The future of grape breeding: theory and technology

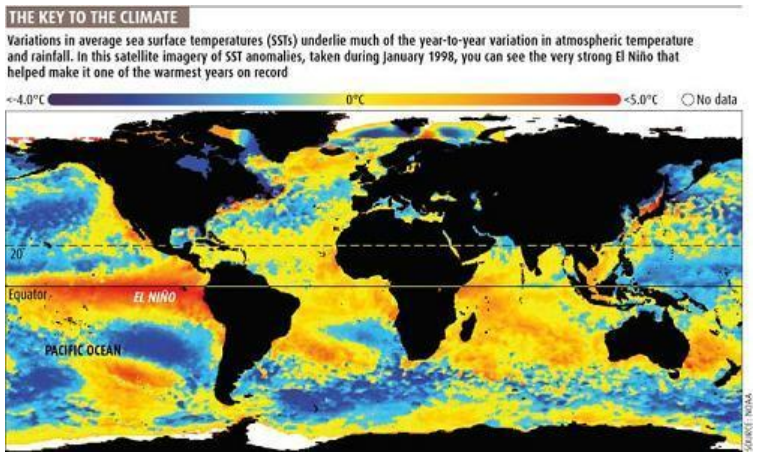
Zhenchang Liang

ZL249@ibcas.ac.cn

Insitute of Botany, Chinese Academy of Sciences

XII GBG, Bordeaux, 07/17/2018

Challenges


THE KEY TO THE CLIMATE
Variations in average sea surface temperatures (SSTs) underlie much of the year-to-year variation in atmospheric temperature and rainfall. In this satellite imagery of SST anomalies, taken during January 1998, you can see the very strong El Niño that helped make it one of the warmest years on record
<-4.0°C
0°C
>5.0°C
No data
20°
Equator
EL NIÑO
PACIFIC OCEAN


中国气象网
CMA.GOV.CN


中华现代农业技术服务中心

The ways to deal with challenges

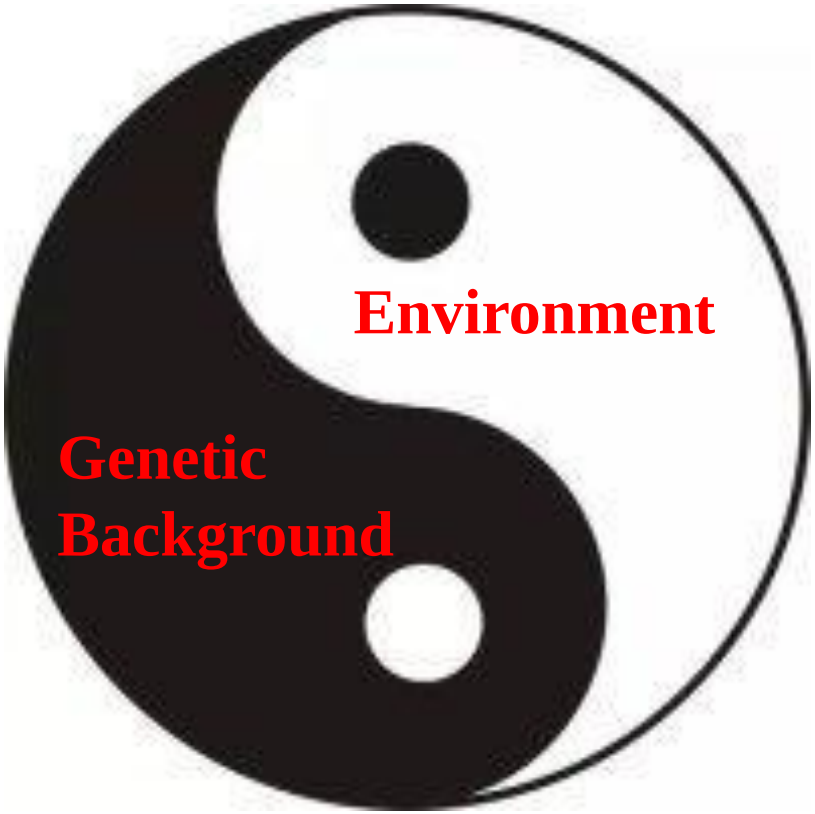
The problem…..

- High cost, e.g. buried and dug fee accounted for 50% total cost (-2,000$/Ha in north of China)
- Low quality, decrease yield, berry quality *et. al*.
- Food safety, pesticide residue

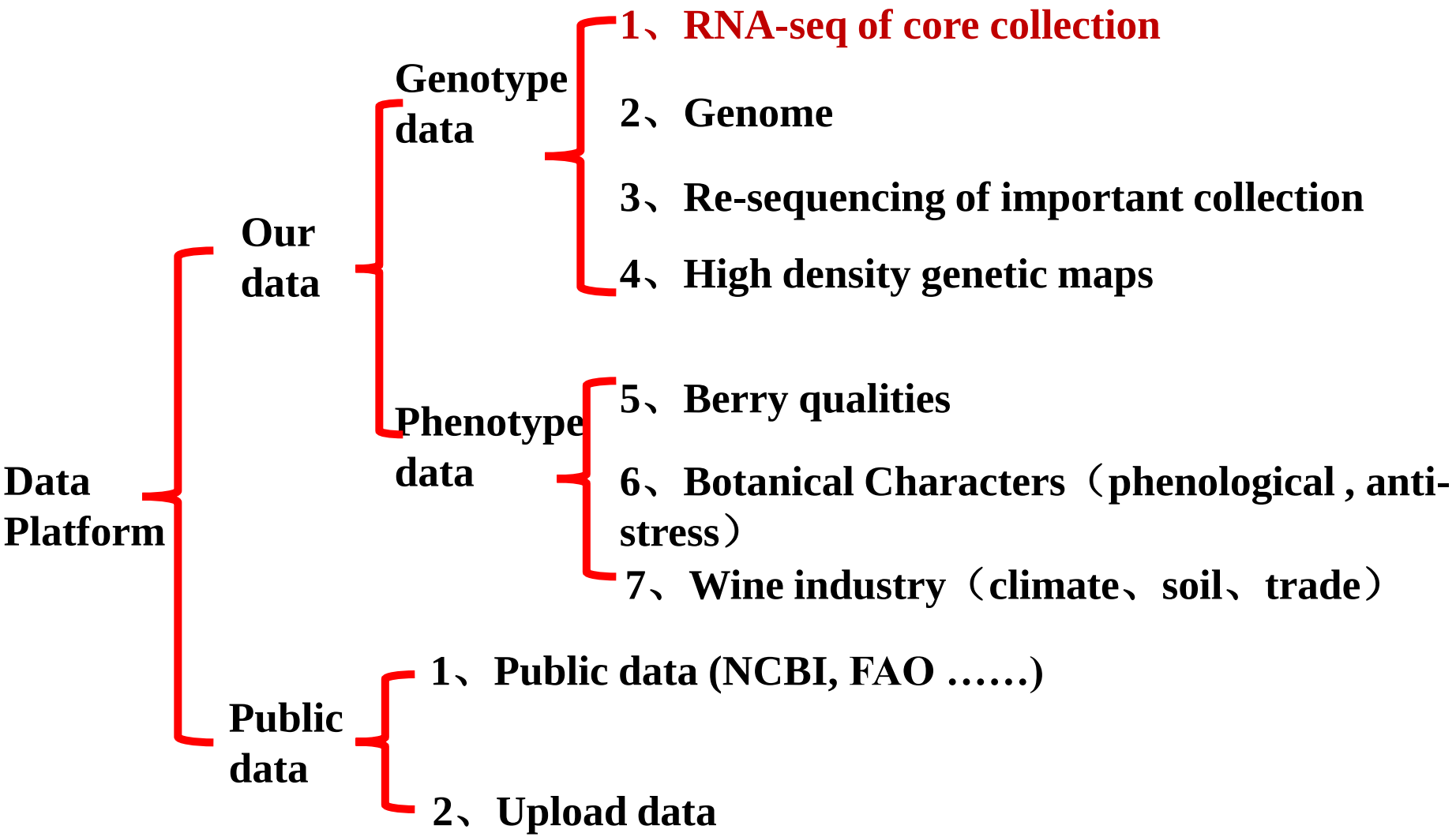
……

The improvement of crop system can not fundamentally address these problem.

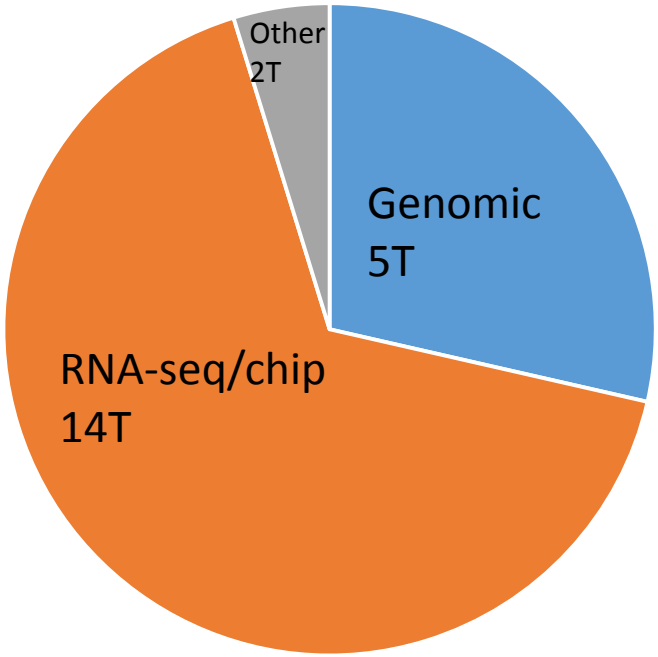
Harmonious growth and good phenotypes



To elucidate grape genetic diversity, we are building a grape data platform

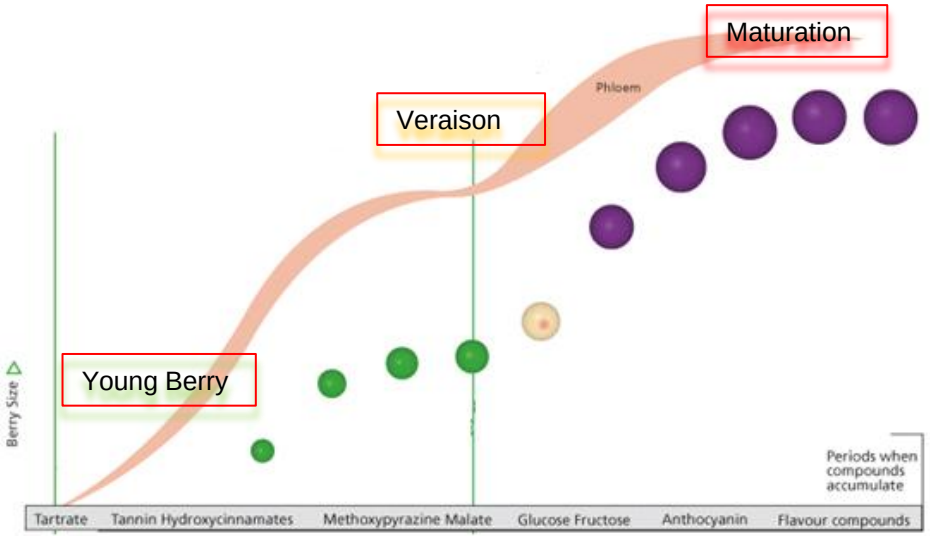


www.grapeworld.org



The English version is on the way…….

