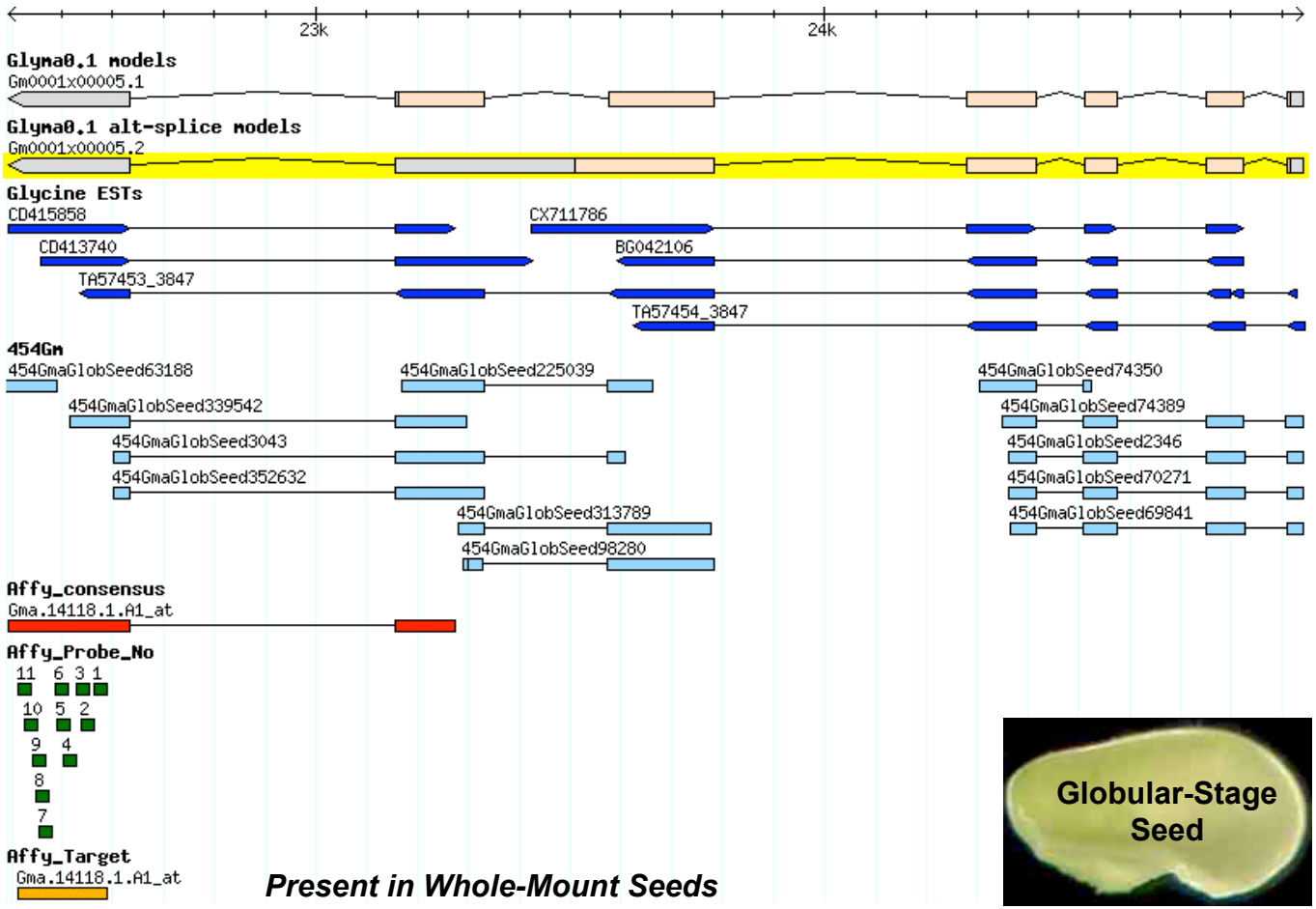
Soybean Public ESTs

Soybean Globular Seed 454 ESTs

Affymetrix Consensus

Affymetrix Probes

Affymetrix Target

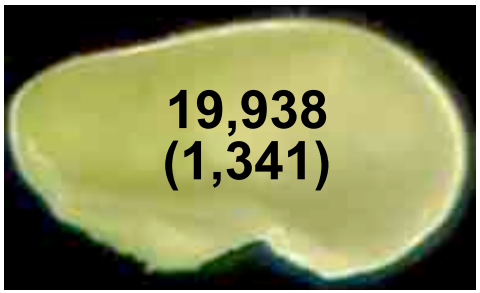


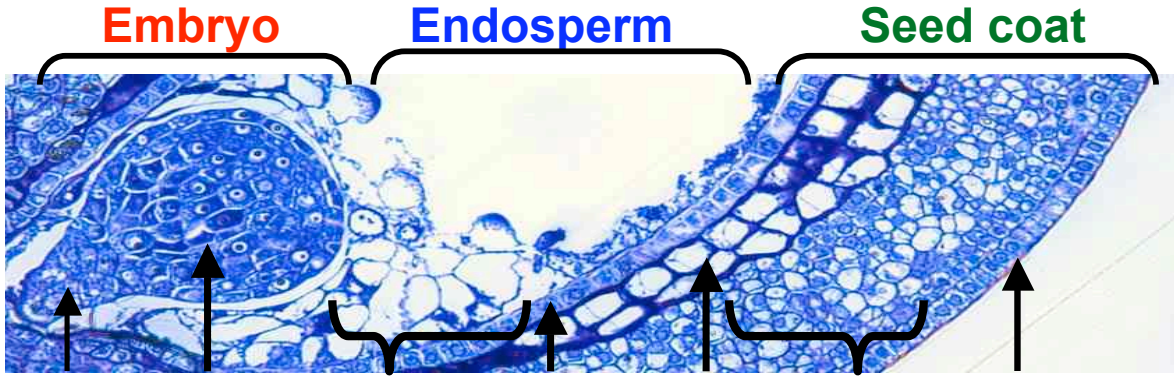
Present in Whole-Mount Seeds

http://estdb.biology.ucla.edu/seeds

Gene Activity in an Entire Globular-Stage Seed Immediately After Fertilization

Whole Mount



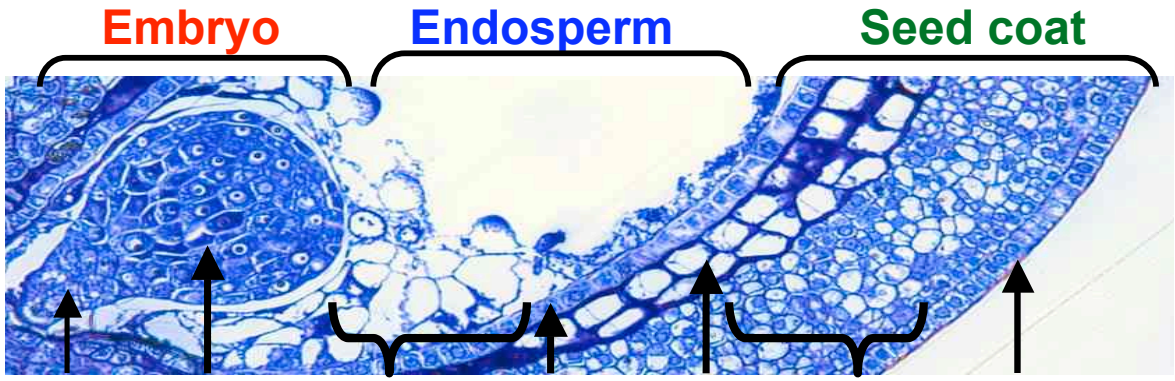


	S	EP	ES	EN	II	OI	EPD	HI
Total mRNAs*	14,177	16,998	13,880	15,274	14,767	16,402	13,451	16,153
TF mRNAs	909	1,100	848	950	927	1,073	837	1,057
Unique mRNAs	74	96	98	39	37	36	23	49
Unique TF mRNAs	5	17	11	3	4	2	2	6
Shared mRNAs	9025							

*** The present call in globular stage is defined as "present" at least in two biological replicates. One factor ANOVA p>0.05**

http://www.estdb.biology.ucla.edu/seed

Gene Activity in an Entire Globular-Stage Seed Immediately After Fertilization



	S	EP	ES	EN	II	OI	EPD	HI
Total mRNAs*	14,177	16,998	13,880	15,274	14,767	16,402	13,451	16,153
TF mRNAs	909	1,100	848	950	927	1,073	837	1,057
Unique mRNAs	74	96	98	39	37	36	23	49
Unique TF mRNAs	5	17	11	3	4	2	2	6
Shared mRNAs	9025	→						

*** The present call in globular stage is defined as "present" at least in two biological replicates**

Note: Unique genes are specific at the level of the GeneChip and within the seed

Globular-Stage Seed Compartments Have a Unique Set of Transcription Factor Genes

Outer Integument: 36 (2)

1 bZip
1 Zinc Finger

Inner Integument: 37 (4)

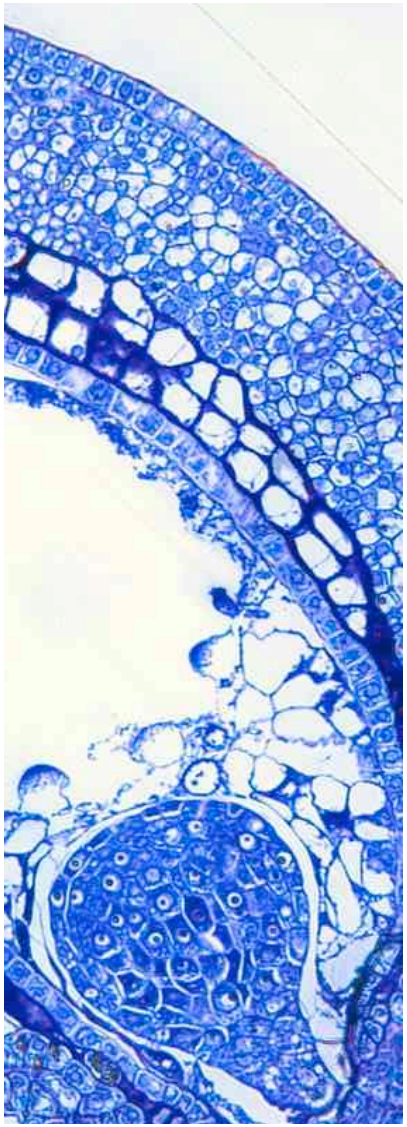
1 General (TFIIA-L)
1 Homeodomain
1 MADS, MYB

Endosperm: 98 (11)

3 Zinc Finger (Dof 2)
2 WRKY
1 ARF (ARF8), bHLH, JUMONJI
1 MADS-Box (PISTILLATA)
1 NAC Domain (NAM)
1 Polycomb Group (TRX1)

Embryo Proper: 96 (17)

4 MYB (MYB 44)
3 Homeodomain (STM, WOX2)
2 IAA (IAA8), bHLH, Zinc Finger
1 bZip (OBF4), WRKY, G2-like
1 GRAS (Scarecrow-like)



Hilum: 49 (6)

3 bHLH (PIF4)
1 Zinc Finger (IDD11),
1 MYB (MYB 111)
1 Aux/IAA (SOLITARY ROOT)

Epidermis: 23 (2)

1 AP2/EREBP
1 bZip

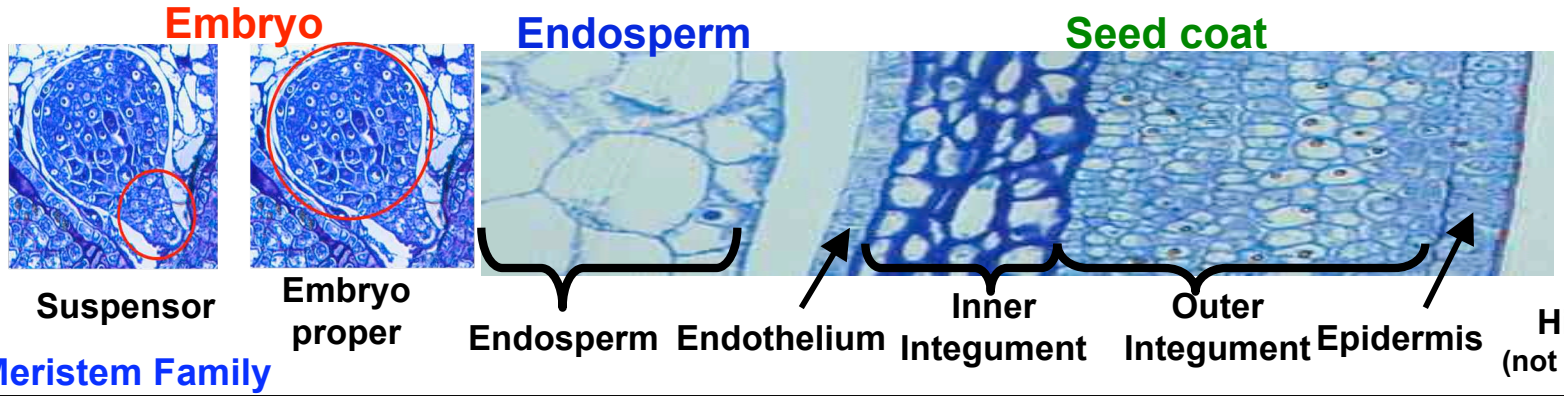
Endothelium: 39 (3)

2 Zinc Finger
1 WRKY

Suspensor: 74 (5)

2 WRKY
1 ARF (ARF16)
1 NAC Domain
1 Zinc Finger

Quantitative RT-PCR Validation of Putative Globular-Stage "Seed Region-Specific" mRNAs