Data 1: RNA-seq of core collection (51 accessions，1.5T)



Stages	Species	color	Seedless	Purpose	Sugar	Acids	Berry.Size	Flavour
Young Berry	8 species	Colored	Seeded	Table	High	High	Big	High
Veraison	5 hybrids	White	Seedless	Wine	Low	Low	Small	Low
Maturation	1 unknown			Rootstock				

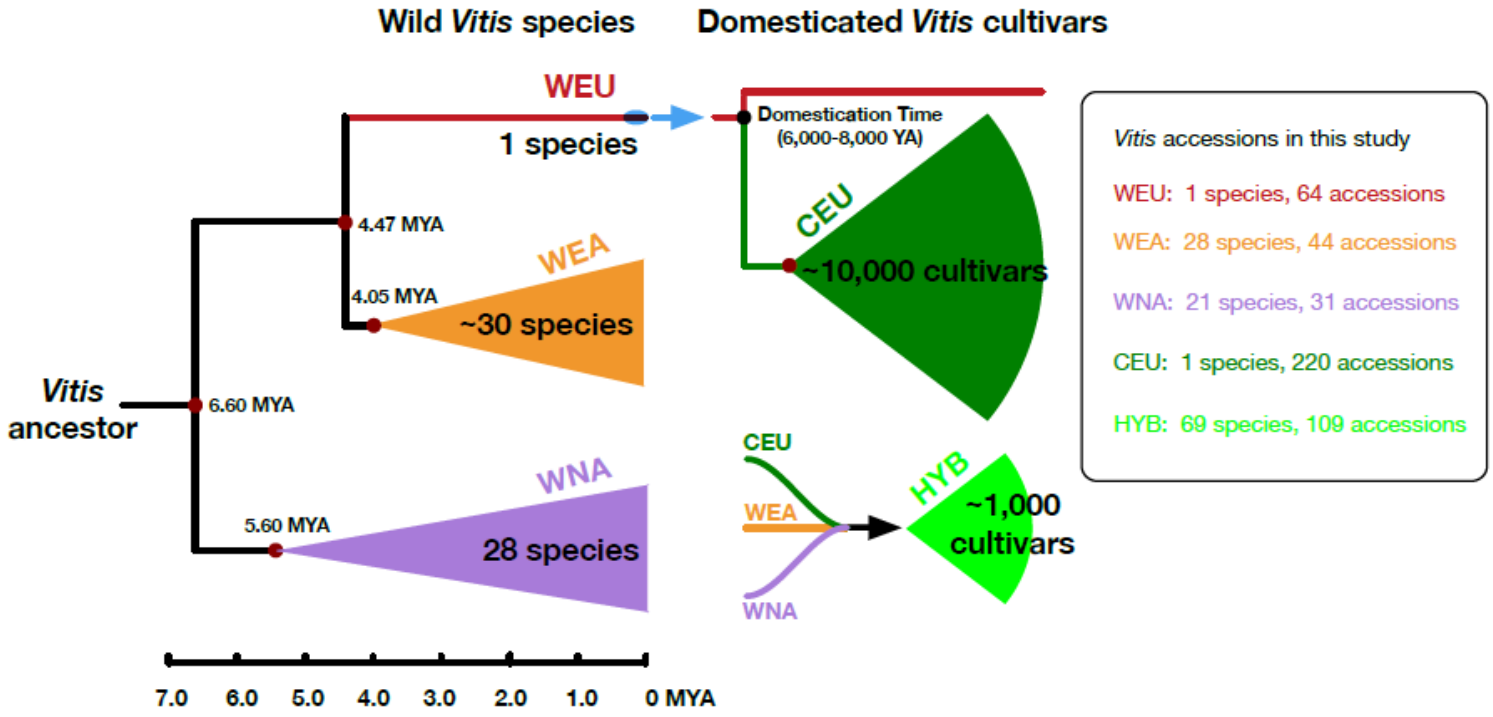
Data 2: Genome Data (*V. amurensis*)

Why is *V. amurensis*?

- **Originate in North-east Asia**
- **High anti-cold and disease**
- **Close to *V. vinifera***
- **Used in production**

Hiseq2500 (200X)+Pacbio (30X)+Bionano（>150KB, 200G)+HiC (100X)

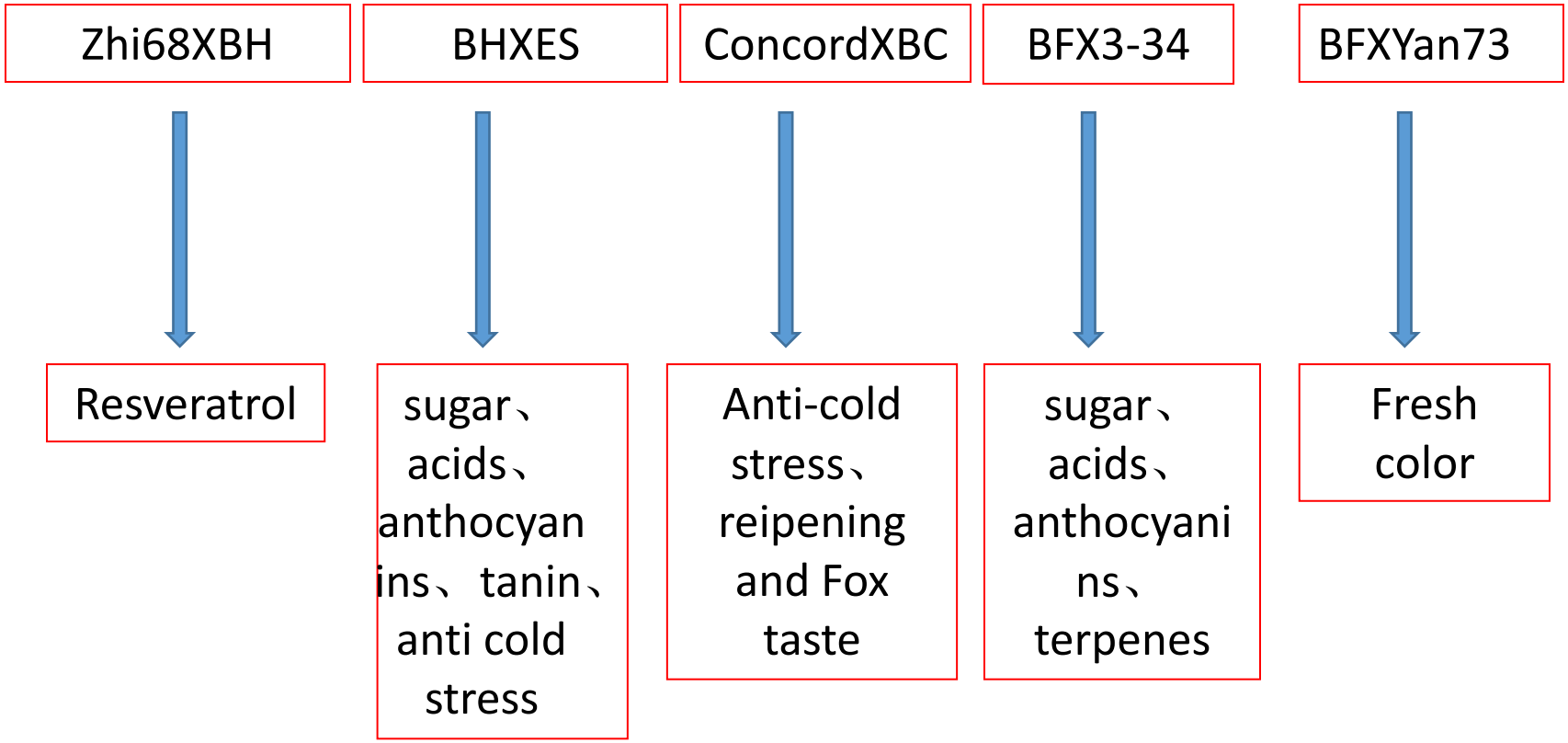
Data 3: Resequencing (472 accessions, Coverage >15X)



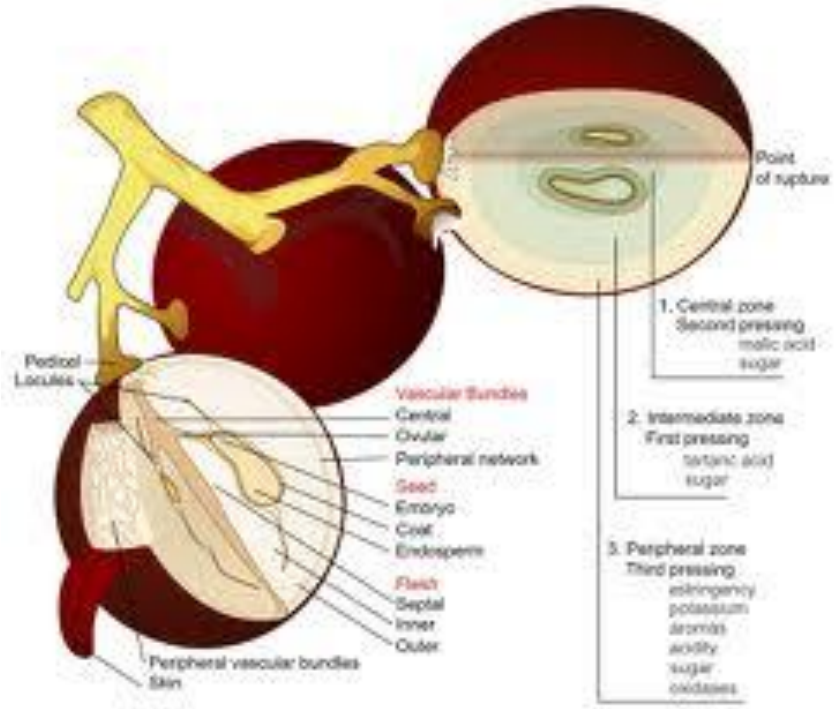
wild Eurasian species **WEU**(Group-1),
domesticated grapevine cultivars **CEU** (Group-2),
interspecific-hybrid grapevine cultivars were classified into two separate clusters: Group-3 and Group-5 **HYB**
wild North American *Vitis* species **WNA** (Group-4)
wild East Asian *Vitis* species **WEA** (Group-6),

Data 4: High density genetic maps

- Methods: RADseq or GBS
- High quality SNP: >2000



Data 5: Berry traits (>200 accessions)



Outer of berry：

shape、color、weight；

Metabolites：

sugar、acids、volatiles、polyphenols、ascorbic acid、animo acids etc.

Including：variety、data、people、time、position、methods

Data 6: Botanical traits

Phenology：

growth potential、budding、

flowering、fruit setting 、

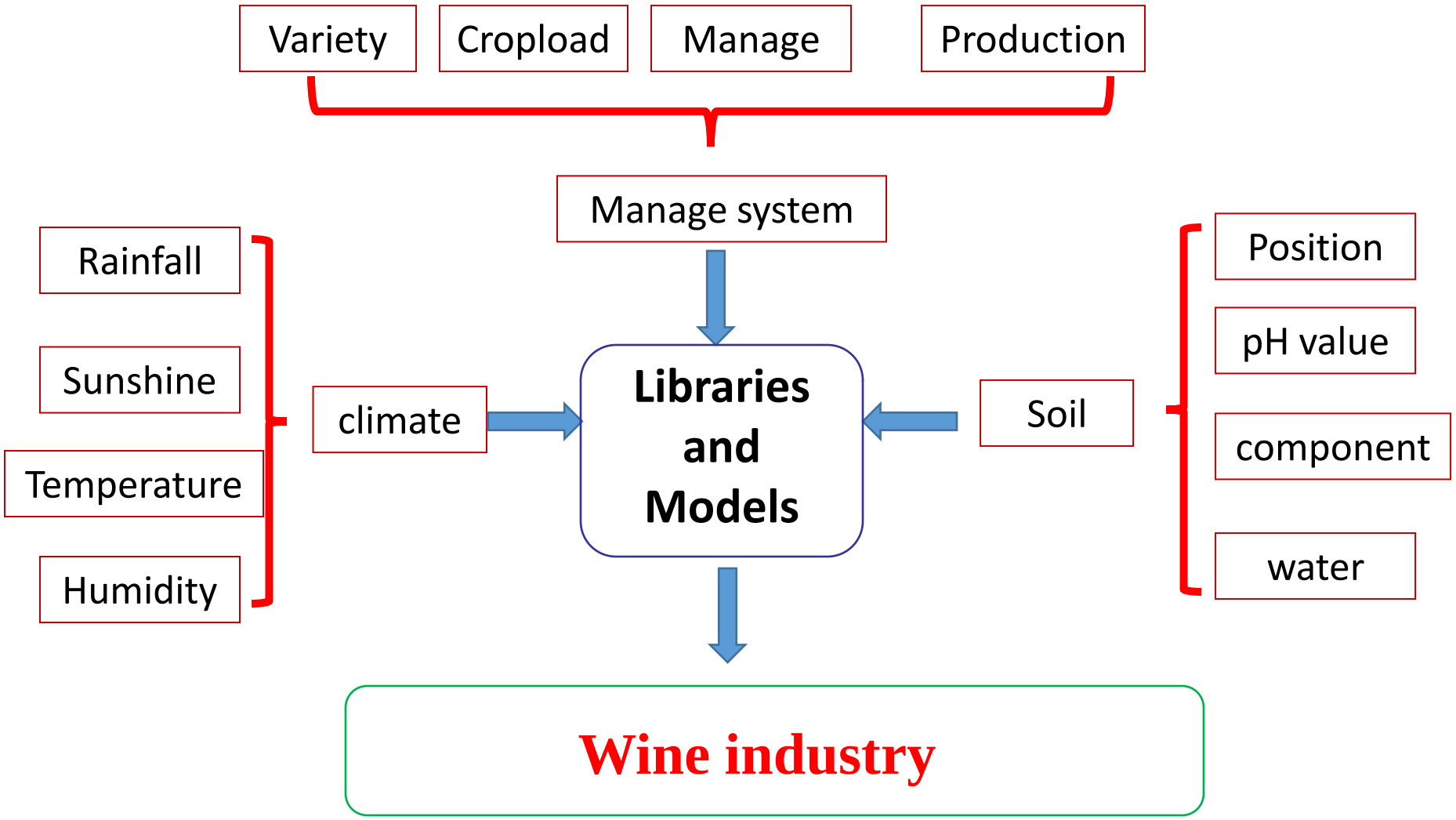
verasion、ripening etc；

Anti stress：

cold、heat、drought、

disease etc.

Data 7: Wine industry



Data 8: Public data

142 plant genome including 4 grape genome

1992 grape RNA-seq (A grape transcriptome database http://grapeworld.org/gt/)

136 grape microRNA

24 grape DNA methlylation

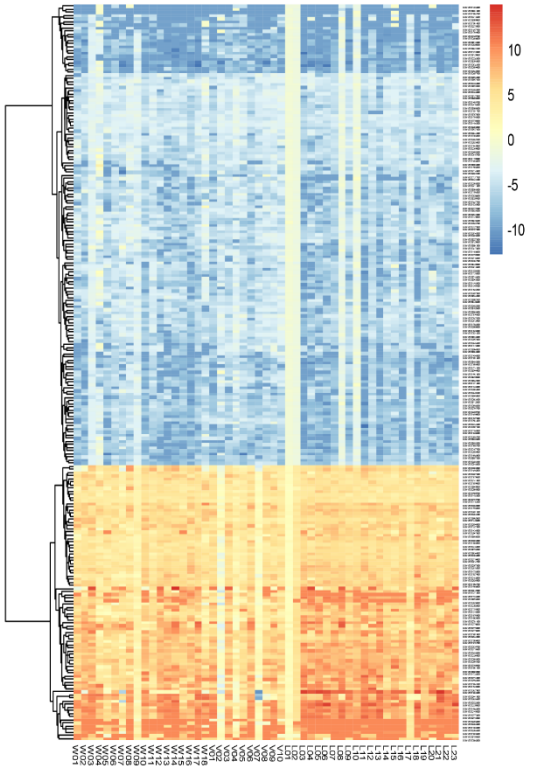
Grape industry data

Developing many new DNA markers/candidate genes linked with traits. e.g.

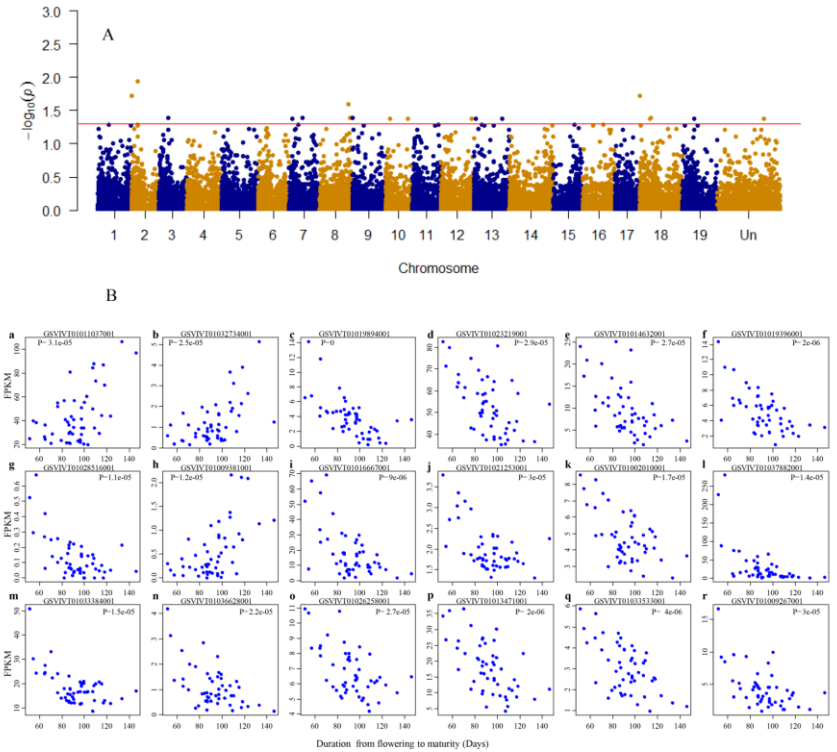
- Many candidate genes linked with onset of ripening

Comparative transcriptomics analysis (GH ***VS*** Veraision):
2077 DEGs in 80% accessions,
1482 down-regulated, 595 up regulated

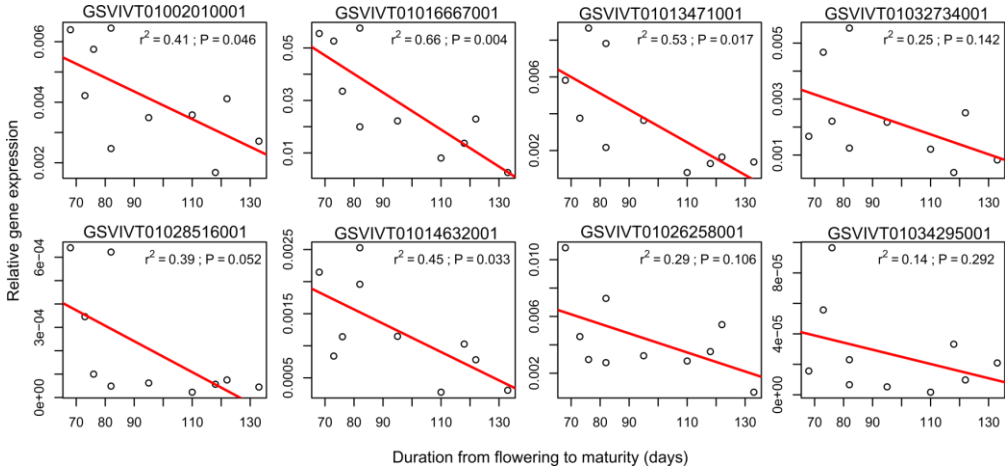
325 HDEGs（>3fold）147 down-regulated (related to growth and development), 88 up-regulated (related to metabolism and stress responses)



Expression association study (EAS)



The relationship between gene expression level and the time for flowering to maturation

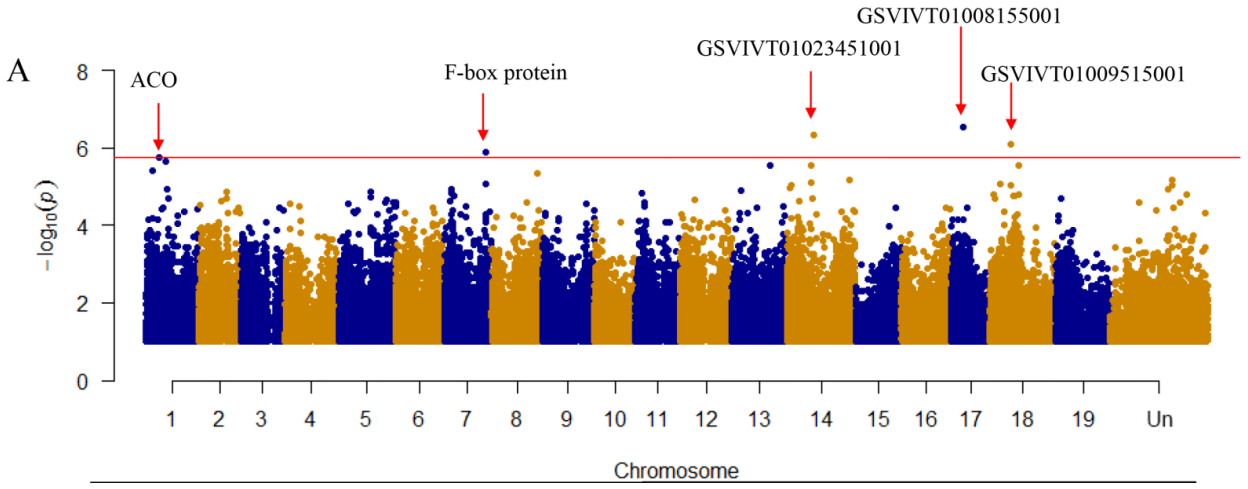


The 8 candidate genes qRT-PCR results in 10 other accessions

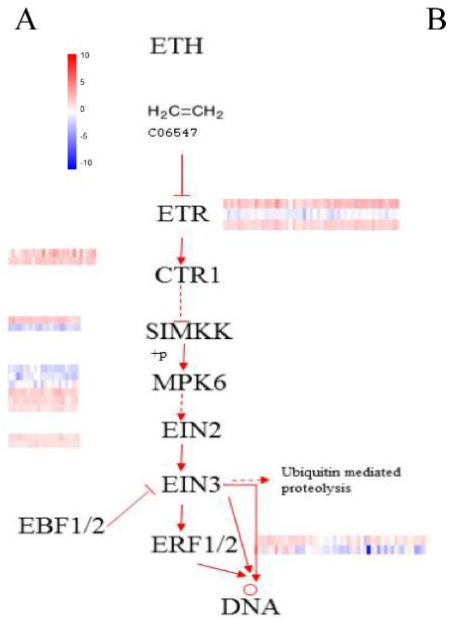
The 18 candidate genes were identified, 13 were down-regulated and 5 of them were related cell wall modification

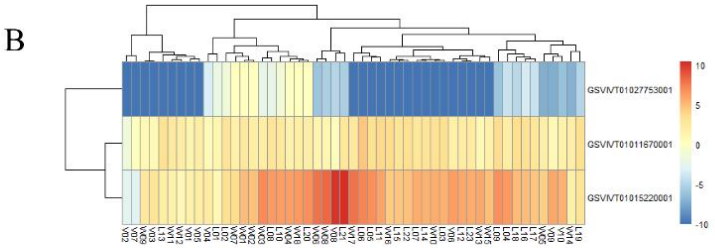
The eGWAS results showed that there were 5 loci related to the flowering-to-maturity duration. These high-effect loci were located in 5 genes

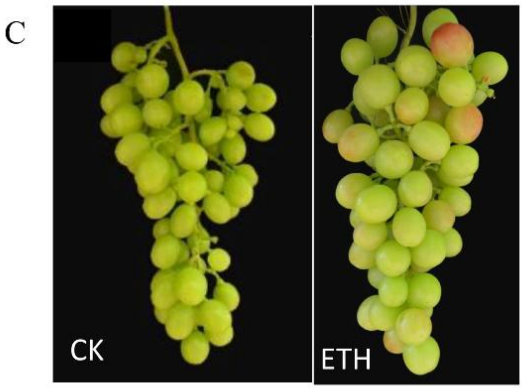
The ethylene sigal pathway showed significantly up-regulated, and ethephon could color berries earlier.