No Apical Meristem Family

GeneChip	Present	A	A	A	A	A	A	A
qRT-PCR	30.9±3.2	ND	ND	ND	ND	ND	ND	ND

Cytochrome P450 family protein

GeneChip	A	A	A	Present	A	A	A	A
qRT-PCR	ND	ND	ND	35.0±0.6	ND	ND	ND	ND

Protease inhibitor/seed storage

GeneChip	A	A	A	A	Present	A	A	A
qRT-PCR	ND	ND	ND	ND	32.6±0.7	ND	ND	ND

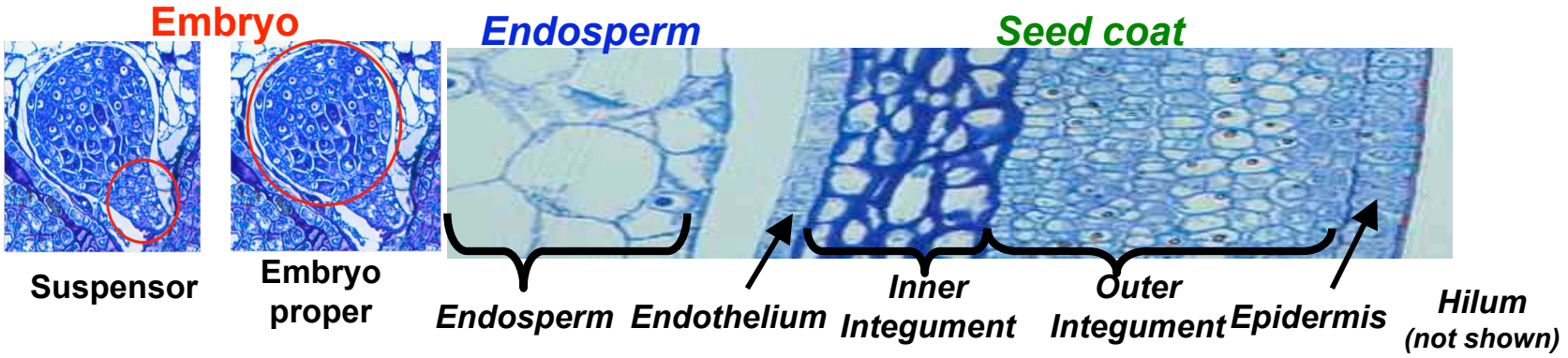
Zinc Finger Family (C2H2-type)

GeneChip	A	A	A	A	A	A	A	Present
qRT-PCR	ND	ND	ND	ND	ND	ND	ND	35.8±2.4

A=Absent = Not Detected at GeneChip Level

ND = Not Detected at Level of qRT-PCR

Quantitative RT-PCR Validation and GUS Activity of a Globular-Stage "Seed Region-Specific" mRNA

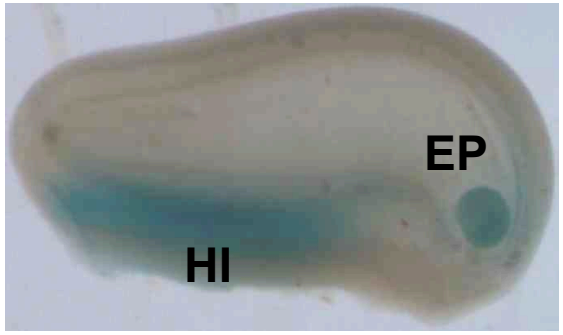


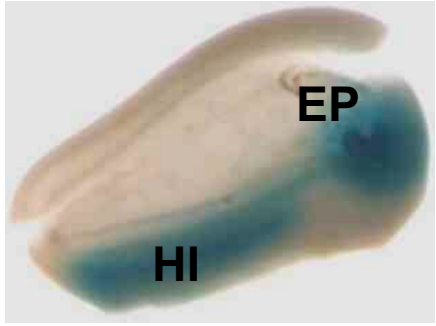
GmSTM-Like (Shoot Meristemless-like) Transcription Factor (GmaAffx.63050.1.S1_at)

GeneChip	A	P	A	A	A	A	A	A
qRT-PCR	ND	30.0±0.6	ND	ND	ND	ND	ND	39.5±1.1
Fold Reduction	-	1	-	-	-	-	-	- 687

A=Absent = Not Detected at GeneChip Level

ND = Not Detected at Level of qRT-PCR

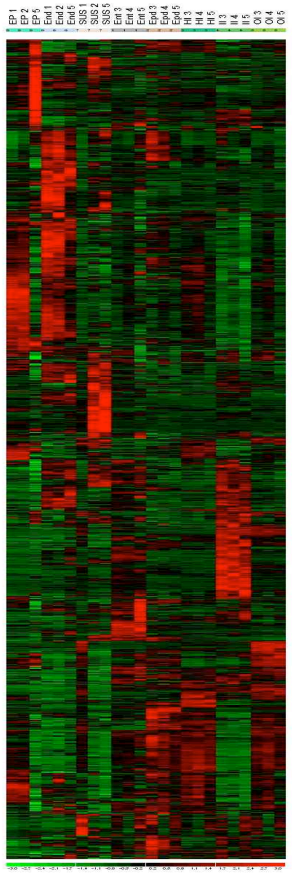


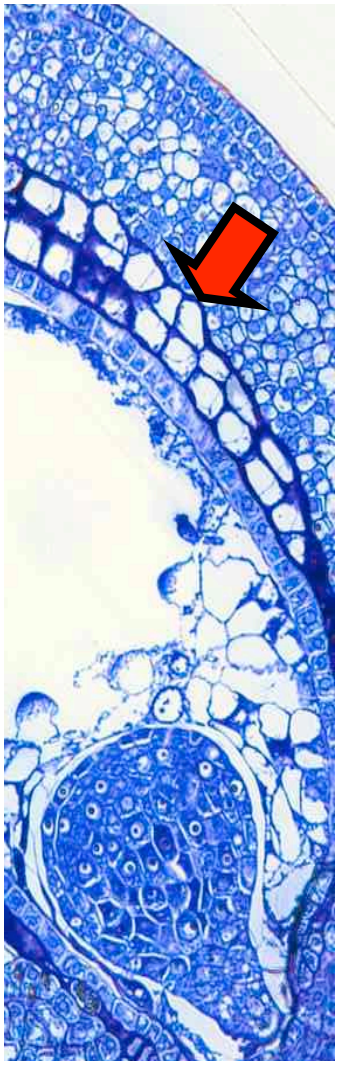


GmSTM-like promoter/GUS **Gene Activity in Soybean Seeds**

Quantitative Regulation of mRNAs Shared by Soybean Globular-Stage Seed Compartments

Top 2000 Varying mRNAs





> 2-fold (p<0.05)

EP ES S EN EPD HI II OI

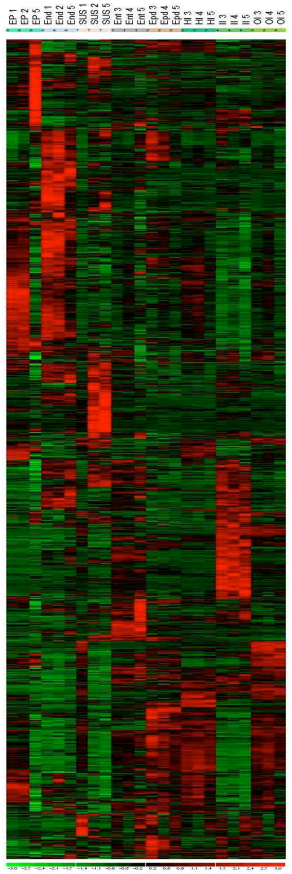
HI	4/5	Protein Storage Transcription
EPD	12/12	Signaling Energy Metabolism
OI	28/62	Energy Metabolism Transporter
II	86/120	Intracellular Traffic Transporter Metabolism
EN	17/31	Metabolism Cell Structure Protein Storage
ES	23/73	Energy Protein Synthesis Protein Storage
EP	4/13	Cell Growth Protein Storage Post-Transcription
S	1/66	Energy Metabolism Protein Storage

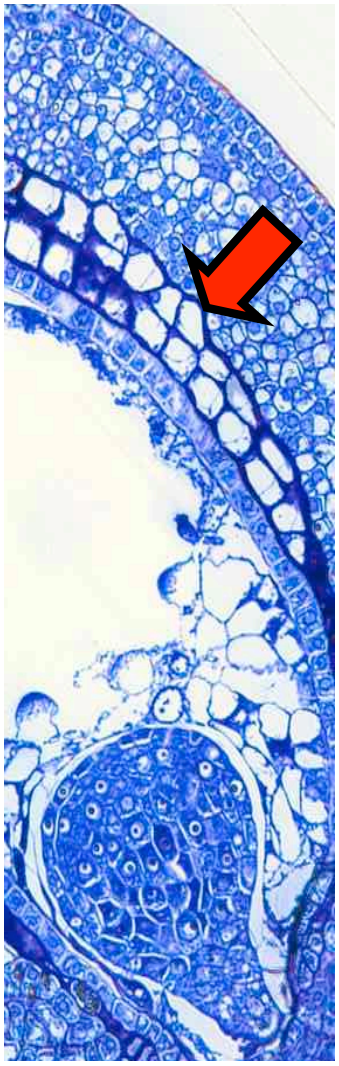
-3.0 -2.7 -2.4 -2.1 -1.7 -1.4 -1.1 -0.8 -0.5 -0.2 0.2 0.5 0.8 1.1 1.4 1.7 2.1 2.4 2.7 3.0

FDR < 5%

Quantitative Regulation of mRNAs Shared by Soybean Globular-Stage Seed Compartments

Top 2000 Varying mRNAs

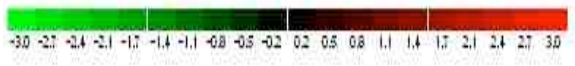




> 2-fold (p<0.05)

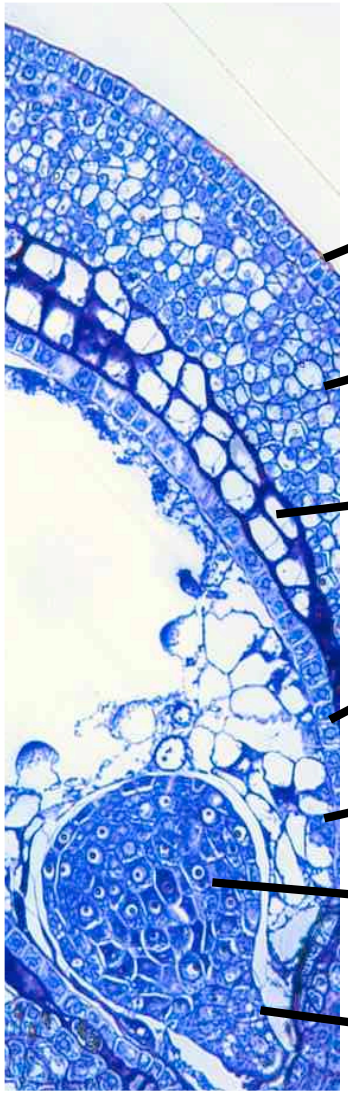
EP ES S EN EPD HI II OI

HI	4/5	Protein Storage Transcription	
EPD	12/12	Signaling Energy Metabolism	
OI	28/62	Energy Metabolism Transporter	
II	86/120	Vacuolar ATPase STP1 - CAT7 SIS - GATL7 - SAH7	
EN	17/31	Metabolism Cell Structure Protein Storage	
ES	23/73	Energy Protein Synthesis Protein Storage	
EP	4/13	Cell Growth Protein Storage Post-Transcription	
S	1/66	Energy Metabolism Protein Storage	



FDR < 5%

How Many Genes Are Active in a Globular-Stage Soybean Seed?



Hilum
16,153 (1,057)

Epidermis
13,451 (837)

Outer Integument
16,402 (1,073)

Inner Integument
14,767 (927)

Endothelium
15,274 (950)

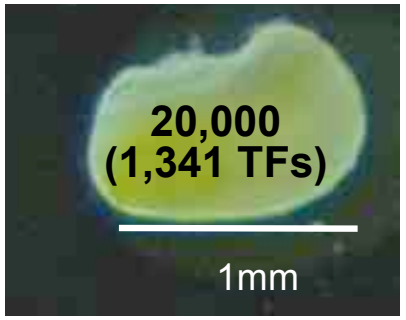
Endosperm
13,880 (848)

Embryo Proper
16,998 (1,100)

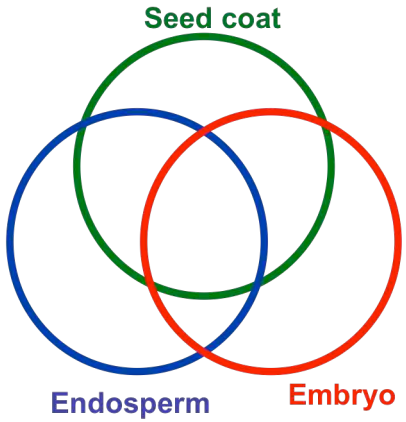
Suspensor
14,177 (909)

At the Level of GeneChip Detection!

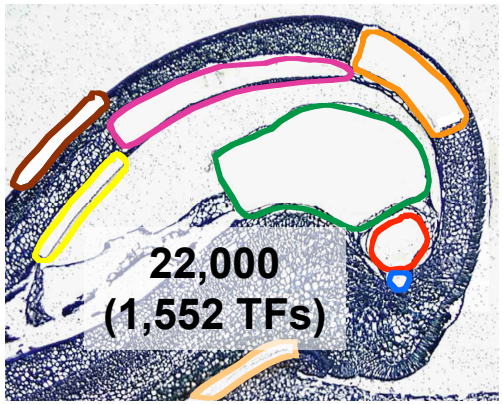
Whole Mount



Union of Seed Compartments



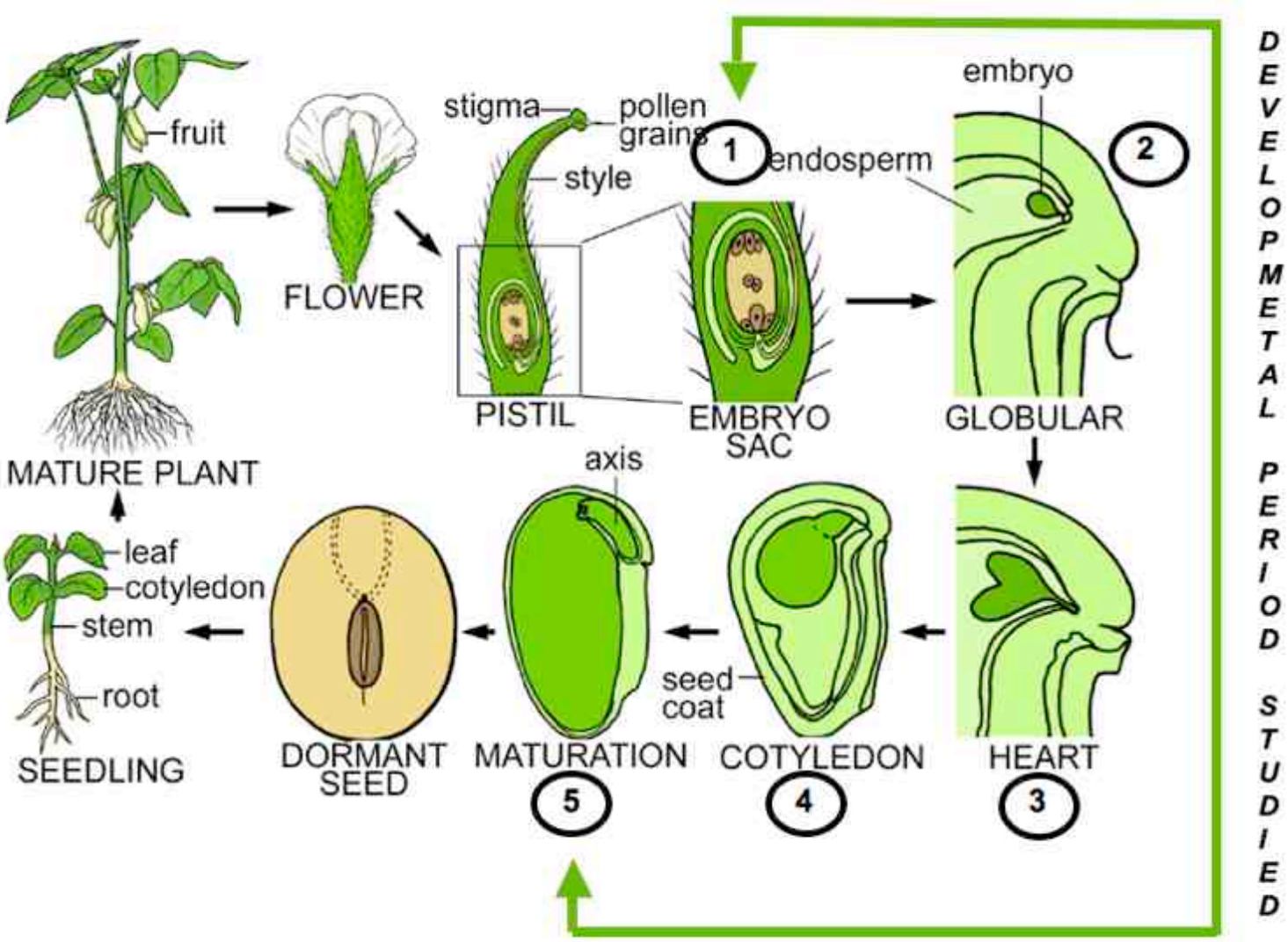
Union of LCM mRNA Sets



(Transcription Factors)

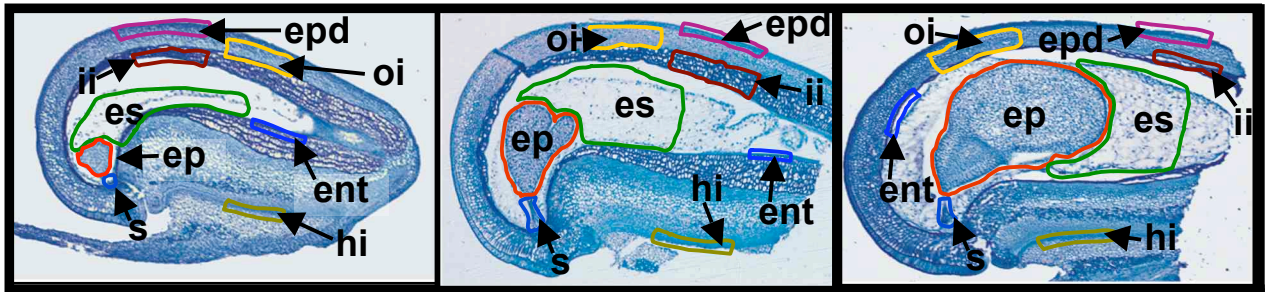
Note: Minimum Numbers!!

What Are The Genes Required to Program Every Compartment, Tissue, and Cell Type During Soybean Seed Development?

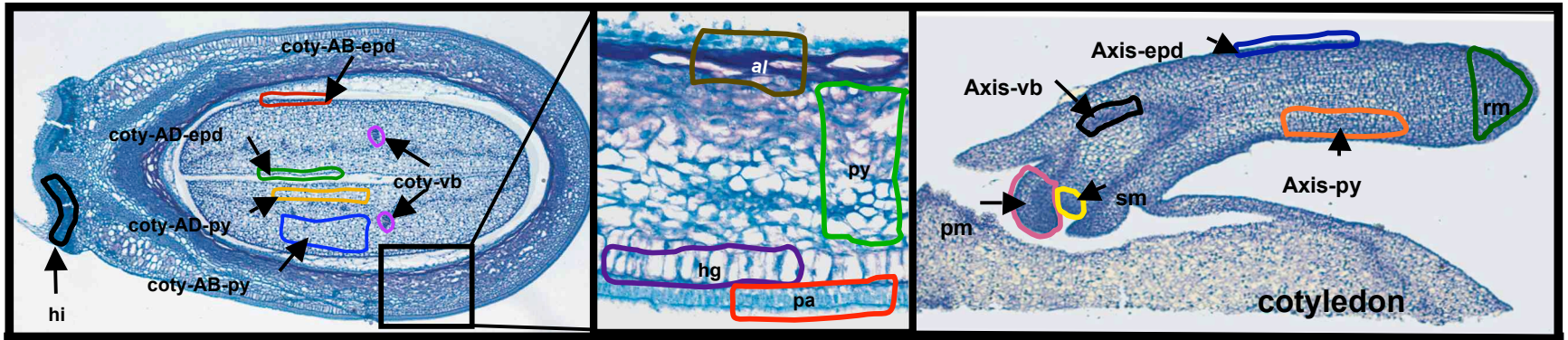


***Note**: Numbers Refer to Specific Seed Stages Studied*

More Than 40 Soybean Seed Compartments Have Been Studied Using GeneChip and LCM Technology



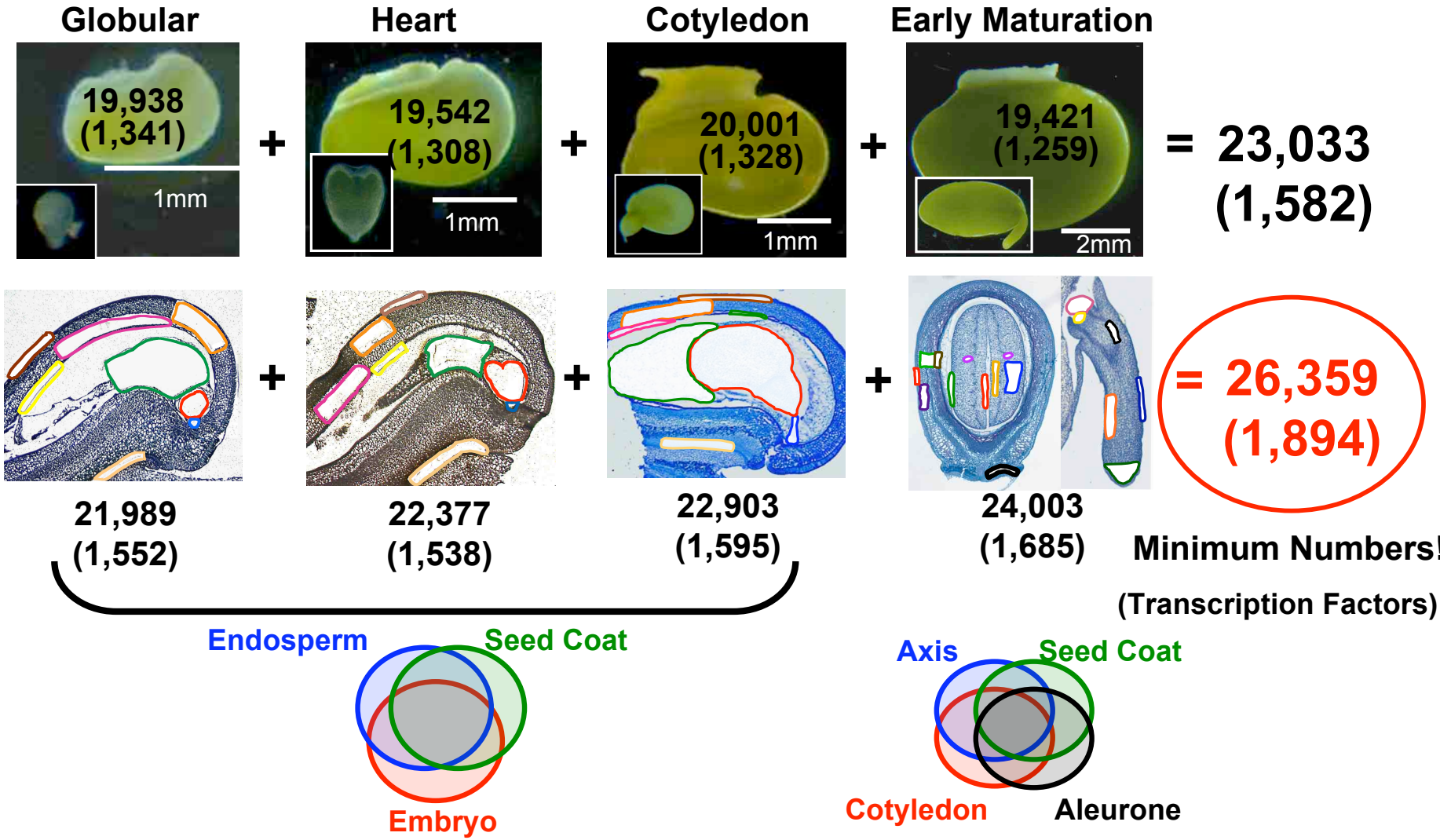
Globular	Heart	Cotyledon
Embryo proper (ep)	Endosperm (es)	Endothelium (ent)
Inner Integument (ii)	Epidermis (epd)	Hilum (hi)
Outer Integument (oi)	Suspensor (s)	



Early- , Mid- , Late-Maturation Stages

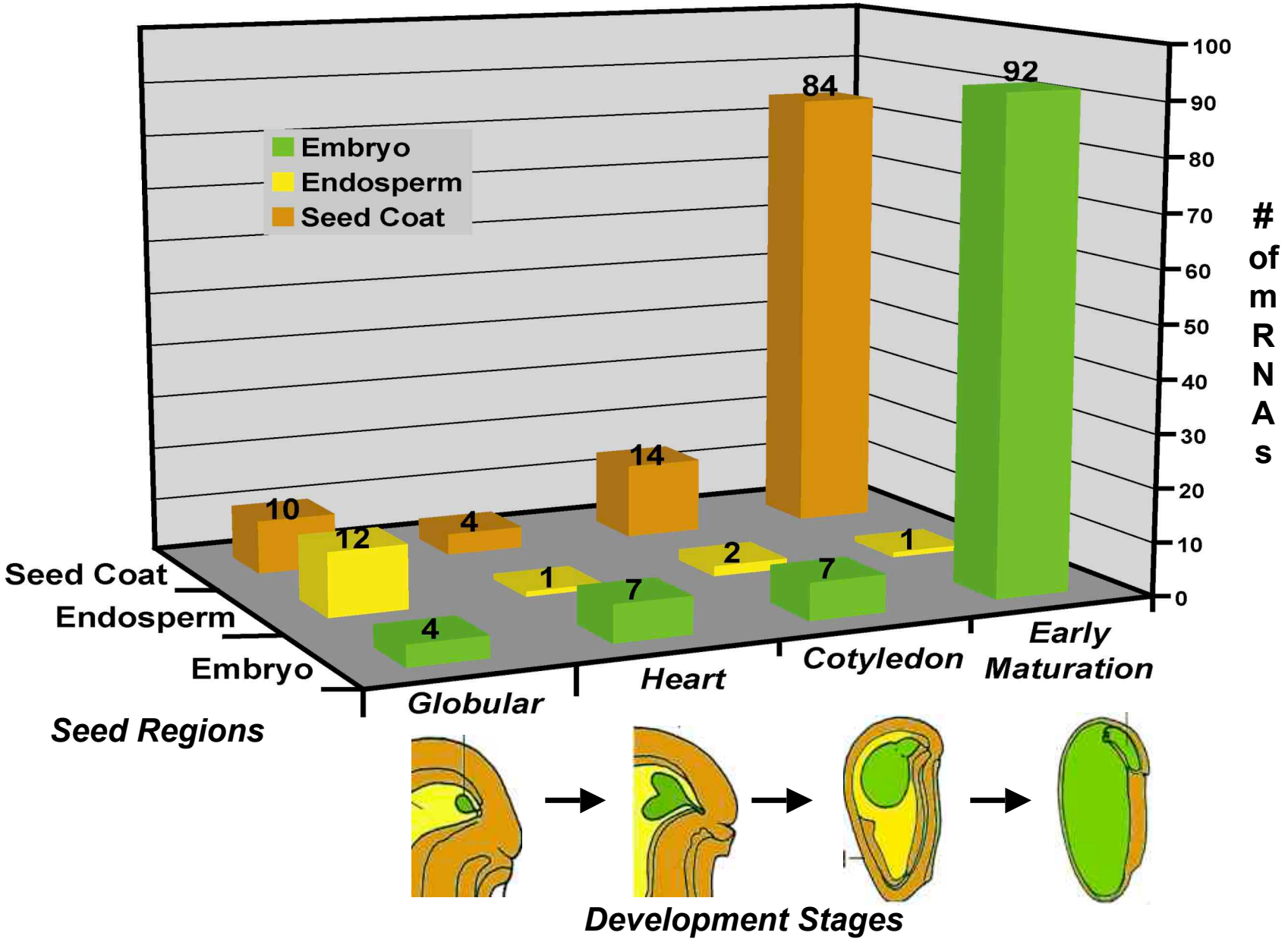
Cotyledon-vascular bundle (coty-vb)	Hilum (hi)	Axis-plumule (pm)
Cotyledon-Abaxial epidermis (coty-AB-epd)	Aleurone (al)	Axis-root meristem (rm)
Cotyledon-Adaxial epidermis (coty-AD-epd)	Seed coat-palisade (pa)	Axis-shoot meristem (sm)
Cotyledon-Abaxial parenchyma (coty-AB-py)	Seed coat-hour glass (hg)	Axis-epidermis (Axis-epd)
Cotyledon-Adaxial parenchyma (coty-AD-py)	Seed coat-parenchyma (py)	Axis-parenchyma (Axis-py)
		Axis-vascular bundle (Axis)

How Many Genes Are Required to Program Soybean Seed Development?