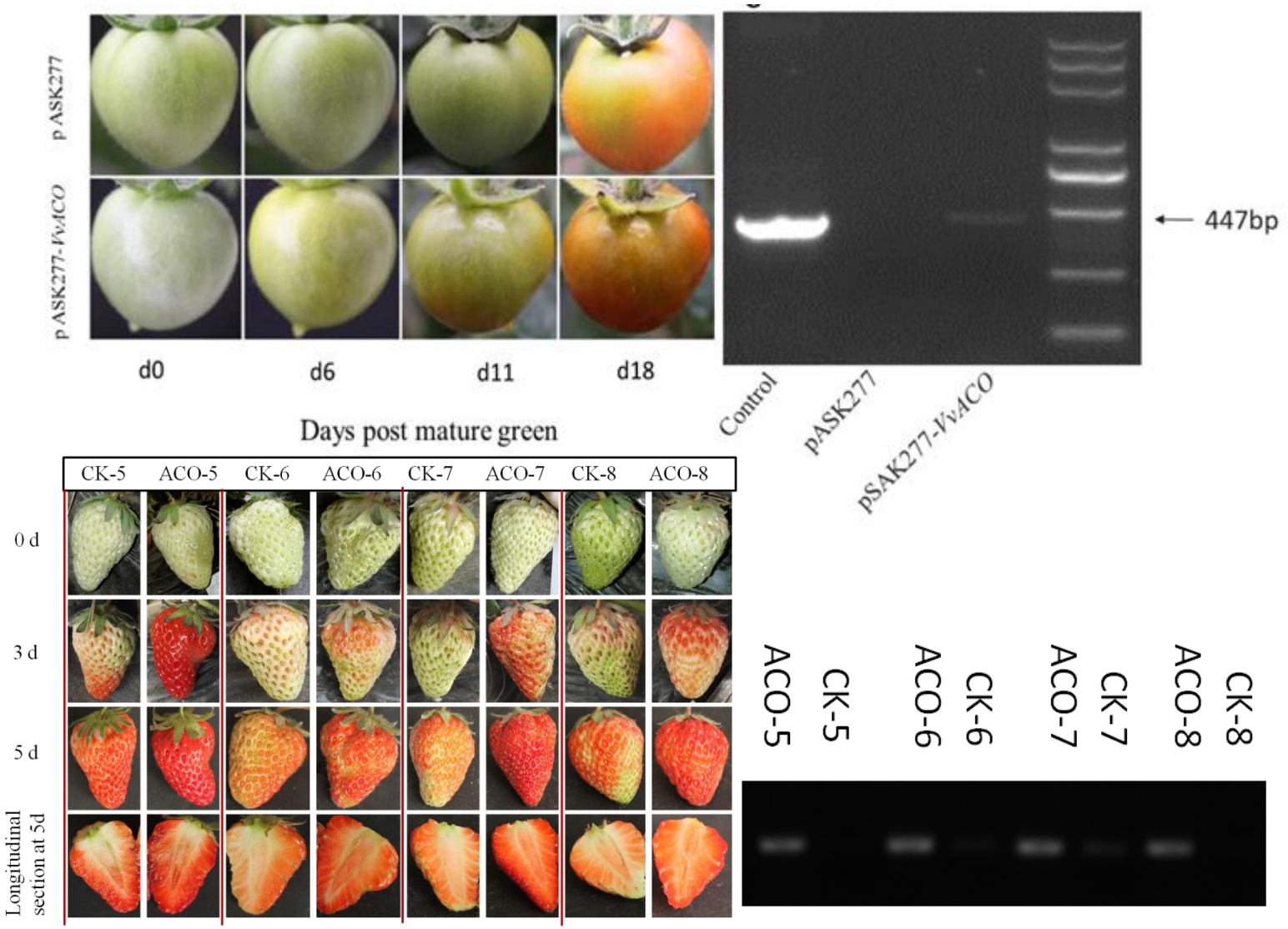


GWAS results with exome SNPs and the flowering-to-maturity



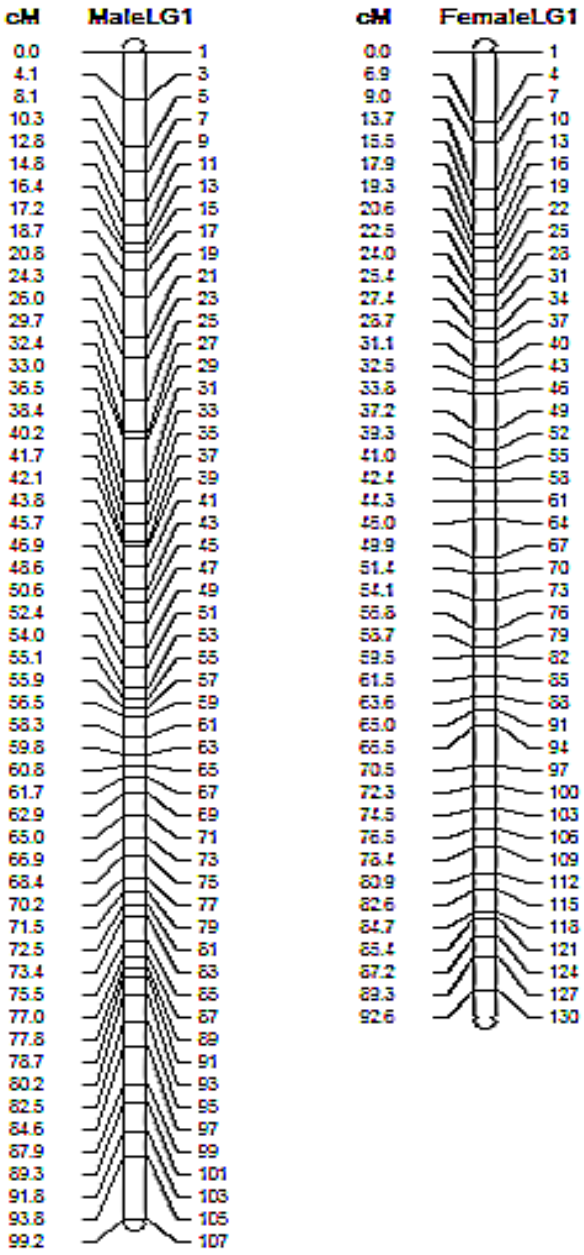






The transient over-expression of *VvACO* in strawberry (non-climacteric) and tomato (climacteric) accelerated the onset of ripening.

Mapping QTL+RNA-seq



Beifeng×3-34

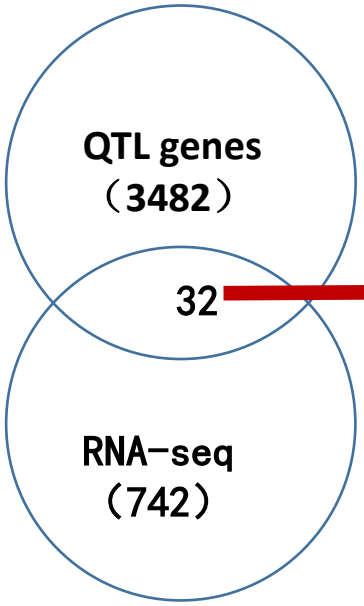
population：150

	BF	3-34	Total
Linkage group	19	19	
Map length	1799.1 CM	1974.6 CM	
SNP	2769	1921	4690
Density	0.65	1	

The number of QTLs in the genetic maps.

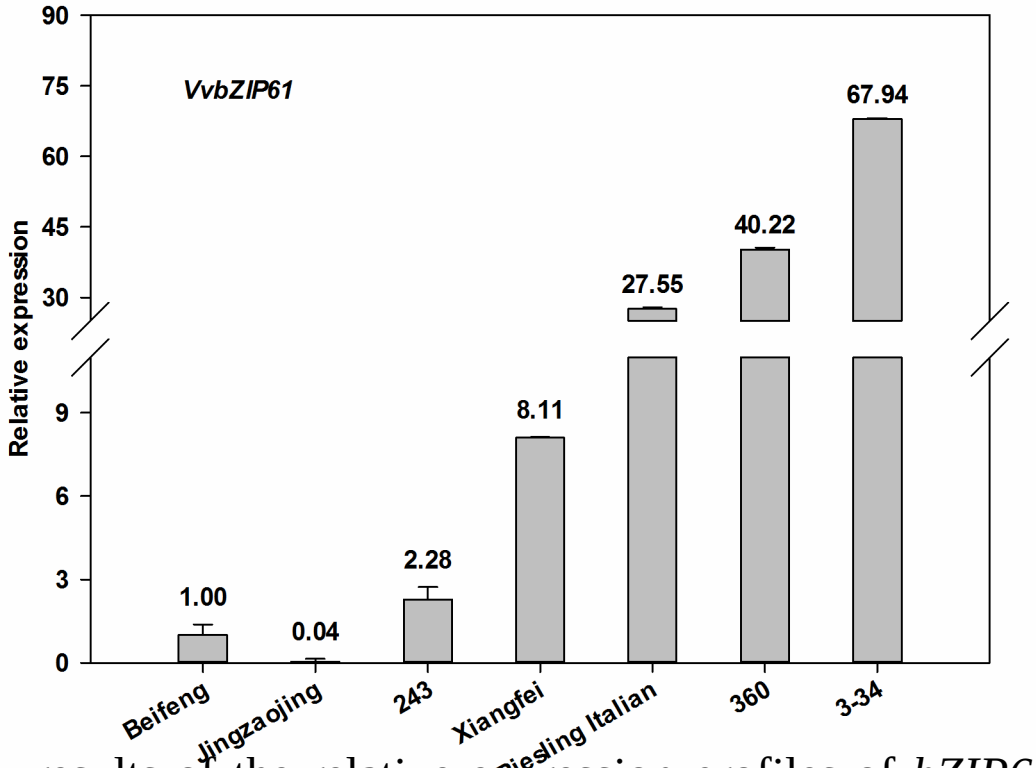
Compound	3-34	Beifeng	Total
Limetol	10	6	16
.beta.-Myrcene	13	3	16
D-Limonene	8	2	10
Ocimene	10	4	14
Linalool oxide	3	4	7
Terpinolene	9	3	12
Linalool	11	7	18
Rose oxide	17	4	21
Alloocimene	16	5	21
Hotrienol	2	nd	2
Nerol oxide	14	4	18
4-Terpinenol	1	nd	1
(E)-Pyranoid linalool oxide	1	3	4
α-Terpineol	16	6	22
Nerol	4	2	6
Geraniol	12	4	16
Geranial	5	nd.	5

Among these QTLs, 73 stable QTL related to monoterpenes were identified in two successive years.

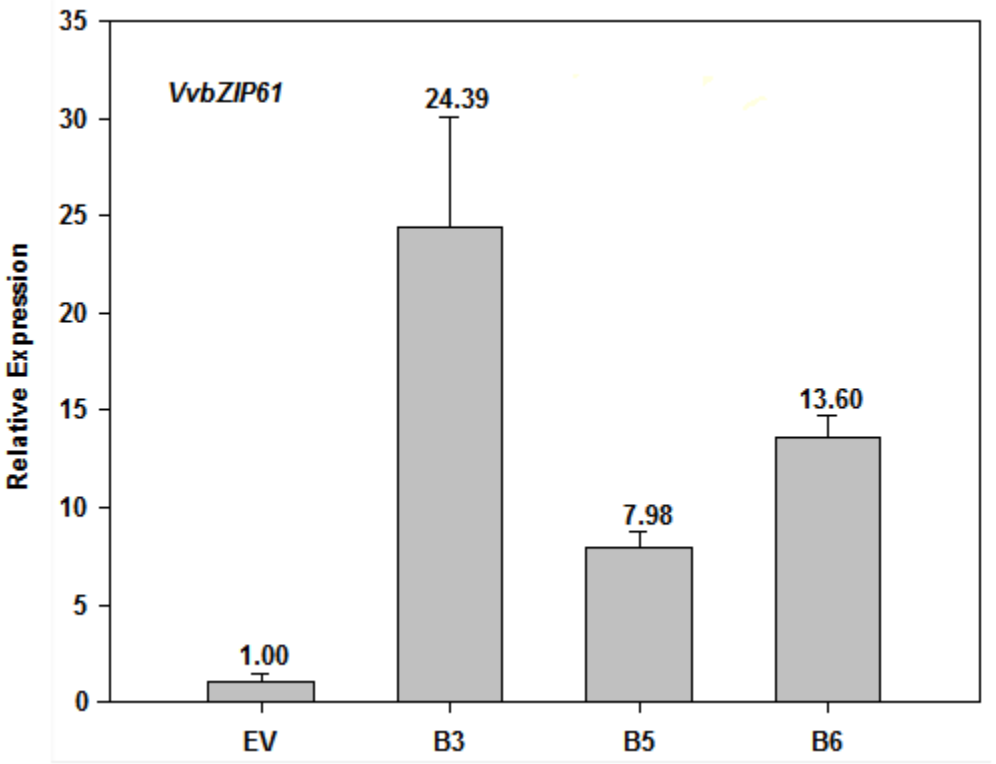


2TF，26 Construction genes，4 unknown

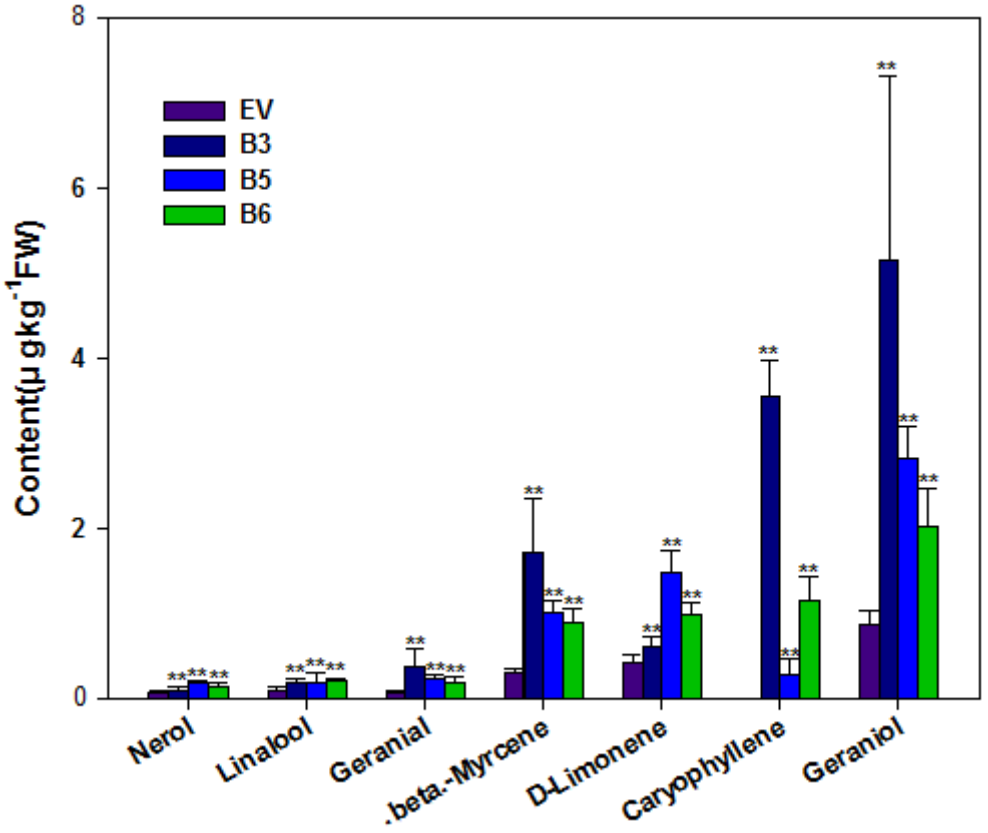
Gene ID	Chromosome	Position	Functional Annotation
GSVIVG01010274001	chr1	18696200-18697377	defensin Ec-AMP-D2
GSVIVG01019878001	chr2	4150396-4152247	NDP-L-rhamnose synthase
GSVIVG01022723001	chr2	15587960-15590452	Nitrilase 4B
GSVIVG01019873001	chr2	4110678-4113795	primary amine oxidase
GSVIVG01019849001	chr2	3822829-3825957	Osmotin
GSVIVG01018921001	chr4	18625063-18629091	dolichyldiphosphatase 1-like
GSVIVG01018767001	chr4	20090879-20092677	thaumatin-like protein 1b
GSVIVG01017899001	chr5	4484043-4487627	aluminum induced protein with YGL and LRDR motif-like
GSVIVG01017796001	chr5	3549320-3551930	UPF0497 family
GSVIVG01017718001	chr5	2877561-2880014	unknown
GSVIVG01017757001	chr5	3181985-3182859	protein RALF-like 33
GSVIVG01031486001	chr6	18028142-18029883	Xyloglucan endotransglucosylase/hydrolase 32
GSVIVG01031418001	chr6	18818715-18822205	UDP-D-apiose/UDP-D-xylose synthase 2
GSVIVG01025223001	chr6	3073920-3076705	Kiwellin Ripening-related protein grip22
GSVIVG01024994001	chr6	5418080-5420685	heat shock cognate 70 kDa protein 2-like
GSVIVG01037249001	chr6	16912617-16920281	magnesium transporter NIPA6
GSVIVG01024970001	chr6	5615788-5616779	GTP-binding protein OBGM
GSVIVG01037417001	chr6	14131230-14133129	calmodulin-7
GSVIVG01015991001	chr9	16881846-16883252	Co-chaperone-curved DNA binding protein A
GSVIVG01017125001	chr9	4405819-4406667	thionin-like protein 2
GSVIVG01022901001	chr12	18178938-18179555	glutelin type-A 3-like
GSVIVG01020584001	***chr12***	***4004274-4005147***	***Ethylene-responsive transcription factor 9***
GSVIVG01023236001	chr12	20811694-20814701	transcript variant X3
GSVIVG01020678001	chr12	3173093-3175003	trichome birefringence-like 39
GSVIVG01020658001	***chr12***	***3365914-3367561***	***basic leucine zipper 61***
GSVIVG01033020001	chr14	25100075-25101021	UPF0497 family
GSVIVG01018579001	chr16	12961189-12962599	acyl-coenzyme A oxidase 3
GSVIVG01028882001	chr16	17827579-17830662	protein TRANSPARENT TESTA 12-like
GSVIVG01008344001	chr17	3086514-3092680	unknown
GSVIVG01029329001	chr17	14425929-14430764	metal-nicotianamine transporter YSL7
GSVIVG01008617001	chr17	369676-378037	unknown
GSVIVG01036848001	chr18	18007180-18008566	unknown