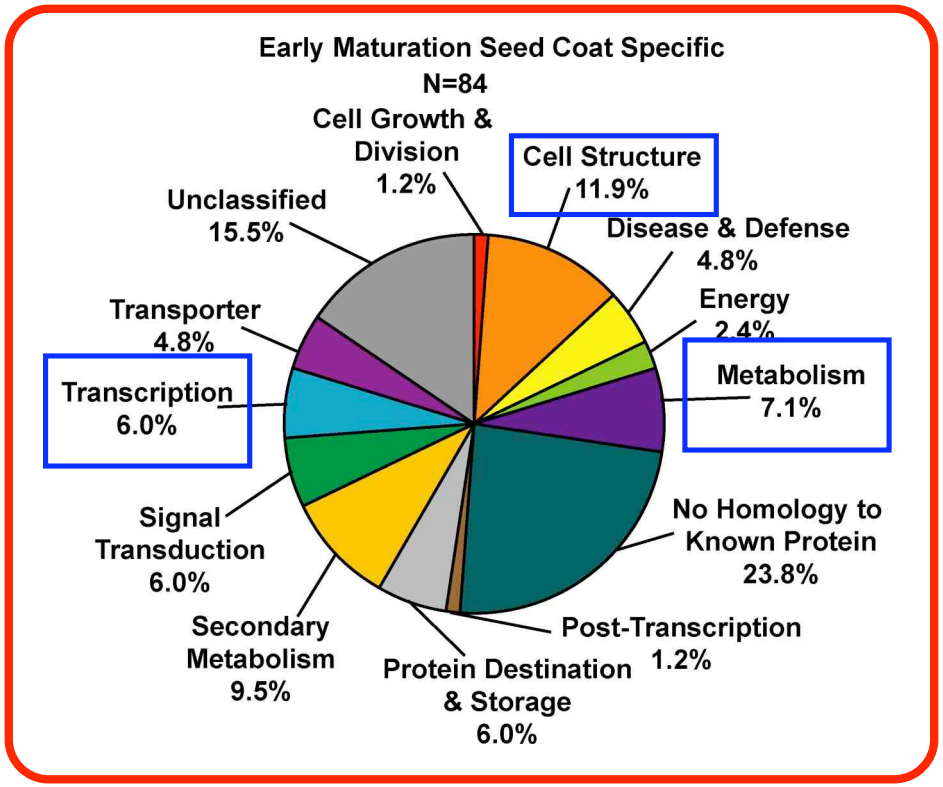


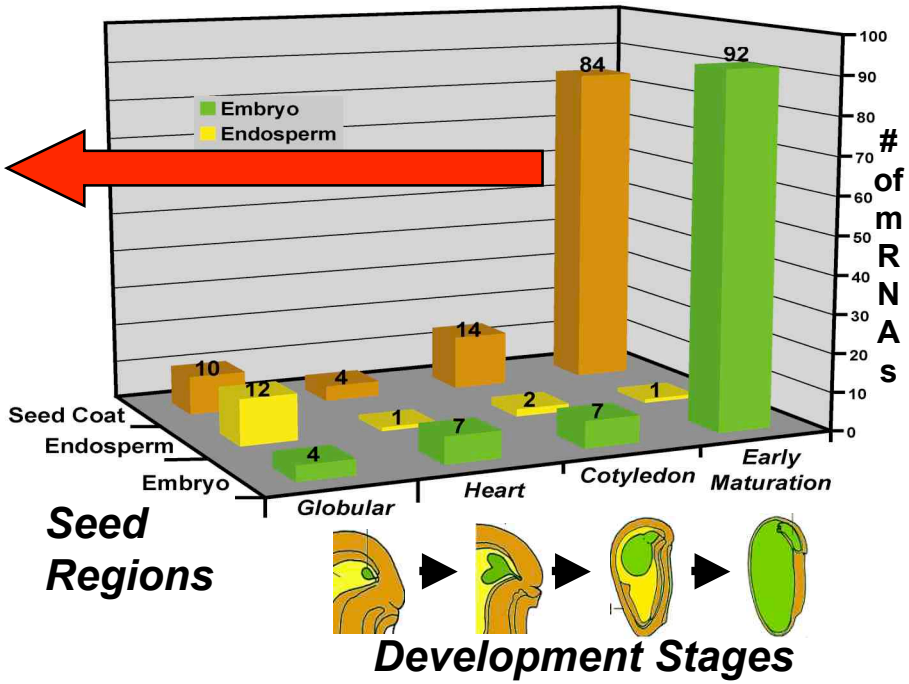
40 Compartments & Tissues Profiled- More than 3.7 Million Data Points!!

Each Soybean Seed Compartment Has a Unique Set of mRNA at Different Developmental Stages



Each Soybean Seed Compartment Has a Unique Set of mRNA at Different Developmental Stages (e.g., Seed Coat)





Cell Structure
Cellulose Synthase
(CESA8, CESA4, IRX3)
Pectinacetylesterase Family
Glucan 1,3-beta-glucosidase

Secondary Metabolism
BANYULS (DFR)
TRANSPARENT TESTA 10
Isoflavone Reductase Homolog
2OG-Fe(II) Oxygenase

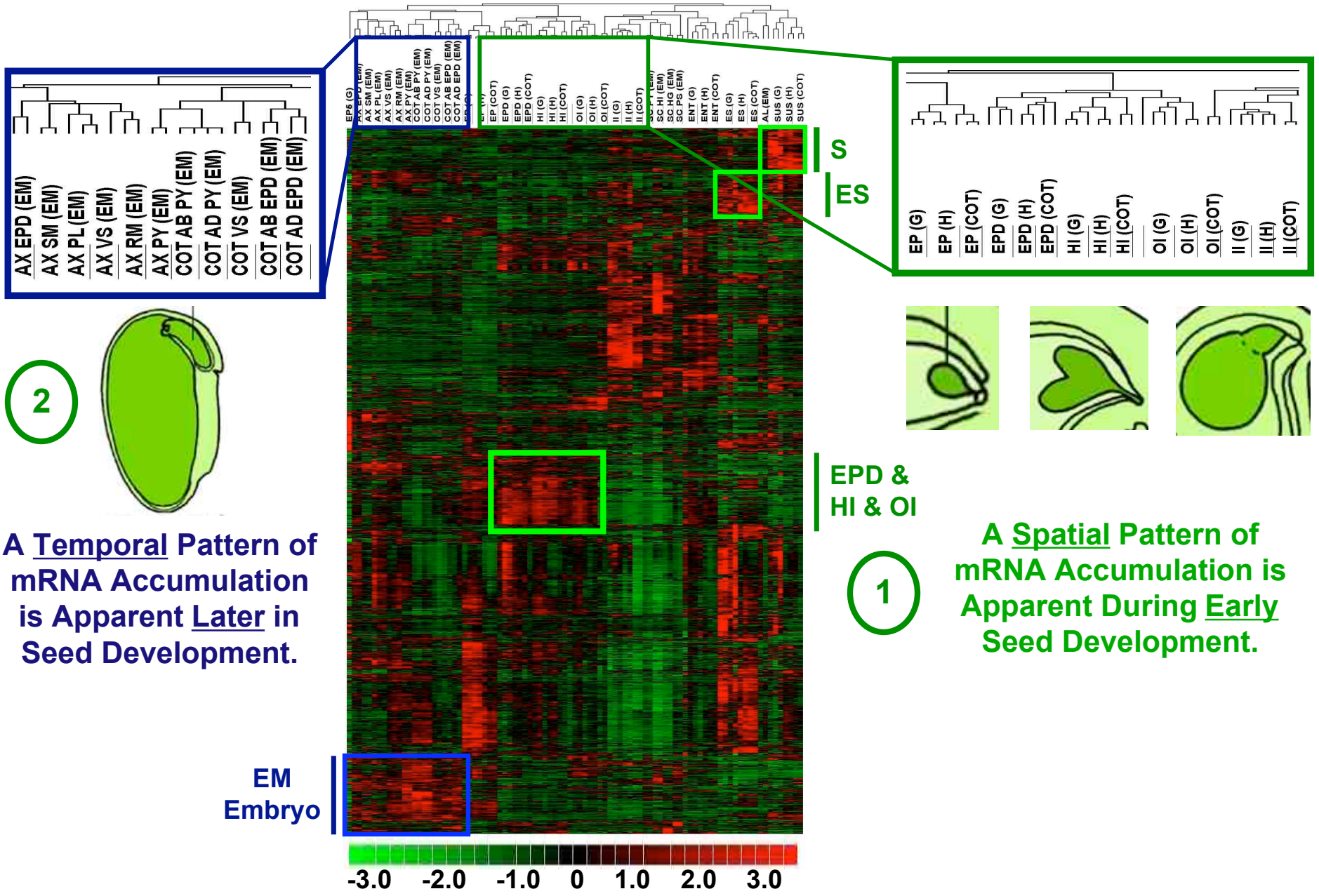
Transcription
SCARECROW-LIKE 3
C2H2 Zinc Finger
B-Box Type Zinc Finger

Cell Wall Enzymes

Brown Color Enzymes

Gene Regulators

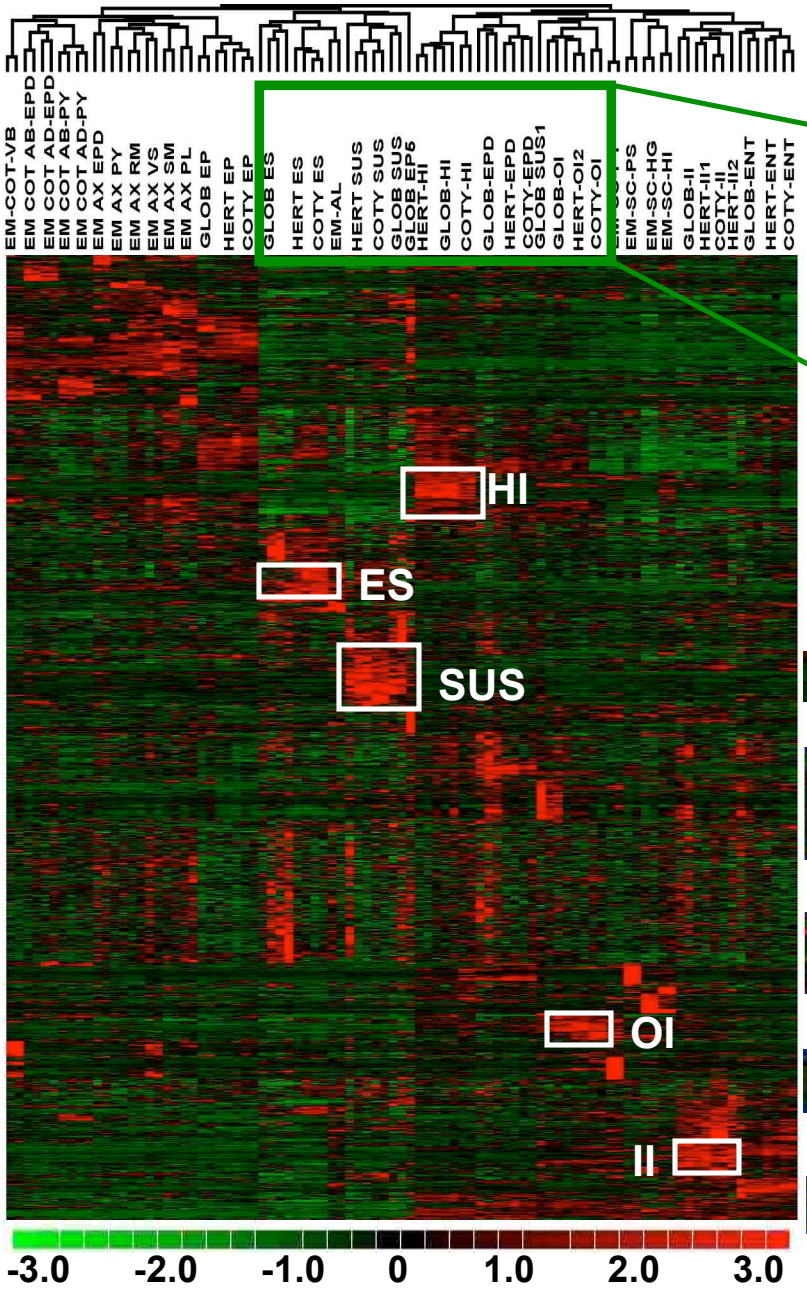
What Are The Major Patterns of Gene Activity During Soybean Seed Development?



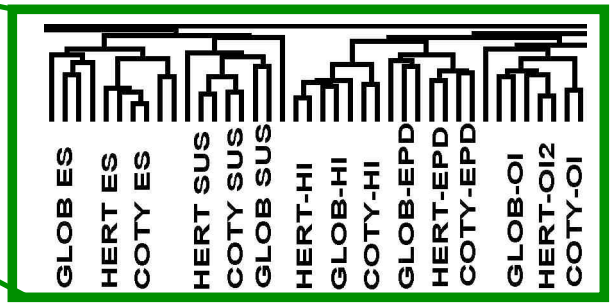
A Temporal Pattern of mRNA Accumulation is Apparent Later in Seed Development.

A Spatial Pattern of mRNA Accumulation is Apparent During Early Seed Development.

Spatial Patterns of Transcription Factor mRNAs Upregulation During Early Soybean Seed Development



Hierarchical Clustering of 1,932 TFs Present in at Least One Compartment Across Development Stages



Over-represented (2-fold, p<0.05)

ES — **22/42** — **ARF (1)**

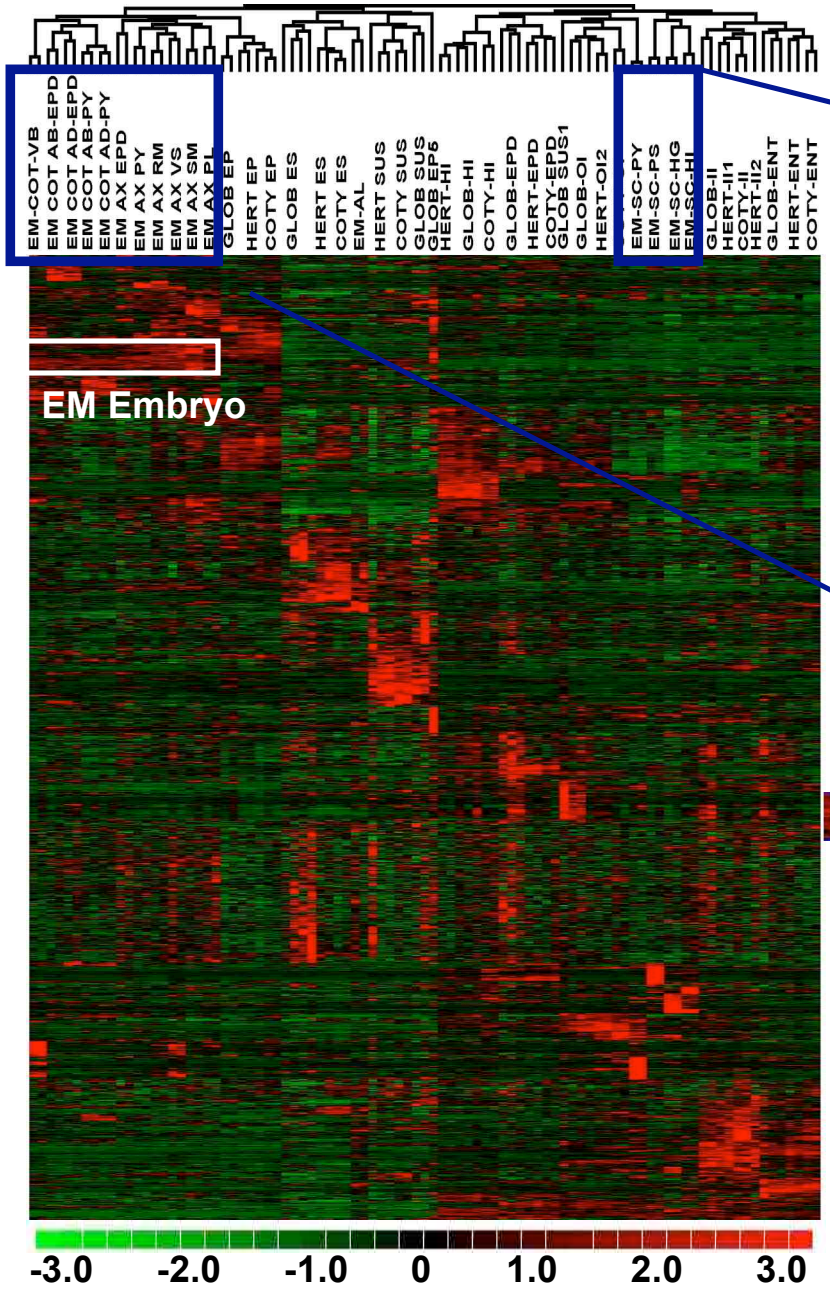
SUS — **51/58** — **WRKY (15)**, **CCAAT-Box (4)**

HI — **39/53** — **bHLH (10)**, **Squamosa (2)**

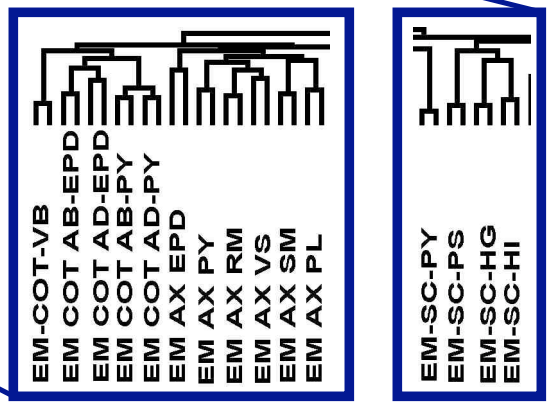
OI — **15/18** — **Unclassified (3)**

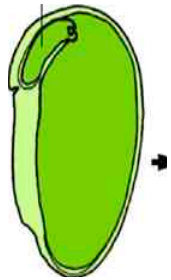
II — **9/9** — **No Over-representation**

Temporal Pattern of Transcription Factor mRNA Upregulation in All Seed Compartments During Early Maturation



Hierarchical Clustering of 1,932 TFs Present in at Least One Compartment Across Development Stages





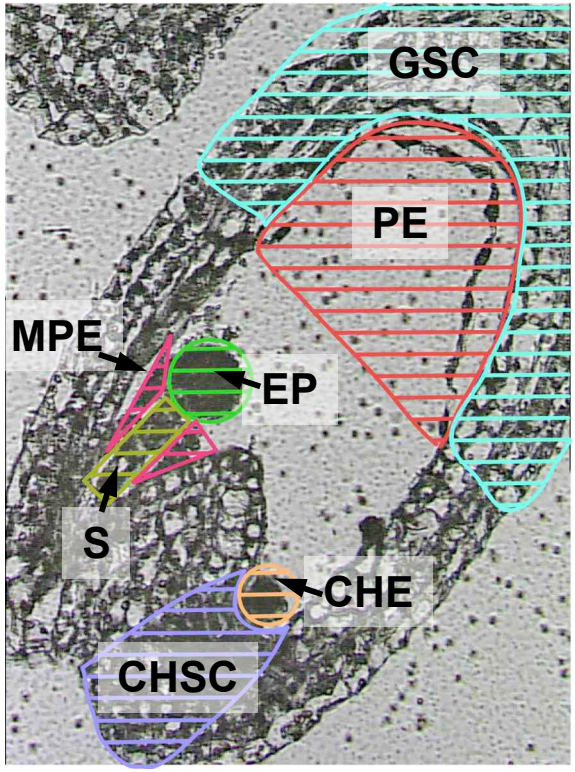
EM Embryo

13/23

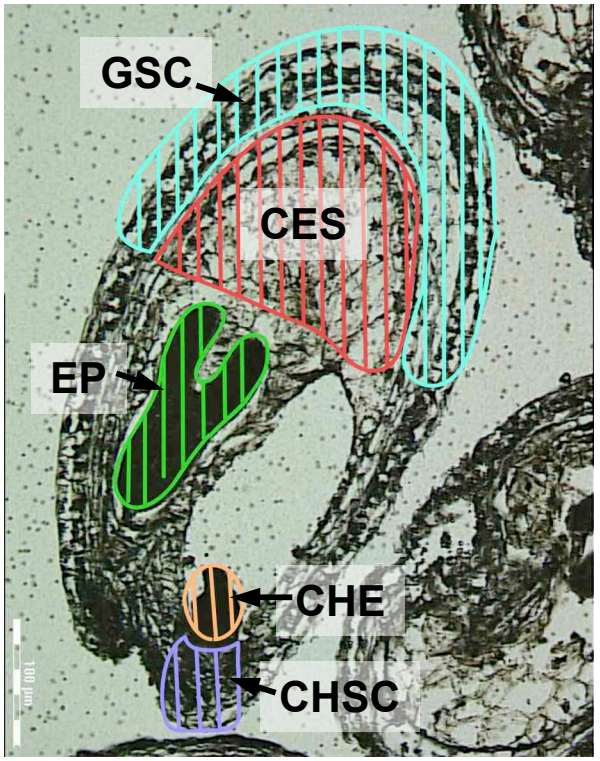
Over-represented (2-fold, p<0.05)

bZIP (2)
CCAAT-Box (3)

Arabidopsis Seed Compartments Captured for RNA Profiling Studies



Globular-Stage



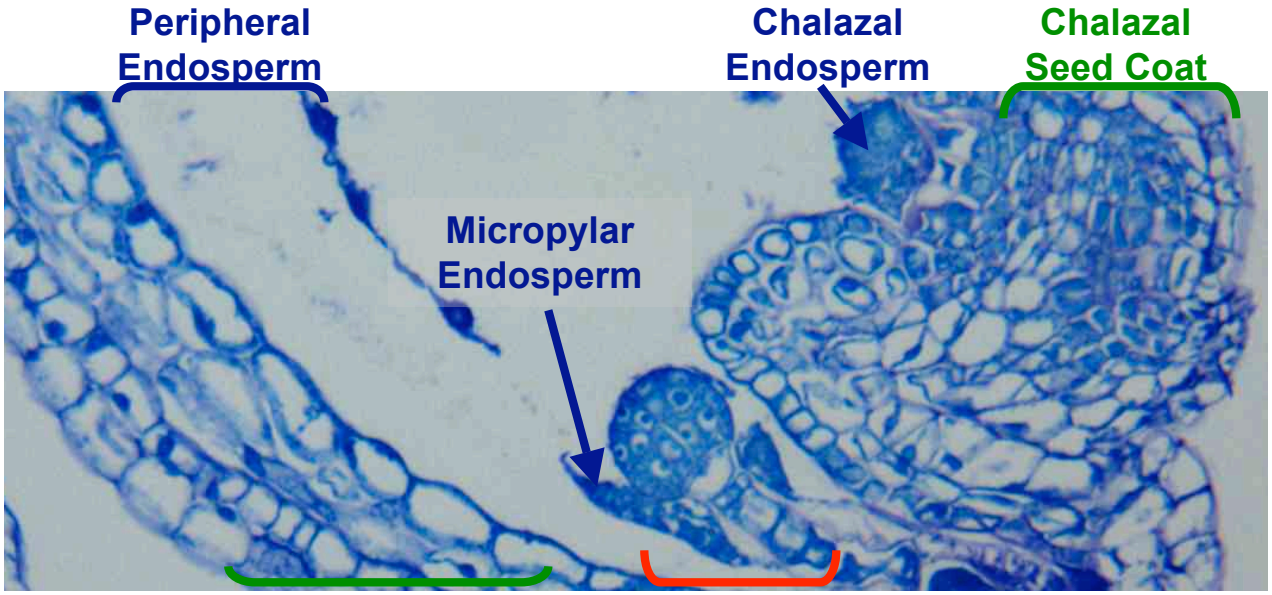
Linear Cotyledon-Stage

General Seed Coat (GSC)	Suspensor (S)
Peripheral Endosperm (PE)	Chalazal Endosperm (CHE)
Embryo Proper (EP)	Chalazal Seed Coat (CHSC)
Micropylar Endosperm (ME)	Cellularized endosperm (CES)

Using Laser Capture Microdissection (LCM) to Profile Gene Activity in Specific Seed Regions

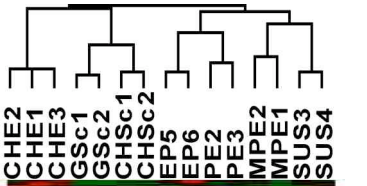
Gene Activity in an Entire Globular-Stage Arabidopsis Seed Immediately After Fertilization

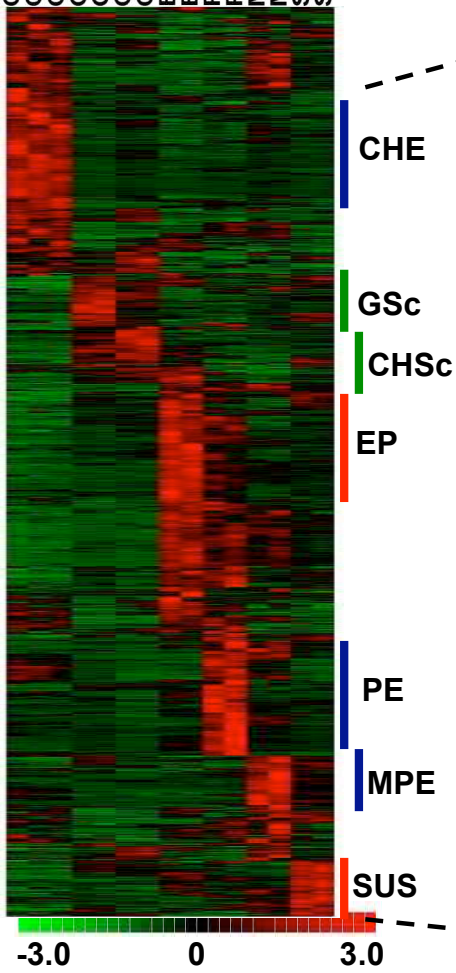
*Unamplified



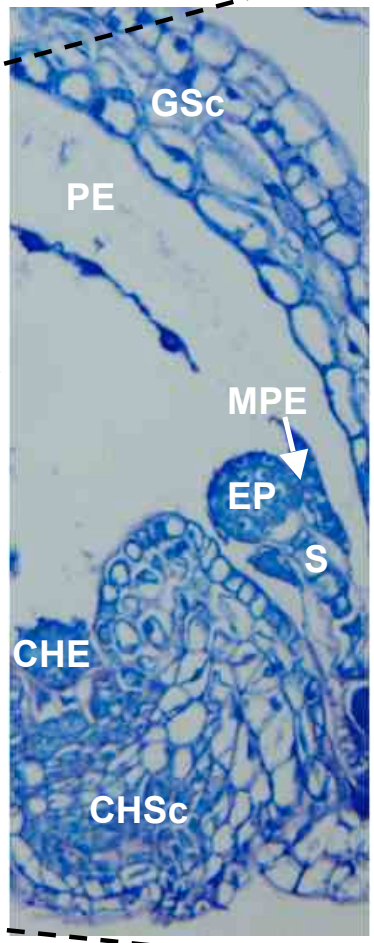
	GSC	PE	MPE	EP	S	CHE	CHSC
mRNA Present BR1	11,402	11,445	7,844	10,532	11,273	7,709	12,440
mRNA Present BR2	12,316	9,972	10,658	10,674	12,965	7,002	11,818
mRNA Present BR3	-					9,793	-
Total mRNAs	10,513	9,136	7,216	9,130	10,692	7,776	10,794
Unique mRNA	108	54	14	56	95	133	152
Unique TFs	16	3	0	17	8	10	26
Shared	4,028	→					