qRT-PCR assay results of the relative expression profiles of *bZIP61*.The y-axis indicates the relative folds of gene expression compared with Beifeng, which has low content of monoterpenes and whose expression was standardized as 1; numbers on the x-axis represents various cultivars: A, Beifeng (one of parents in the F1 population, low content of monoterpenes); B, Jingzaojing (low monoterpenes); C, F1 progeny no. 243 (low content of monoterpenes); D, Xiangfei (high content of monotedrpenes); E, Riesling Italian (high content of monoterpenes); and F, F1 progeny no. 360 (high content of monoterpene).

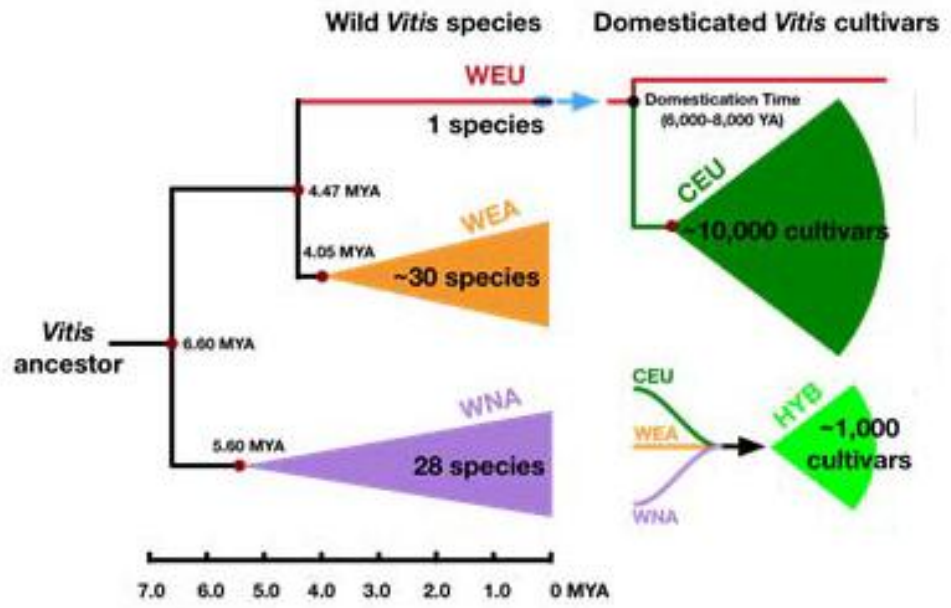


Expression of *bZIP61* in inflorescences of the empty vector (without *bZIP61* sequence, EV) and transgenic callus in different lines(B3,B5 and B6),the y-axis indicates the relative folds of gene expression compared with EV, which expression was standardized as 1.Error bars indicate SD of three biological replicates.



Monoterpenes(Nerol,D-limonene,beta.-myrcene,geranial and geraniol) and sesquiterpene(caryophyllene) content(μg kg-1 FW) from the empty vector (without *bZIP61* sequence, EV) and transgenic callus(from *Vitis amurensis* petiole) in different lines (B3,B5 and B6).EV Error bars indicate SD of nine biological replicates.B3 and B5 Error bars indicate SD of six biological replicates.B6 Error bars indicate SD of seven biological replicates.The mark "**" represent P<0.01,very significant difference.

Grape cultivars in industry



Rank	Cultivar	Color	Species
1	Cabernet Sauvignon	Red	*V. vinifera*
2	Merlot	Red	*V. vinifera*
3	Airen	White	*V. vinifera*
4	Tempranillo	Red	*V. vinifera*
5	Chardonnay	White	*V. vinifera*
6	Syrah	Red	*V. vinifera*
7	Garnacha Tina	Red	*V. vinifera*
8	Trebbiano Toscano	White	*V. vinifera*
9	Sauvignon Blanc	White	*V. vinifera*
10	Pinot Noir	Red	*V. vinifera*

The grape breeding status in last 20 years

	Total	Vinifera	Vinifer/La brusca	Other speices	Ratio of V/LV
World	212	123	51	38	82%
China	75	39	21	15	80%

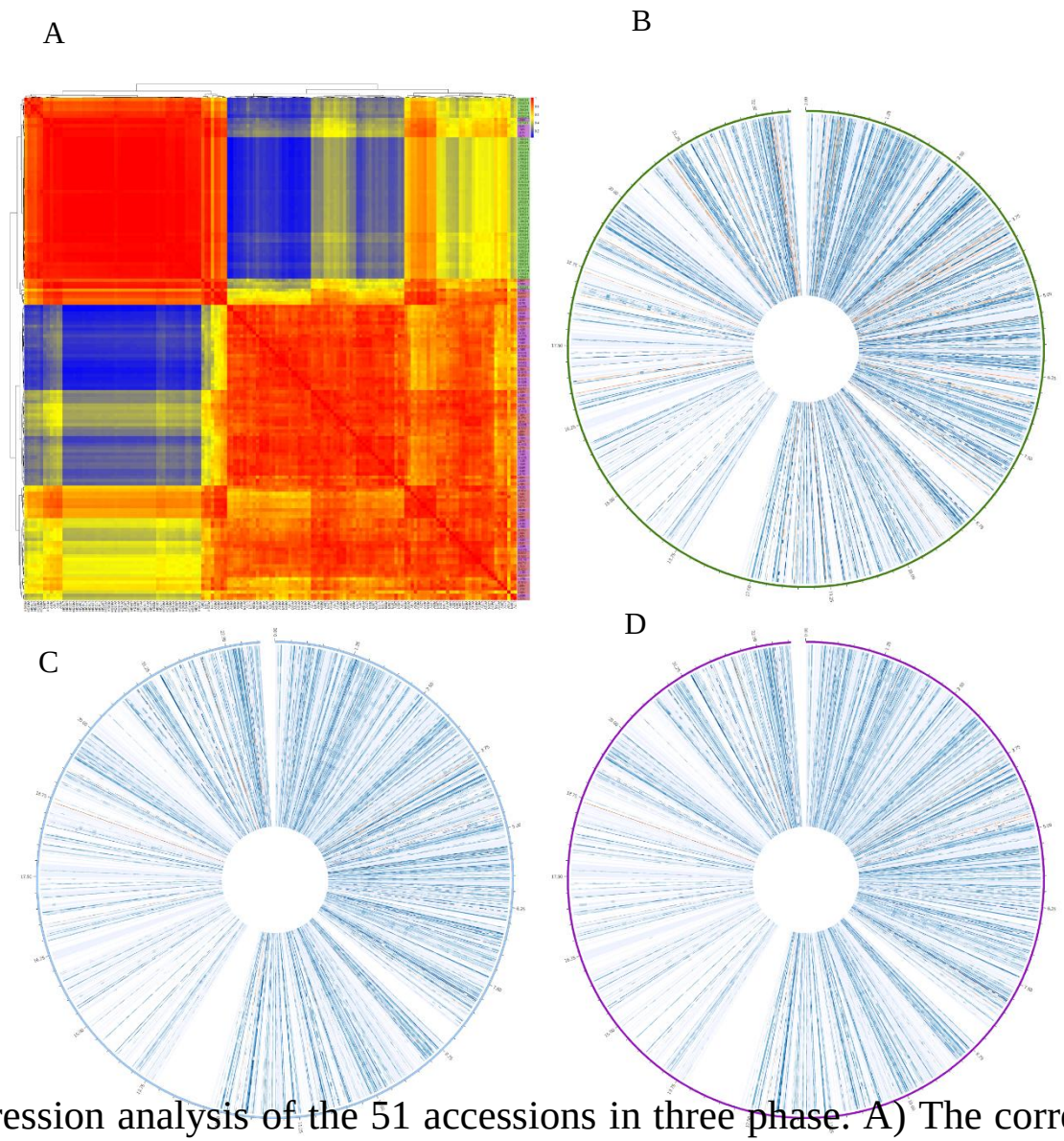
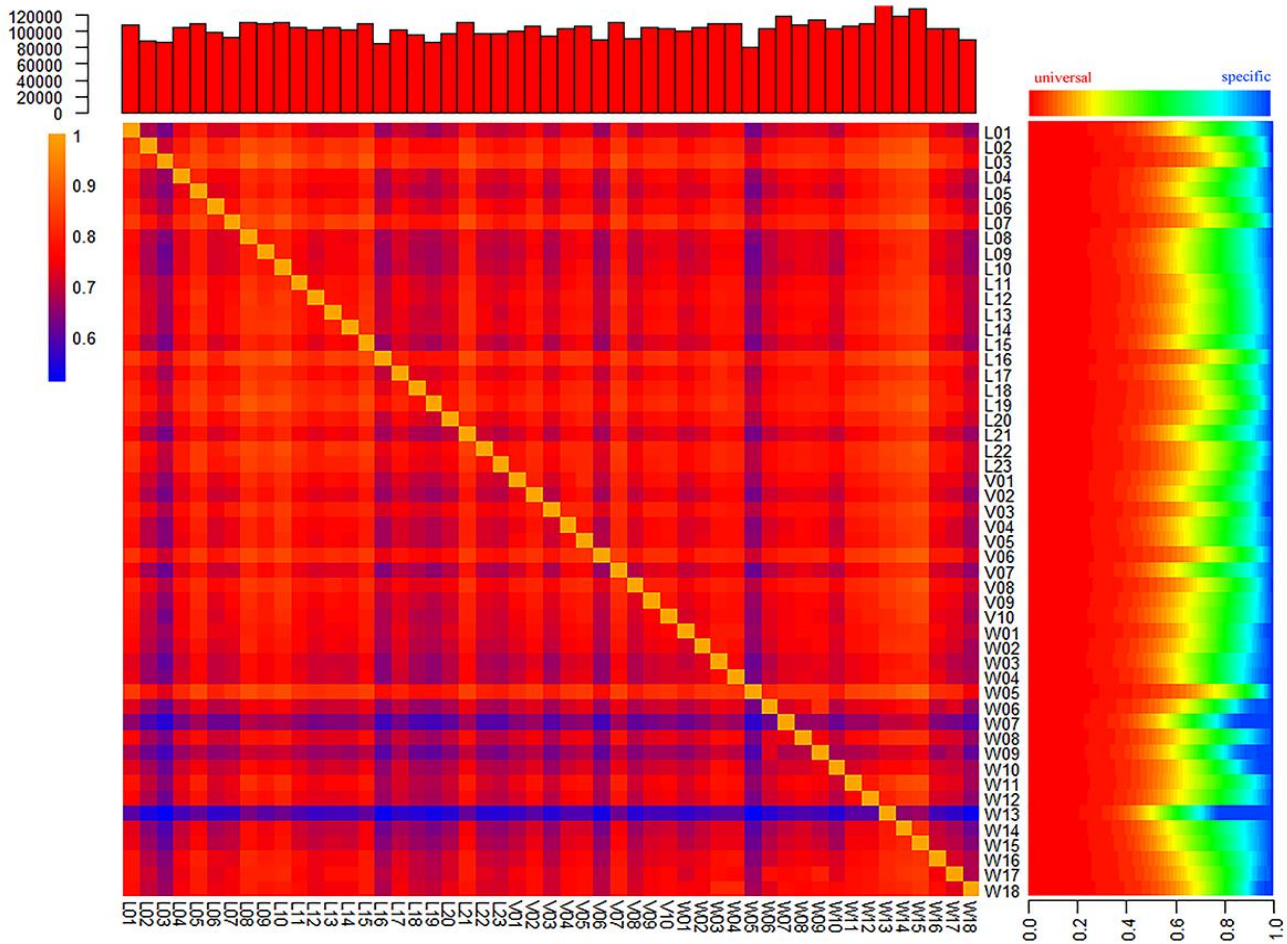
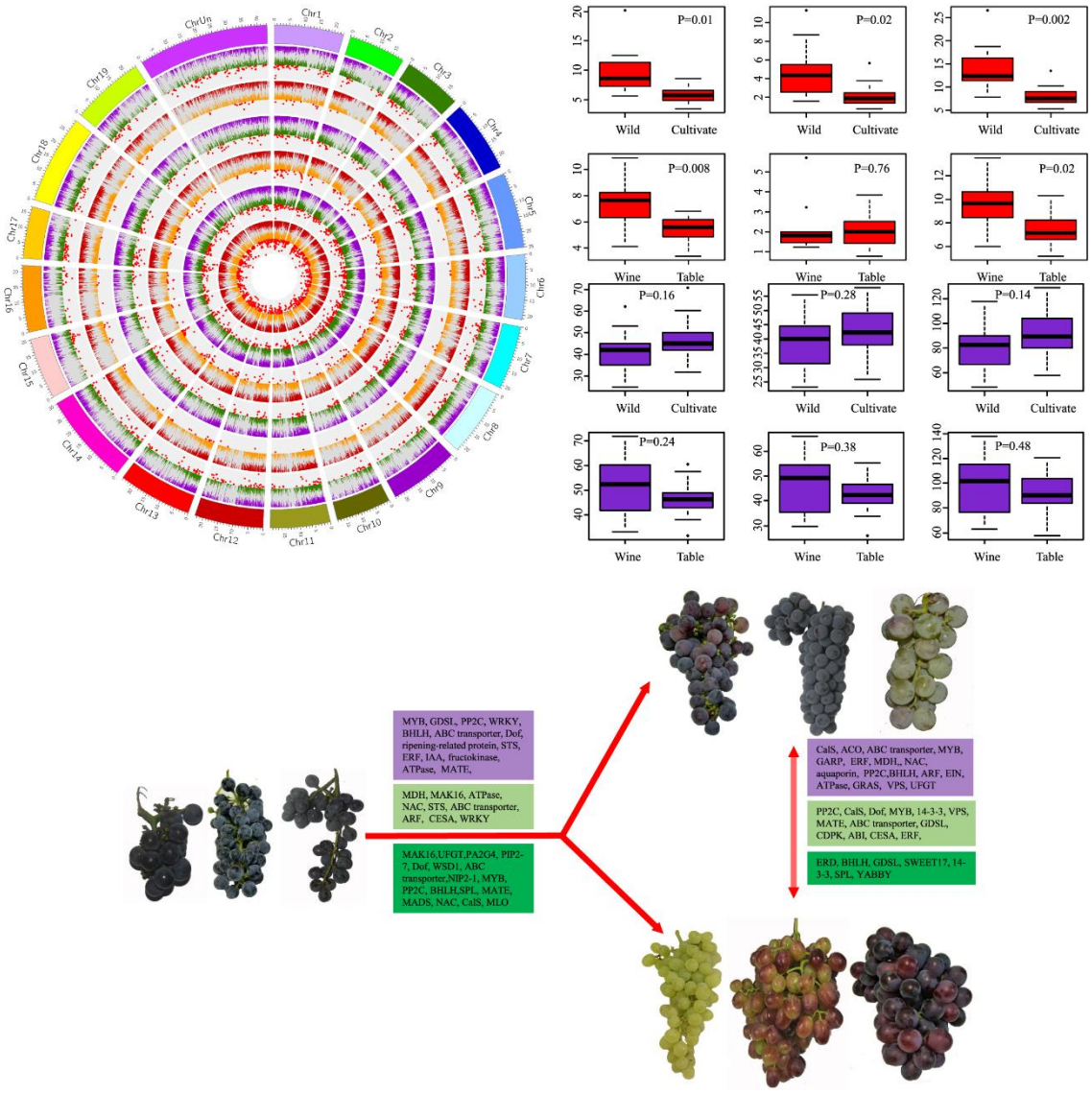