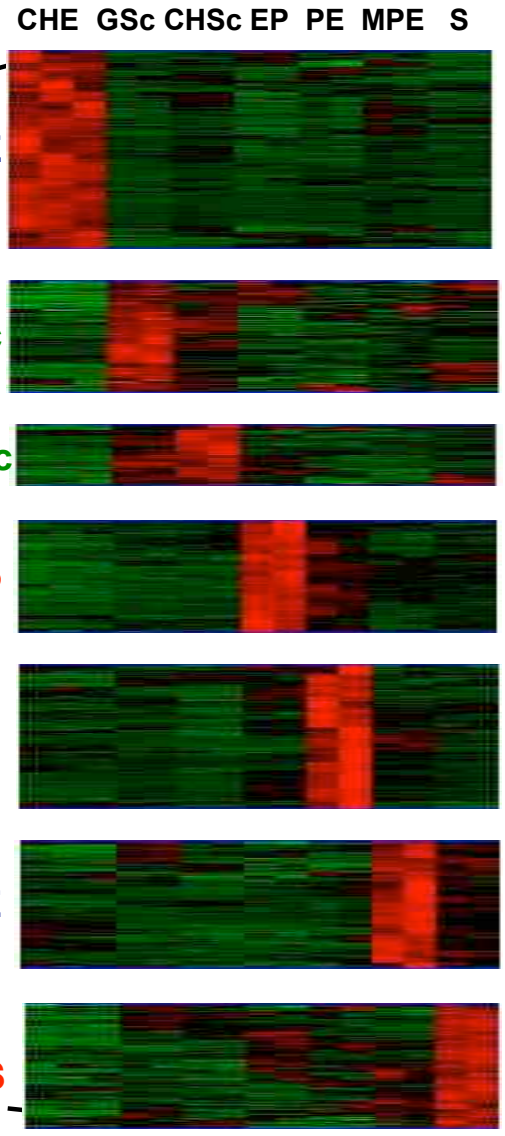
Quantitative Regulation of Transcripts Shared by Arabidopsis Globular Stage Seed Compartments





Top 2000 Varying mRNAs

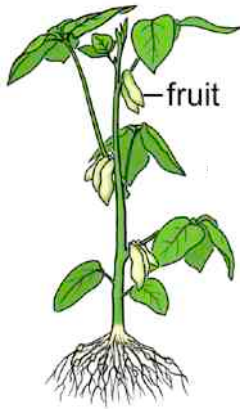




Over-represented (2-fold, p<0.05)

CHE	204/248	Intracellular Traffic (16) Post-Transcription (11)
GSc	45/112	Energy (4) Metabolism (17)
CHSc	40/83	Energy (1)
EP	114/119	Cell Growth & Division (9) Post-Transcription (13) Protein Synthesis (46)
PE	132/167	Energy (24) Protein Synthesis (28)
MPE	59/114	Intracellular Traffic (2) Protein Destination & Storage (9)
SUS	59/124	Energy (5) Metabolism (20) Protein Destination & Storage (8)

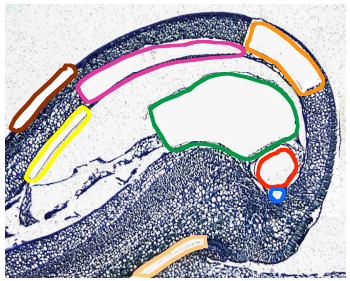
Comparison of Soybean and Arabidopsis Gene Activity at Globular- and Cotyledon-Stage



Soybean

Globular-Stage

LCM

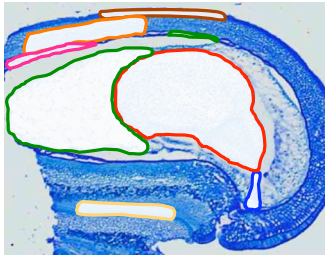


21,989

+

Cotyledon-Stage

LCM



22,903

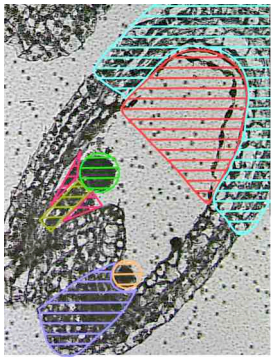
=

24,322

Arabidopsis

Globular-Stage

LCM

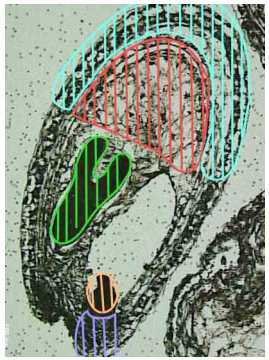


15,156
(18,247)

+

Cotyledon-Stage

LCM

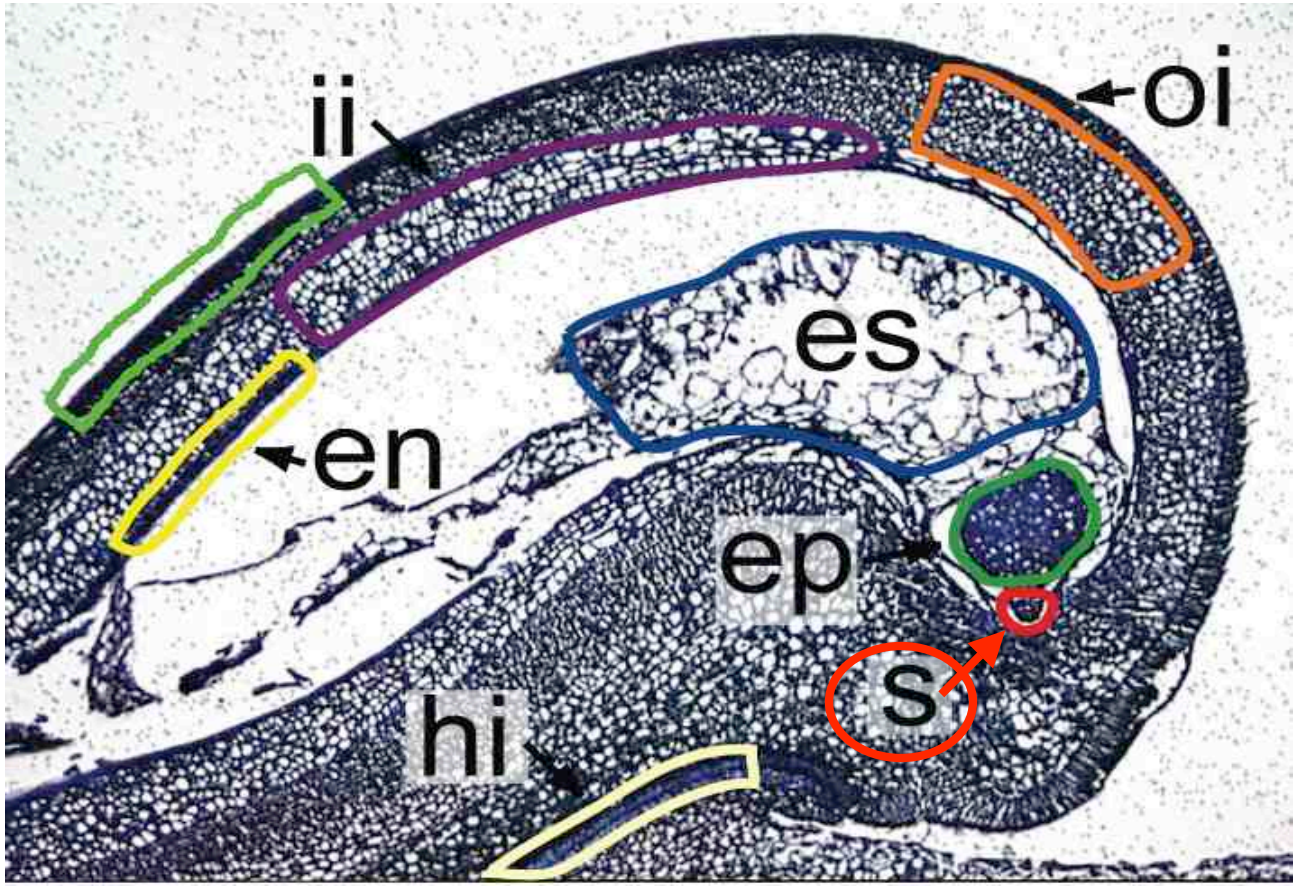


14,456
(17,576)

=

16,248
(19,755)

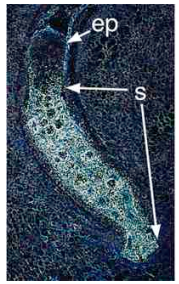
How Are Genes Activated in Different Seed Compartments Following Fertilization?



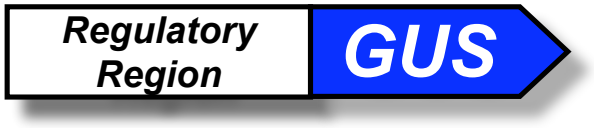
For Example…. in the Suspensor-an Embryonic Nurse Cell

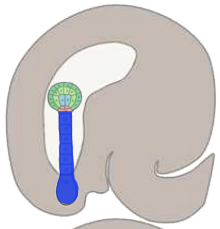
What Are the DNA Regulatory Sequences Important for Activating Transcription in the Suspensor?

Suspensor-Specific Gene

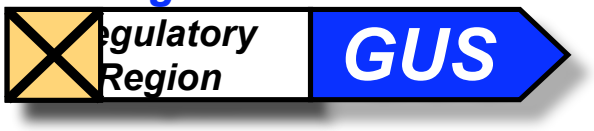


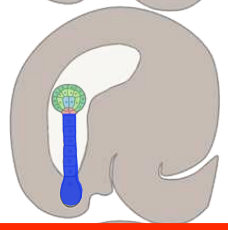
Fuse Regulatory Region to GUS Reporter Gene

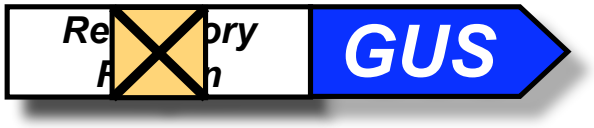


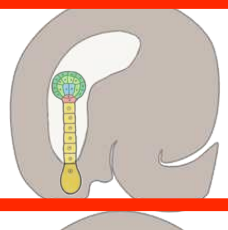


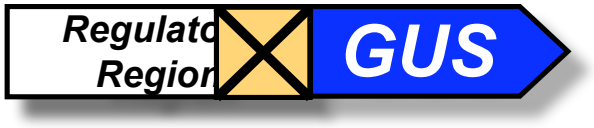
Mutated Region

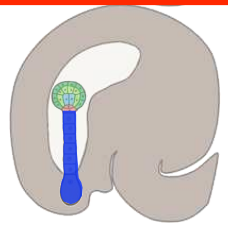




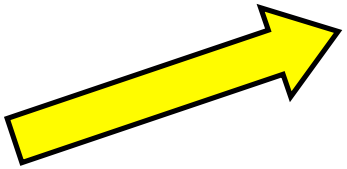


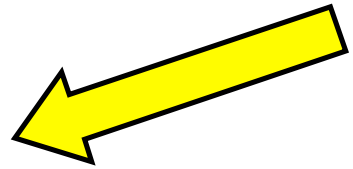






Transformation



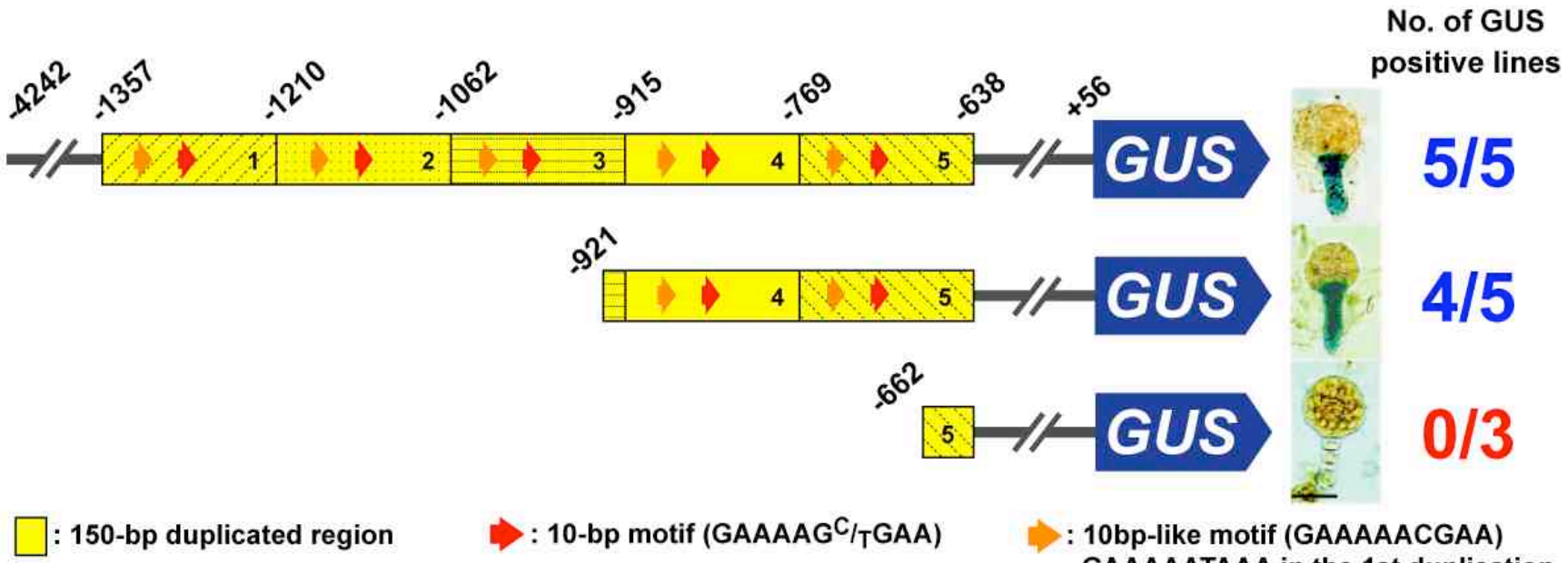


Observe Promoter Activity

Tobacco Plant

Mutation Destroyed Suspensor Control Sequence!!!

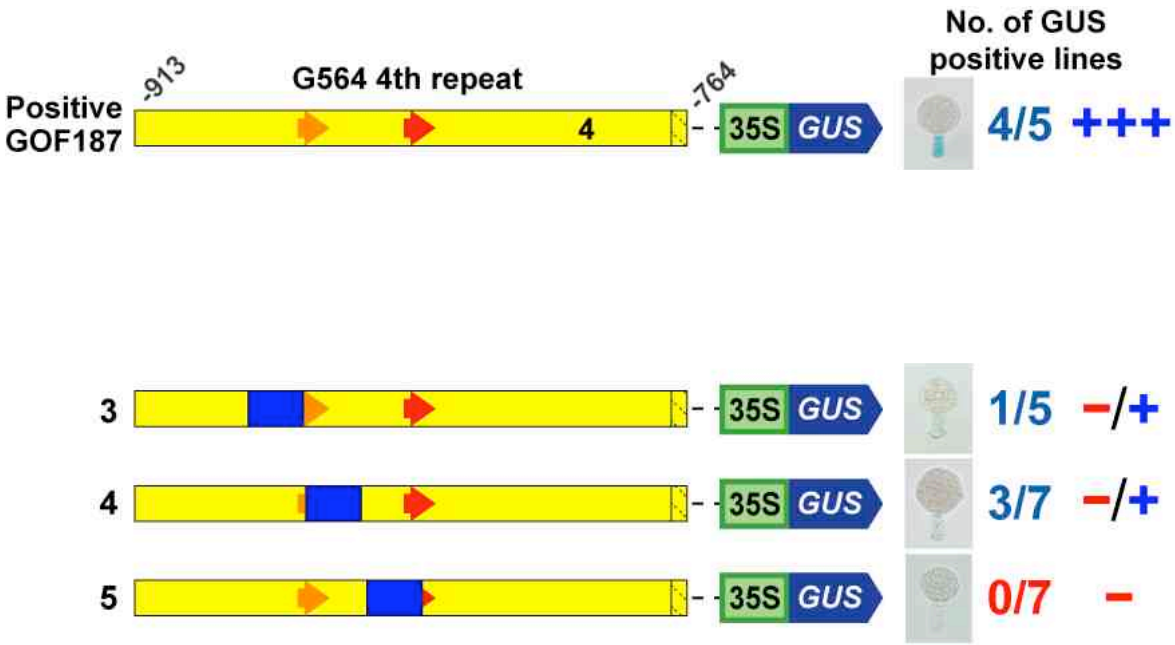
What Region in the G564 Upstream Sequence Is Important for Transcription in the Suspensor?



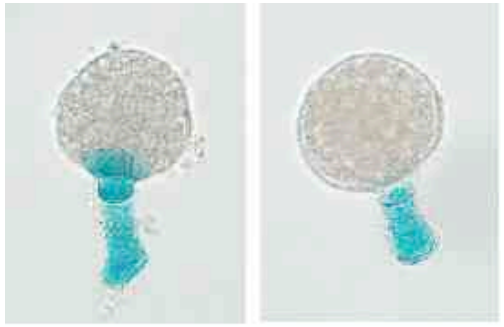

```
1 -1357 AAAATTTATAACACTTTTATGCTG TGTTTGTTTC GAAGCATAGAAAAATAAAAAGTTATTGTTGGGAAAGAAAAGTGAACAAAATCATGTAATAAAAA CAAA    ATGACACGACAATCAAAAAAAAAGTTTTCATGCAAAACTTTTTTC
2 -1210 AAAATTTA   CACTTTTATGATG TGTTTGTTTC GAAGTGTAGAAAAACGAAAAGTTATTATTGGTAAAGAAAAGCGAACAAAATCACGTAATAAAAA CAAAGCAAGATGGCACGACAATCAAAAAAAA GTTTCTACACAAAACTTTATTC
3 -1062 AAAATTTACAACACTTTTATGTTGTTGTTTGTTTCCGAGGTATAGAAAAACAAACAATTAGTGTTGGTAAAGAAAAGTGAACAAAACCATGTAATGAAAA CAAA    ATGGCACGACAATCAAAAAAA  GTTTTCACGCAAAATTTTCTTC
4 -915  AAAATTTATAACATTTTCATGTTG TGTTTGTTTC AAAGCCTAGAAAAACGAAAAGTTACTATTGGTAAAGAAAAGCGAACAAAACCACATAATAAAAA CAAA    ATGGCACGACAATCAAGAAAAA GTTTTCACACAAAACTTTTTTC
5 -769  AAAATTTACTA          TGTTTATTTC GAAATTTAGAAAAACGAAAAGTTATTATTAGTAAAGAAAAGCGAACAAAACTACGTAATAAAAAACAAA    ATGGCACGACAAT AAAAAAA  GTTTTCACGCAAAATTTTCTTG
```

- *Transcriptional Activity in the Suspensor Is Abolished When the Upstream Sequence Is Deleted to -662.*
- *Transcriptional Activity in the Suspensor Is Abolished When All 150-bp Duplications Are Deleted.*

What Regions in the 4th Duplication Are Important for Suspensor Transcription?



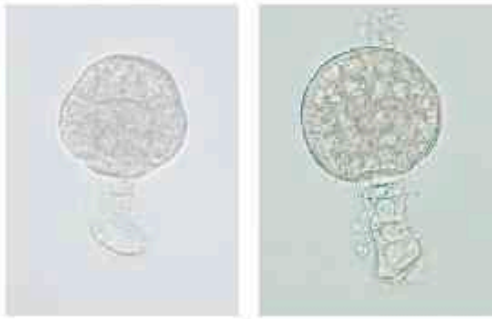
Positive (+++)



3 (−/+)

4 (−/+)

5 (−)



: 150-bp duplicated region

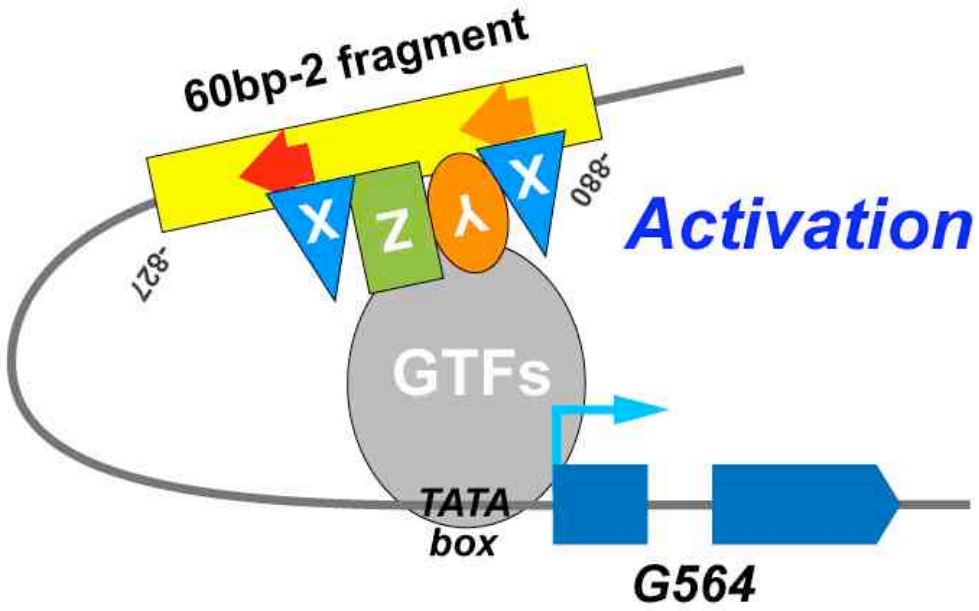
: Replacement sequence (GGCCGCGGGGGGCCC)

: 10-bp motif (GAAAAGCGAA)

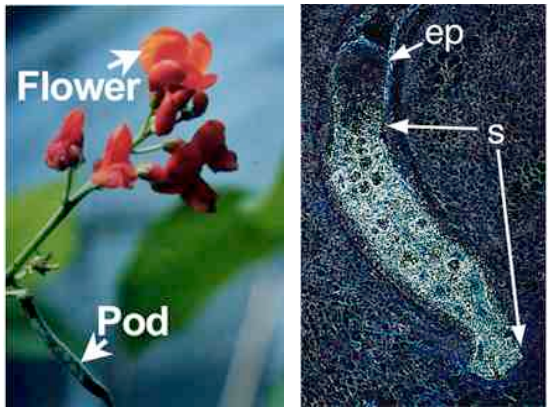
: 10bp-like motif (GAAAAaCGAA)

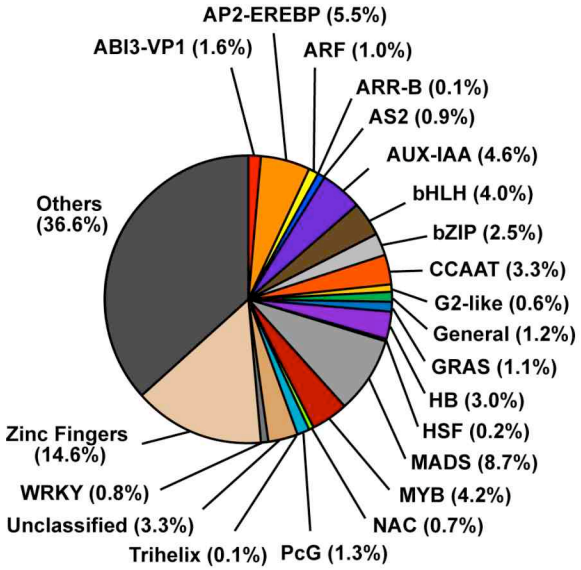
A Model of G564 Transcription in the Suspensor

: 150-bp duplicated region
: 10-bp motif (GAAAAGCGAA)
: 10bp-like motif (GAAAAaCGAA)
: general transcription factor complex
: factor X binding to the 10-bp and the 10bp- like motifs
: factor Y binding to unknown sequences
: factor Z binding to unknown sequences



SRB G564

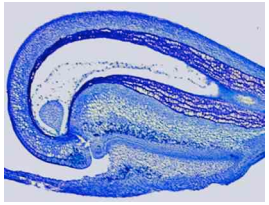


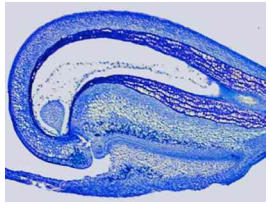


Transcription factors found in SRB SUSPENSOR
(N = 3,107)

Summary: "How To Make A Seed"

- *At Least 15,000 Diverse mRNAs Are Present in Each Seed Compartment, Region, and/or Tissue*
- *At Least 22,000- 24,000 Diverse mRNAs Are Present in a Seed as a Whole Depending on the Stage (i.e., Genes Required to "Make a Seed")*
- *At Least 26,000 Diverse mRNAs Are Required to Program Seed Development*
- *Most Diverse mRNAs are Shared by Different Compartments, Regions, and Tissues -- Many Are Quantitatively Regulated*
- *Each Compartment Region, and Tissue Has a Small Set of "Specific" mRNAs, Including Those Encoding Transcription Factor mRNAs*
- *Sequences Are Beginning to Be Identified That Activate Transcription in Different Seed Regions*

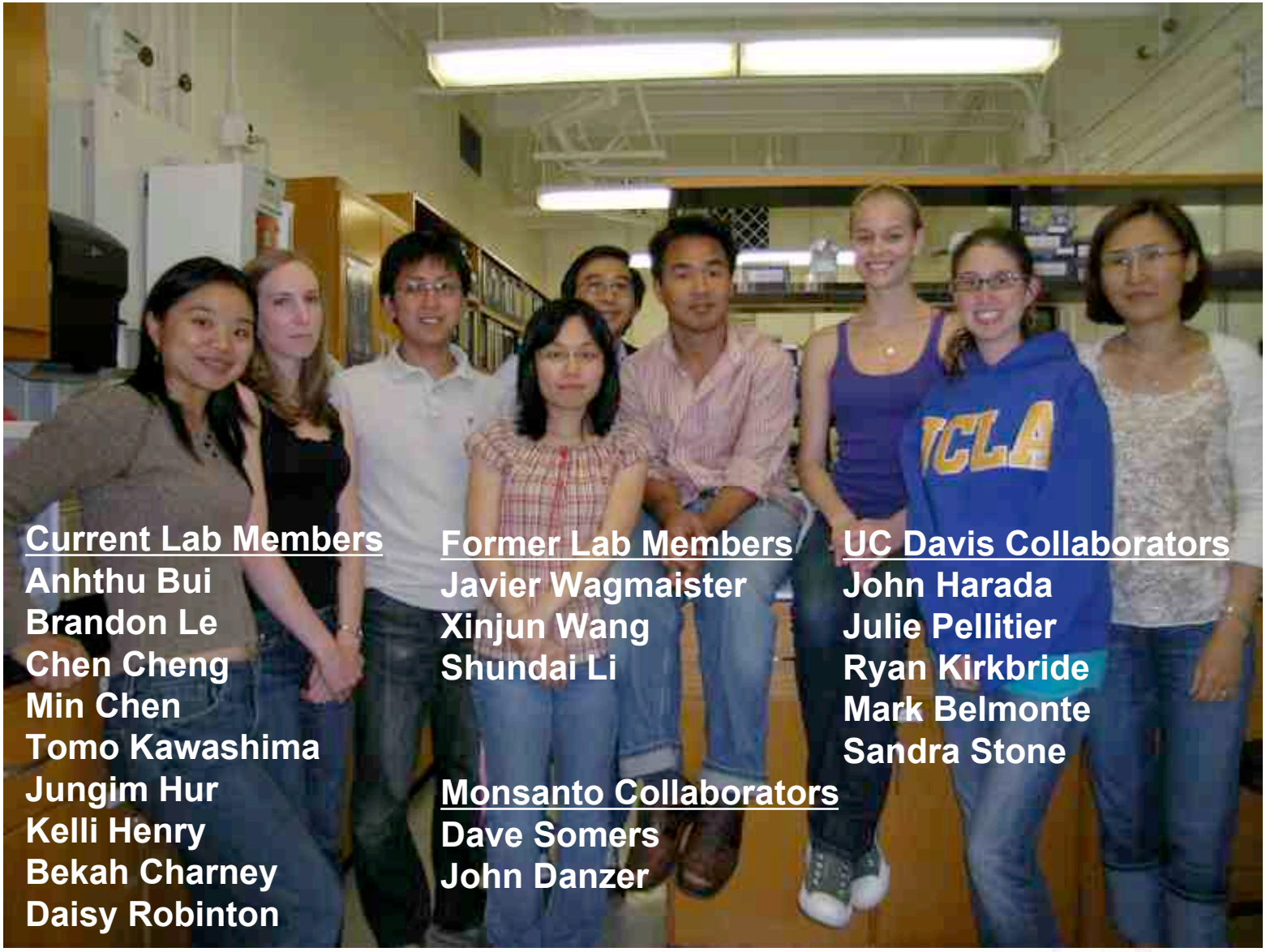




The End……..or Is It the Beginning?