Figure 1 the expression analysis of the 51 accessions in three phase. A) The correlation analysis between the 153 samples. B) The expression level of the chromosome 1 in green hard. The sample position from the outside in is L01- L23, V01-V10, W01-W18, the red area means high expression level and the blue area means low expression level. The expression level were normalized by log2(FPKM/mean FPKM). C) The expression level of the chromosome 1 in veraison phase. D) The expression level of the chromosome 1 in ripening phase.

Transcripts assembling



The abundance, distribution, and similarity of uni-genes in 51 accessions. A) Number of uni-genes in all the 51 accessions. B) Similarity rate of the uni-genes among all the 51 accessions. Each block means the similarity uni-genes proportion of the whole uni-genens in this accession. C) distribution of universal and specific genes, the red color means they can find similarity uni-genes in most accessions, the blue one indicate only one or several accession contain these uni-genes.



The domestication related genes in the grape berry. A) Distribution and selection of the domestication genes on the whole genome. The purple bar is the expression level of cultivated grape, the green bar is the wild grape. The red is wine grape and the orange is the table grape. The red point is the genes which FDR < 0.05, and the blue ones is FDR >= 0.05. From outside in is the GH, V and R phase. B) Some important genes during the domestication. C) The sugar and acid content in each groups. The red box-plot is the acid content in the berry, the first one is tartaric acid, and the second is malic acid. The third box-plot is the total acid content. The purple box-plot is the sugar content in the ripening berry (glucose, fructose and total sugar).

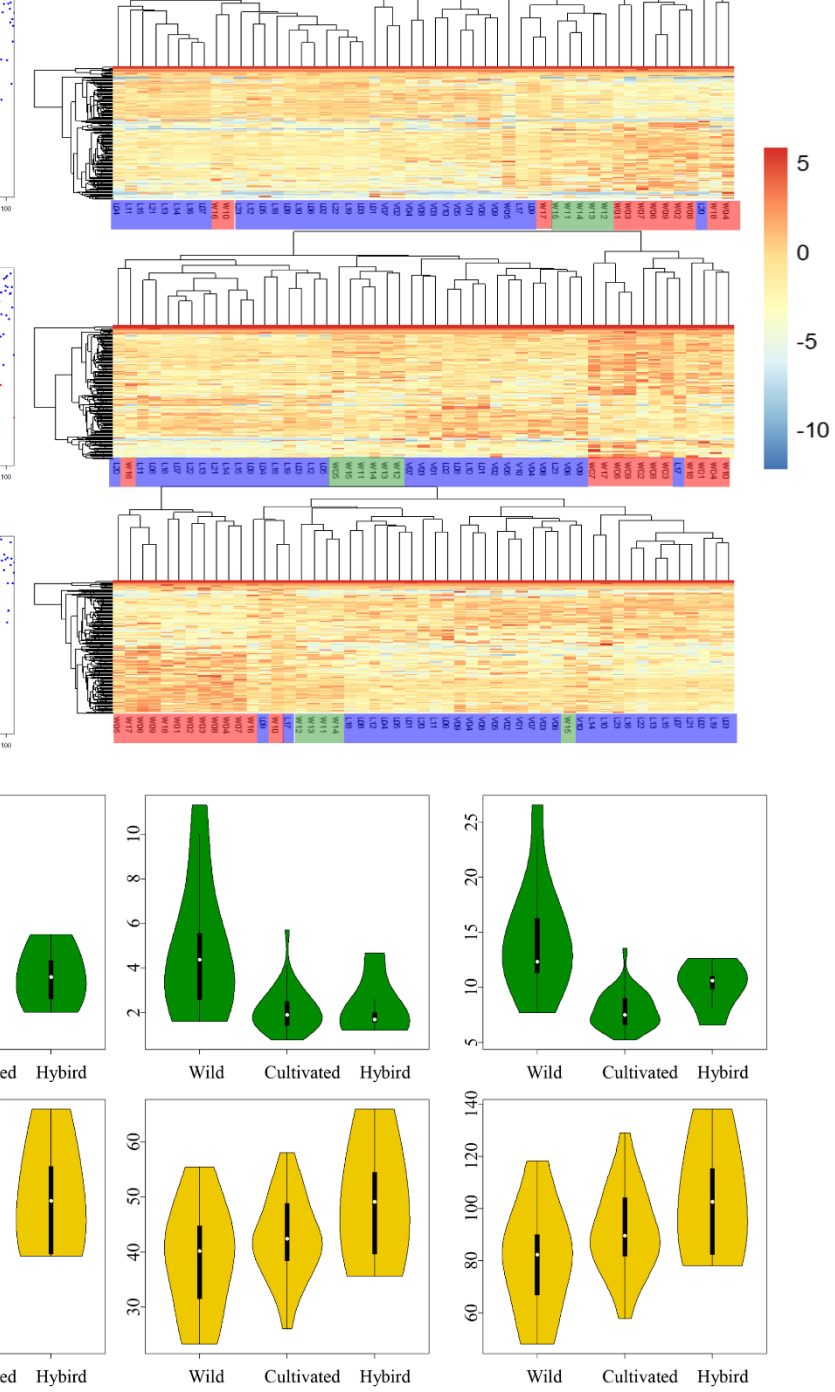
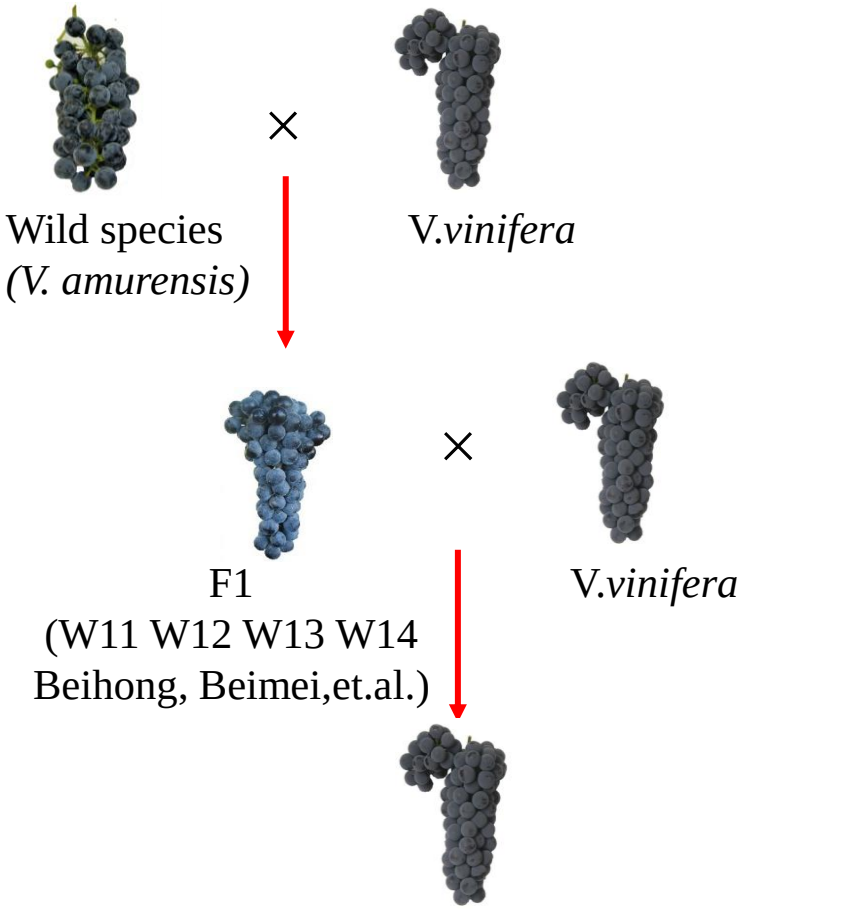
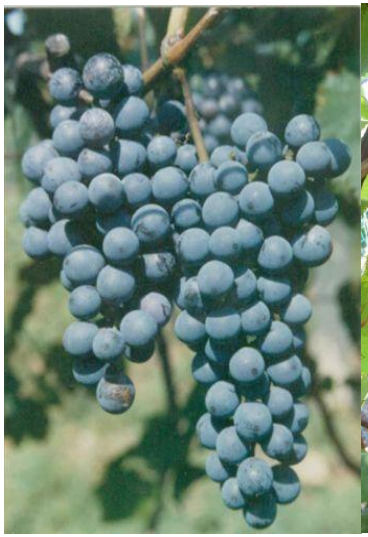


Figure 4 the behavior of the hybrid types. A-C) The selection of the different genes in GH (A) V (B) and R (C) and the cluster of these genes in each accession. The red points is the amid genes (|log2(change fold)| > 1). D) The acid content of each group. The three vioplot is tartaric acid, malic acid and total acid. E) The sugar content of each group. The three vioplot is glucose, fructose and total sugar.