A Giant Seed!

GOLDBERG LAB



Funded By an NSF Plant Genome Grant
To Bob Goldberg and John Harada


How Will Seed Yields Be Increased in the Future?
As We Always Have…….By Using the Best Cutting-Edge Science!!!

Plant Genome Projects Are Identifying Genes Essential For Increasing Crop Yields!!

Plant Genomes Sequenced To Date

- *Arabidopsis*
- *Rice*
- *Poplar Tree*



- *Soybean*
- *Corn*
- *Medicago*



- *Papaya*
- *Grape*
- *Castor Bean*



These Genes Will Help Increase Food Production Significantly in the 21st Century To Feed Our Growing Population

Yield (Developmental Traits)

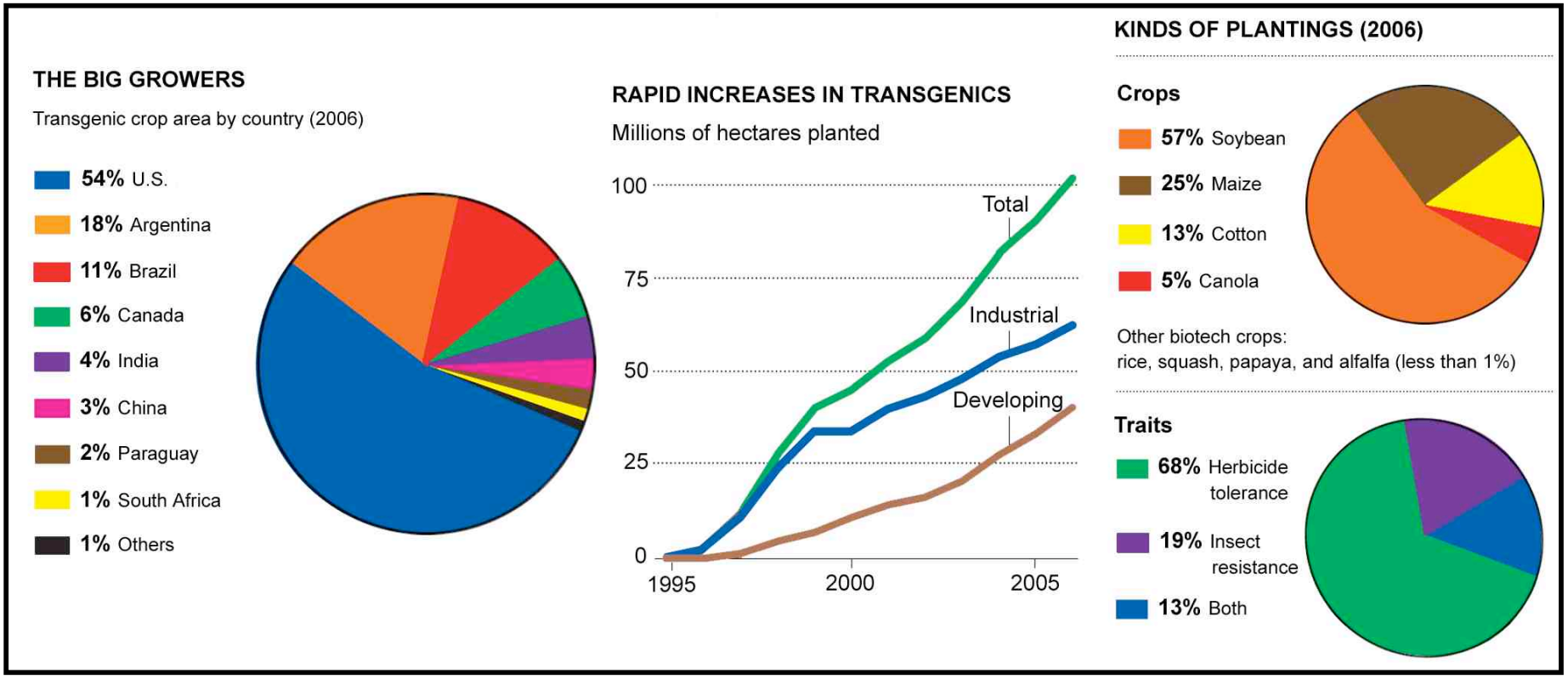
- ***Seed Number***
- ***Seed Size***
- ***Growth Rate***
- ***Organ Size (More Seeds)***
- ***Plant Architecture***
- ***Flowering Time***
- ***Senescence***
- ***Maturity***
- ***Stature***

Yield (Stress Traits)

- ***Nutrient Uptake***
- ***Drought Resistance***
- ***Heat Resistance***
- ***Cold Tolerance***
- ***Salt Tolerance***
- ***Shade Tolerance***
- ***Disease Resistance***

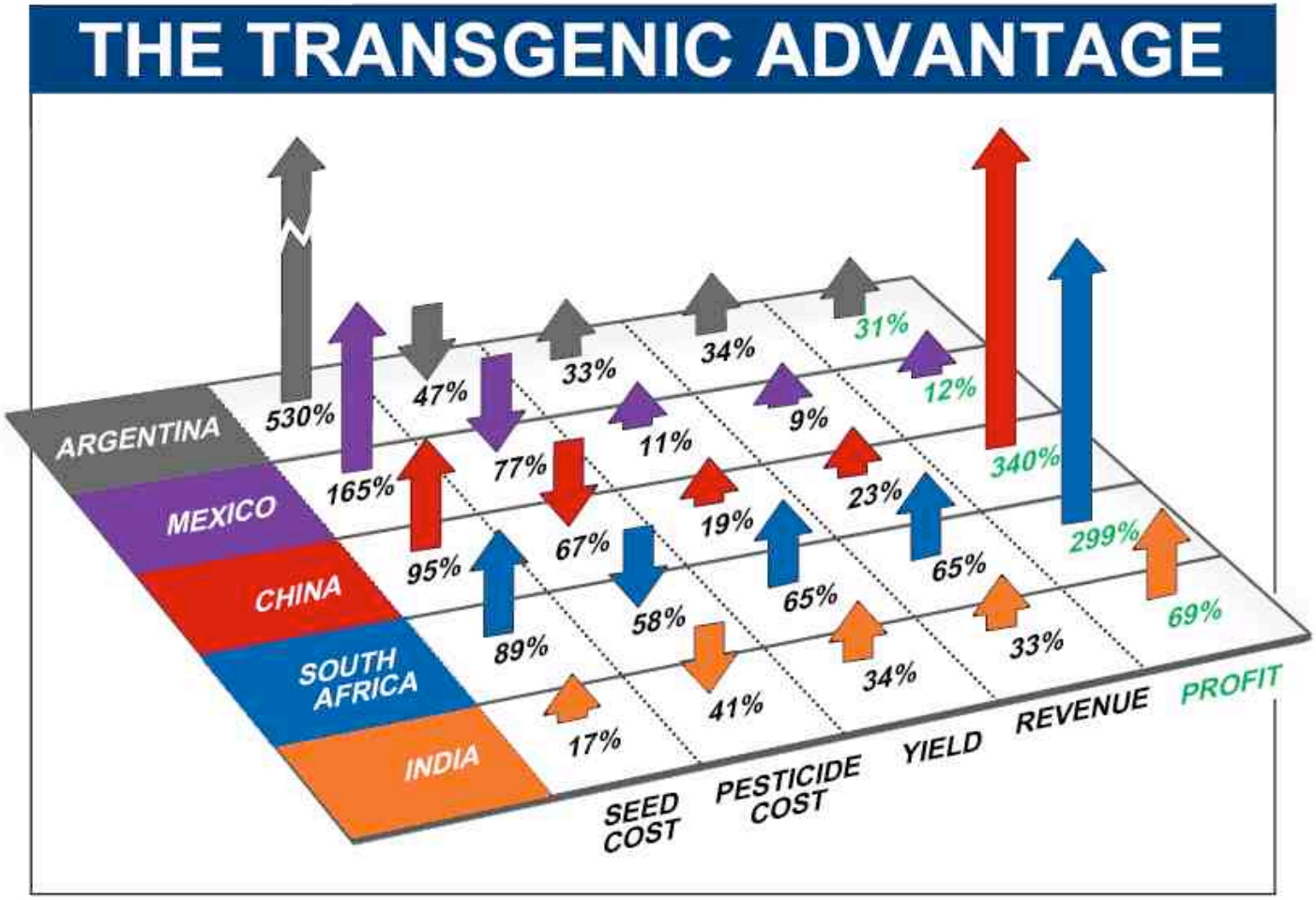
Rice

One Way is to Use These New Traits in Engineered Crops That Farmers Have Adopted Faster Than Any New Agricultural Technology In the Past 100 Years!



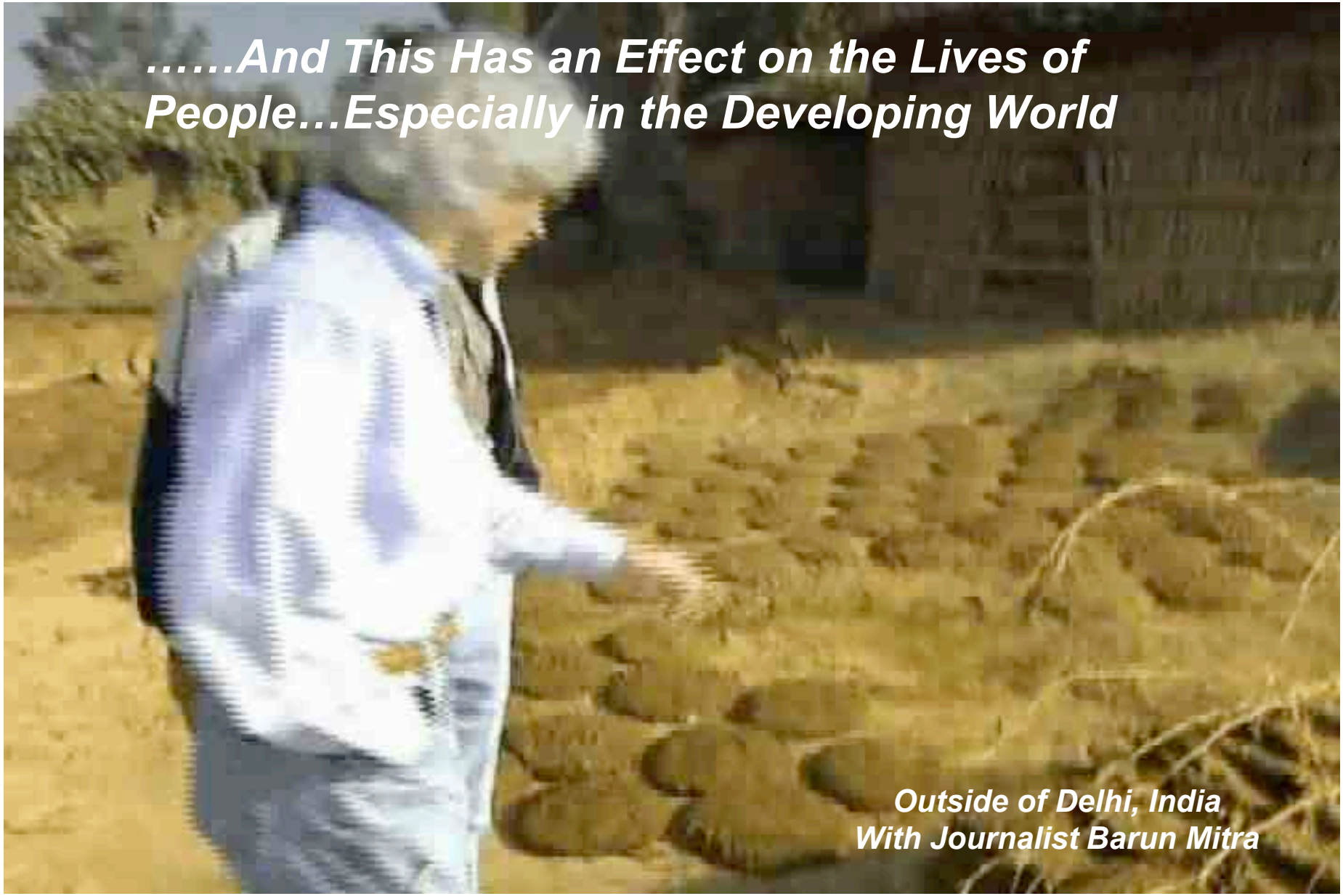
Over One Billion Acres of Bioengineered Crops Have Been Grown World-Wide Since 1996 and 250 Million Acres in 2007

Engineered Crops Have Increased Yields, Reduced Pesticide Use, and Increased Incomes of Farmers in the Developing World



United Nations FAO Report No. 35, 2003-04; Scientific American, September, 2007

However…There's a Battle Raging to Get Bioengineered Crops Adopted in Many Parts of the World


……And This Has an Effect on the Lives of People…Especially in the Developing World
Outside of Delhi, India
With Journalist Barun Mitra

Professor Frank Furedi, University Of Kent, England

The End……..or Is It the Beginning?

A Giant Seed!