×
Wild species
(V. amurensis)
V.vinifera
×
F1
(W11 W12 W13 W14
Beihong, Beimei,et.al.)
V.vinifera
BC1F1 （W15, Beiquan）



'Beicun'

'Beimei'

'Beihong'

'Beiquan'

'Beixi'

'Beixin'

		Brix	Berry weight（g））	Anti-cold
F1	Beihong	25.2	1.57	Strong
F1	Beimei	23.3	2.66	Strong
F1	Beichun	20.1	2.62	Strong
F1	Beixin	23.4	3.62	Strong
F1	Beixi	23.8	2.27	Strong
F2	Beiquan	18.0	4.52	Mid

Technique

Methods for grapevine breeding

- Cross breeding
- Bud mutation
- Radiation induced mutation breeding
- Seedling selection
- **Precision breeding**

Grape flowers (top) and emasculation of clusters (below)

Burger P. et al. 2009

Regeneration of grapevine

- Organogenesis pathway

Adventitious bud formation in leaf explants was first reported by Favre in 1976, and was further applied in several grapevine species, including wine grapes, table grapes and rootstocks.

Explants

- Leaf primordial fragments
- lamina or petioles
- hypocotyls of somatic embryos
- callus

Regeneration of grapevine

- Embryogenesis pathway

Grapevine regeneration via embryogenesis pathway *in vitro* is a well established procedure as early as in 1976 (Mullins and Srinivasan).

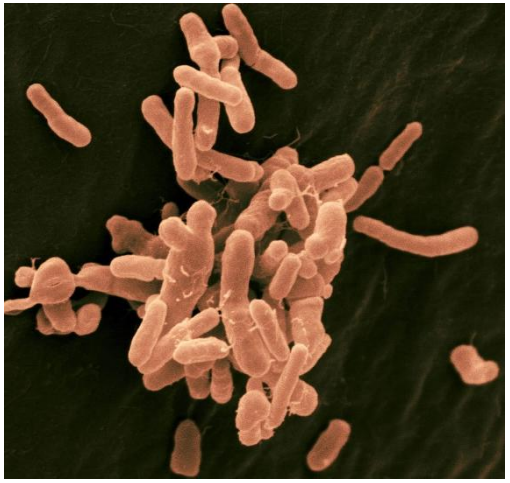
Explants

- Somatic callus
- Somatic suspension cells
- Somatic embryos

Genetic transformation

- *Agrobacterium tumefaciens*-mediated method
  - Embryogenic callus
  - Embryogenic suspension cells
  - Leaf discs

The *Agrobacterium tumefaciens*-mediated transformation system is the predominant technology based on the ability of *Agrobacterium* on insert genes into plant cells.



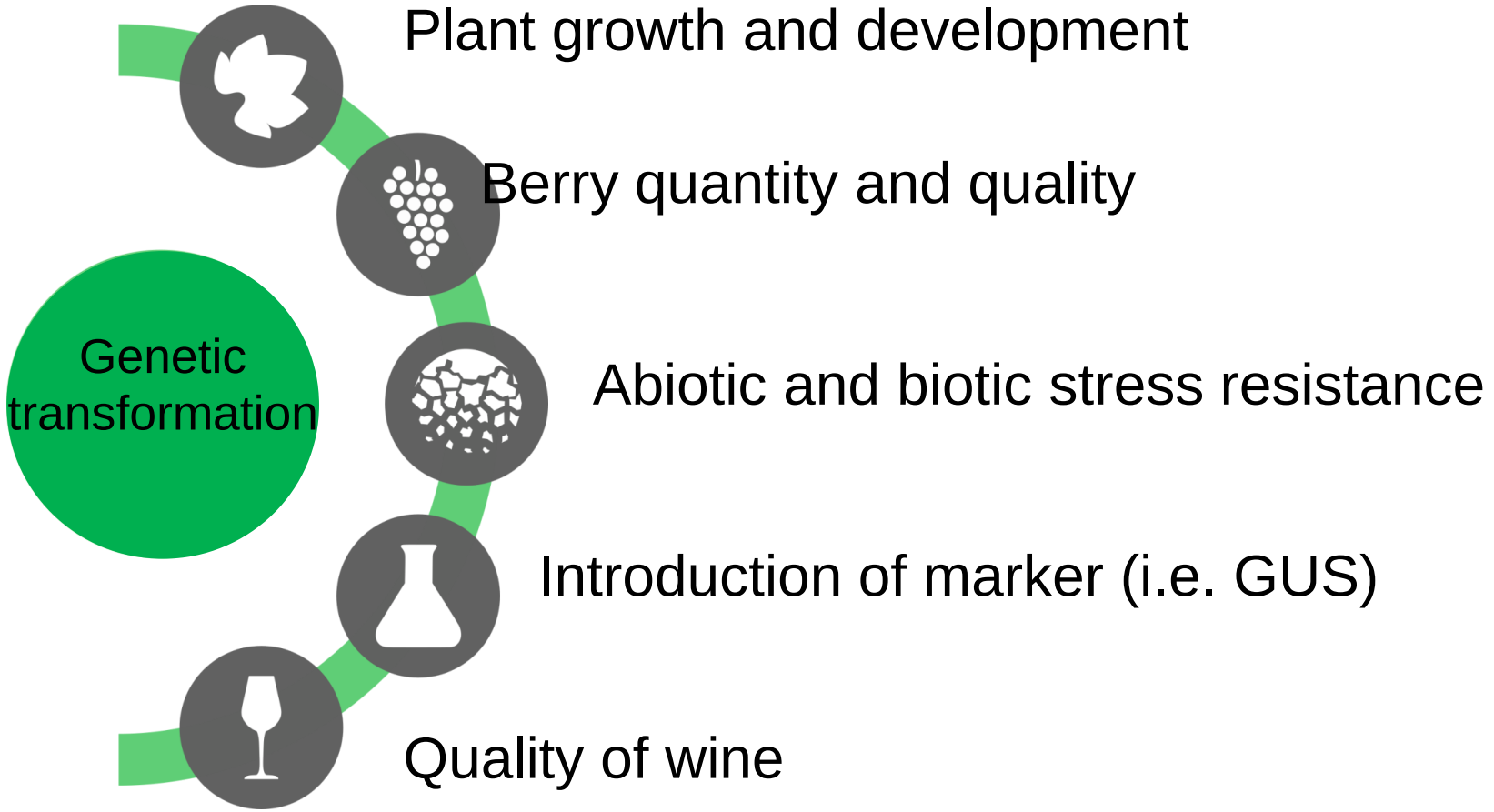
Genetic transformation with new techniques

In 1990, Mullins et al. obtained transformed plants overexpressing *GUS* gene

In 2009, Hanania et al. silenced the ubiquitin extension *S19a* gene by using RNAi

In 2016, Ren et al. knocked out *IdnDH* gene in Chardonnay via CRISPR/Cas9

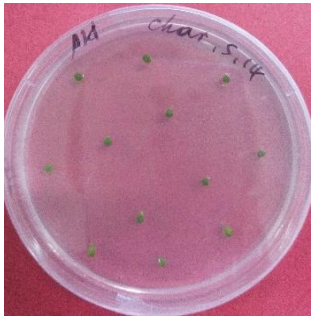
Applications of genetic transformation in grapevine



The procedure of genetic transformation via embryogenesis

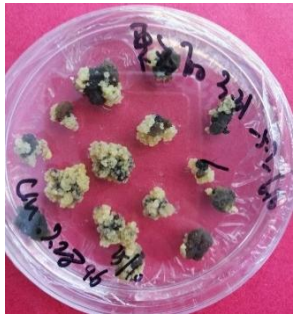
Whole flower

1-2 days



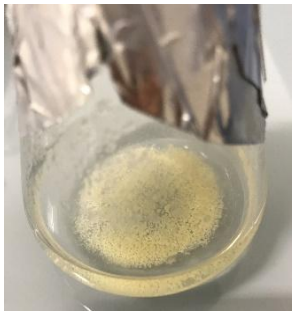
Calli induction

3-6 months



Embryonic calli

2-3 days



Transformation

2-3 months

Selection

15-30 days

Embryo induction

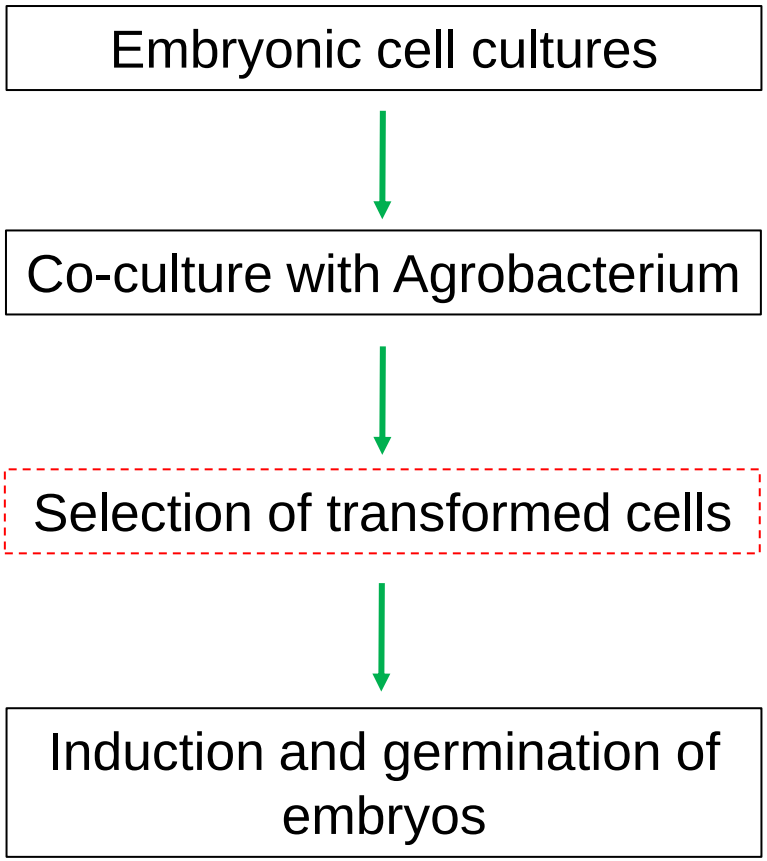
15-30 days

Embryo germination

2-3 months

Generation of shoots

The timing of genetic transformation via embryogenesis



Sub-culture weekly

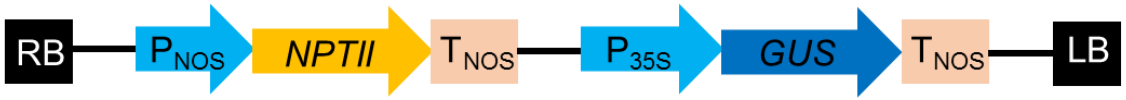
Usually 2 days

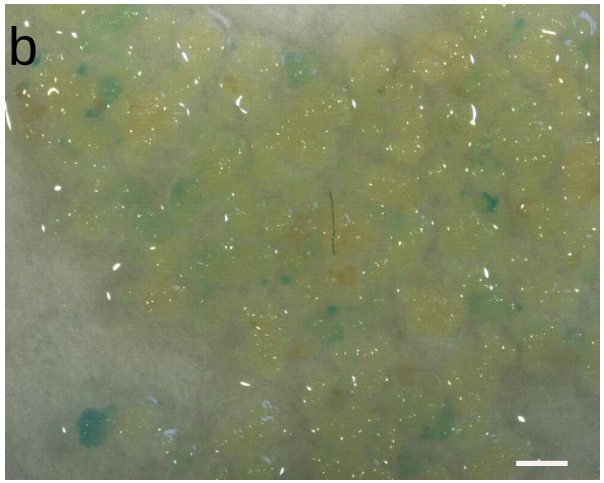
At least 2-3 months. Time-consuming step. Skipped?

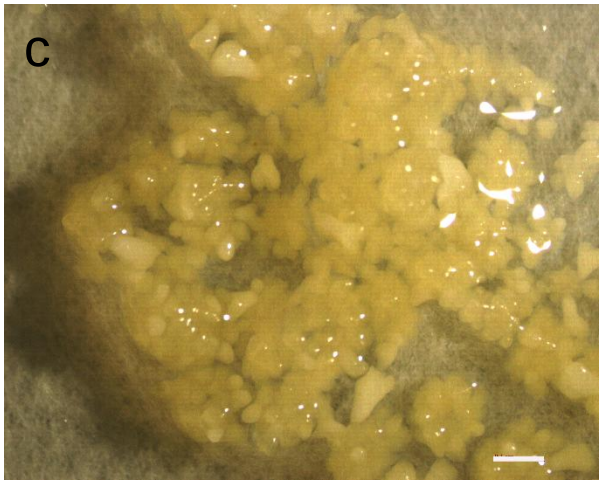
Usually 1-2 months

Directly induction of embryos without selection

a







a, vector construct with *GUS* marker gene; b, GUS staining of cells after Agrobacterium co-culture; c, Embryos induction from grape cells

CRISPR/Cas9 has been the predominant method for genome editing.

Accomplishments of genome editing in grape with CRISPR/Cas9

Plant material	Delivery method	Target gene	Modification type	Effect	Off-target	Reference
Protoplasts	PEG; transient expression	*MLO-7*	Gene knockout		ND	Malnoy et al. 2016
Embryogenic cells	Agrobacterium infection; stable integration	*IdnDH*	Gene knockout	Reduction in tartaric acid content	No	Ren et al. 2016
Proembryonal masses (PEM)	Agrobacterium infection; stable integration	*WRKY52*	Gene knockout	Increased resistance to Botrytis cinerea	No	Wang et al. 2017
Embryonic calli	Agrobacterium infection; stable integration	*PDS*	Gene knockout	Albino phenotype	No	Nakajima et al. 2017

Many factors have been found to have an impact on CRISPR/Cas9 system.

sgRNA/Cas9 expression level;
PAM sequence;
Genetic background
…

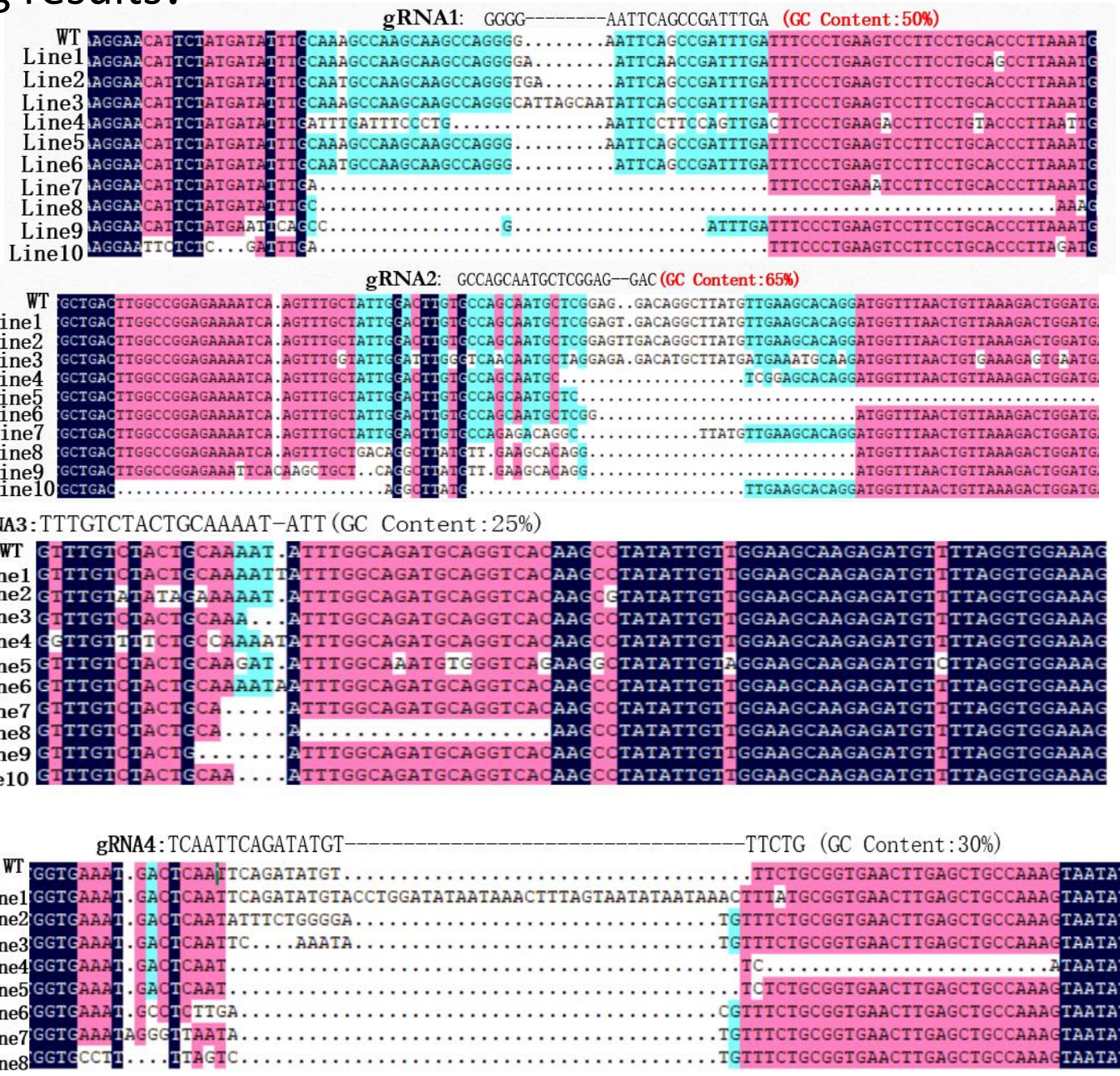
The efficiency of CRCRISPR/Cas9-mediated targeted mutagenesis in grape

Four GC content of sgRNAs were designed to target exon sites of the Vitis vinifera phytoene desaturase (*VvPDS*) gene two varieties ‘Chardonnay’ and ‘41B’ suspension cells were used as the transgenic cell mass.

	SEQUENCE	GC Content
gRNA1	GGGGAATTCAGCCGATTTGA	50%
gRNA2	GCCAGCAATGCTCGGAGGAC	65%
gRNA3	TTTGTCTACTGCAAAATATT	25%
gRNA4	TCAATTCAGATATGTTTCTG	30%

The sequencing results：

The sequencing results shows that four GC content of sgRNAs all worked in the transgenic cell mass.



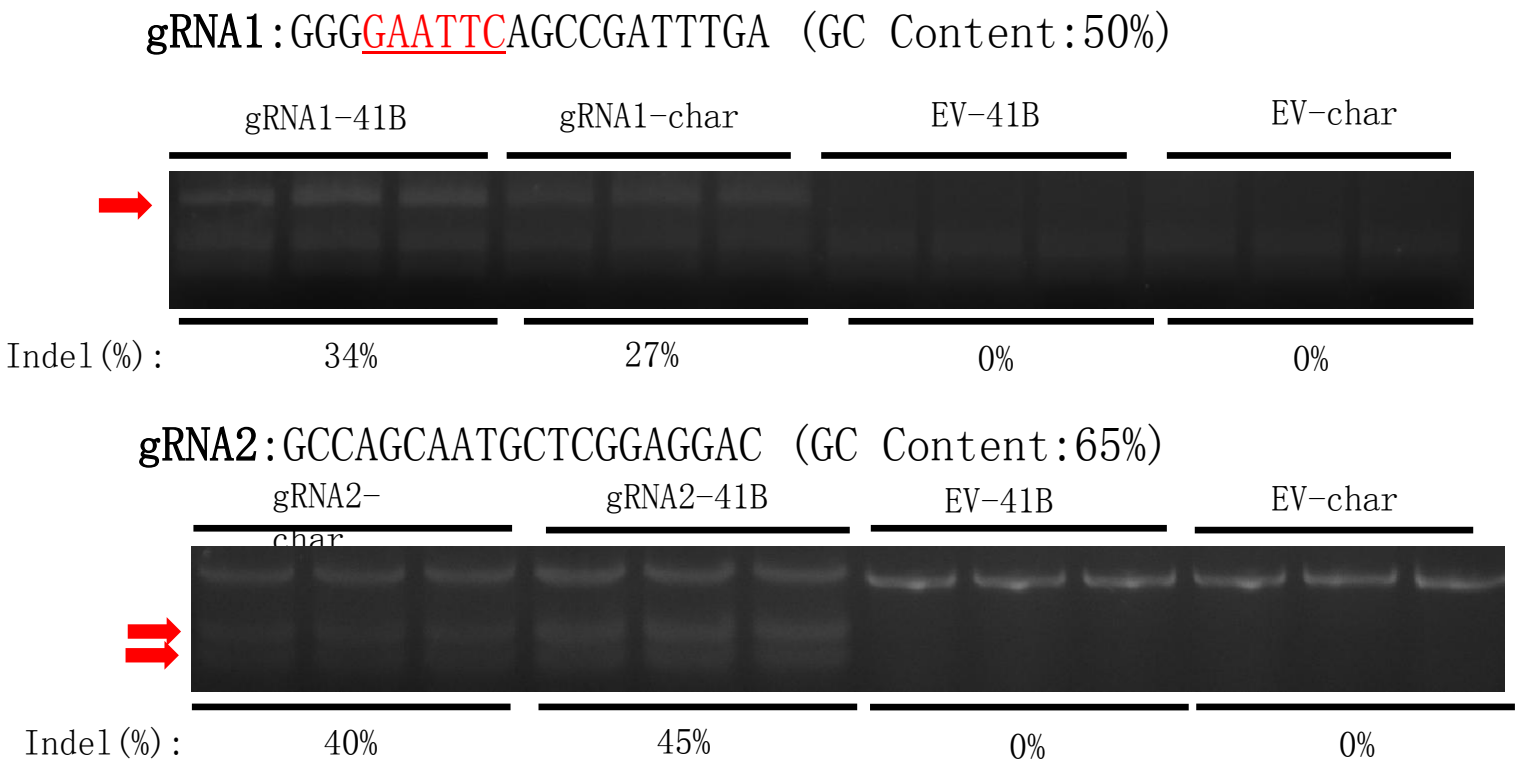
The GC content of guideRNA

	GC content(%)	Indel(%)
gRNA1	50%	34%
gRNA2	65%	45%
gRNA3	25%	10%
gRNA4	30%	15%

- T7EI assay and PCR/RE assay showed that the most efficient one is the 65% GC content sgRNA, followed by 50%. The indel mutations were detected in transgenic CM with the 25% and 30% GC content sgRNA respectively but the efficiency of them is much lower than others.

The variety of the suspension cells for transformation

Two high efficient gRNA were used to detect the efficiency of two varieties (‘Chardonnay’ and ‘41B’) suspension cells (red arrowheads indicate cleaved mutated bands).

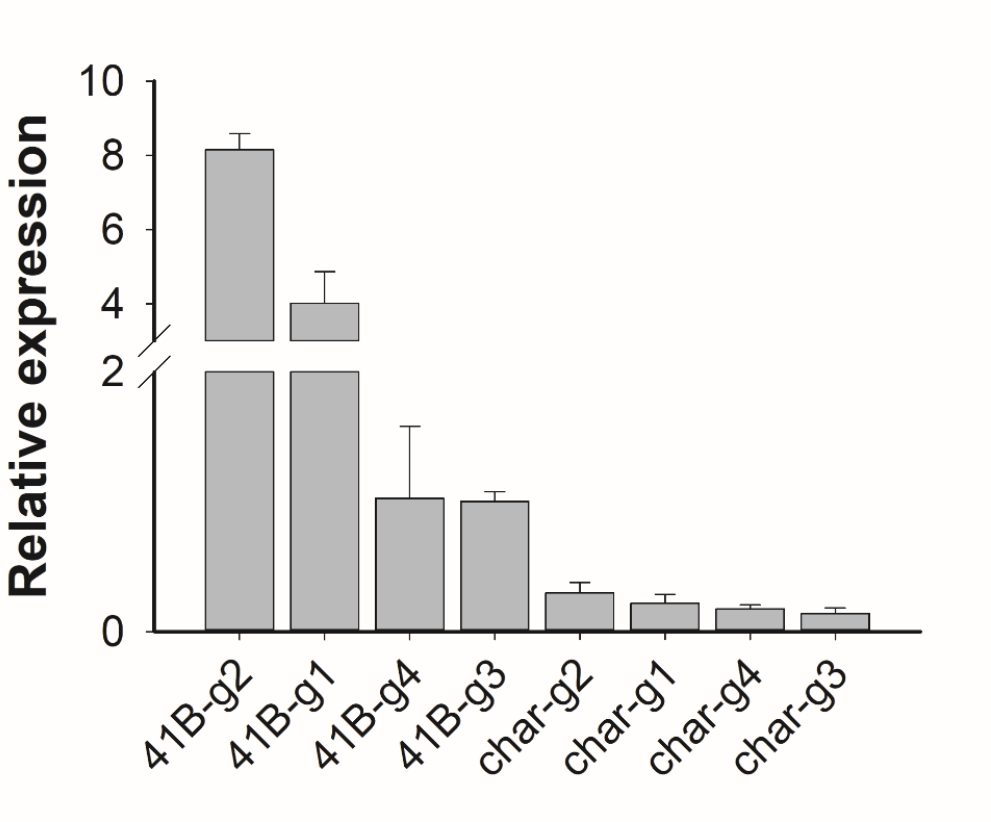


The variety of the suspension cells for transformation

	Char-Indel(%)	41B-Indel(%)
gRNA1	27%	34%
gRNA2	40%	45%

- By comparing the efficiency of CRCRISPR/Cas9 system in ‘Chardonnay’ and ‘41B’ transgenic CM and we found that CRCRISPR/Cas9 system worked more efficiently in ‘41B’ suspension cells.

The expression level of *SpCas9*

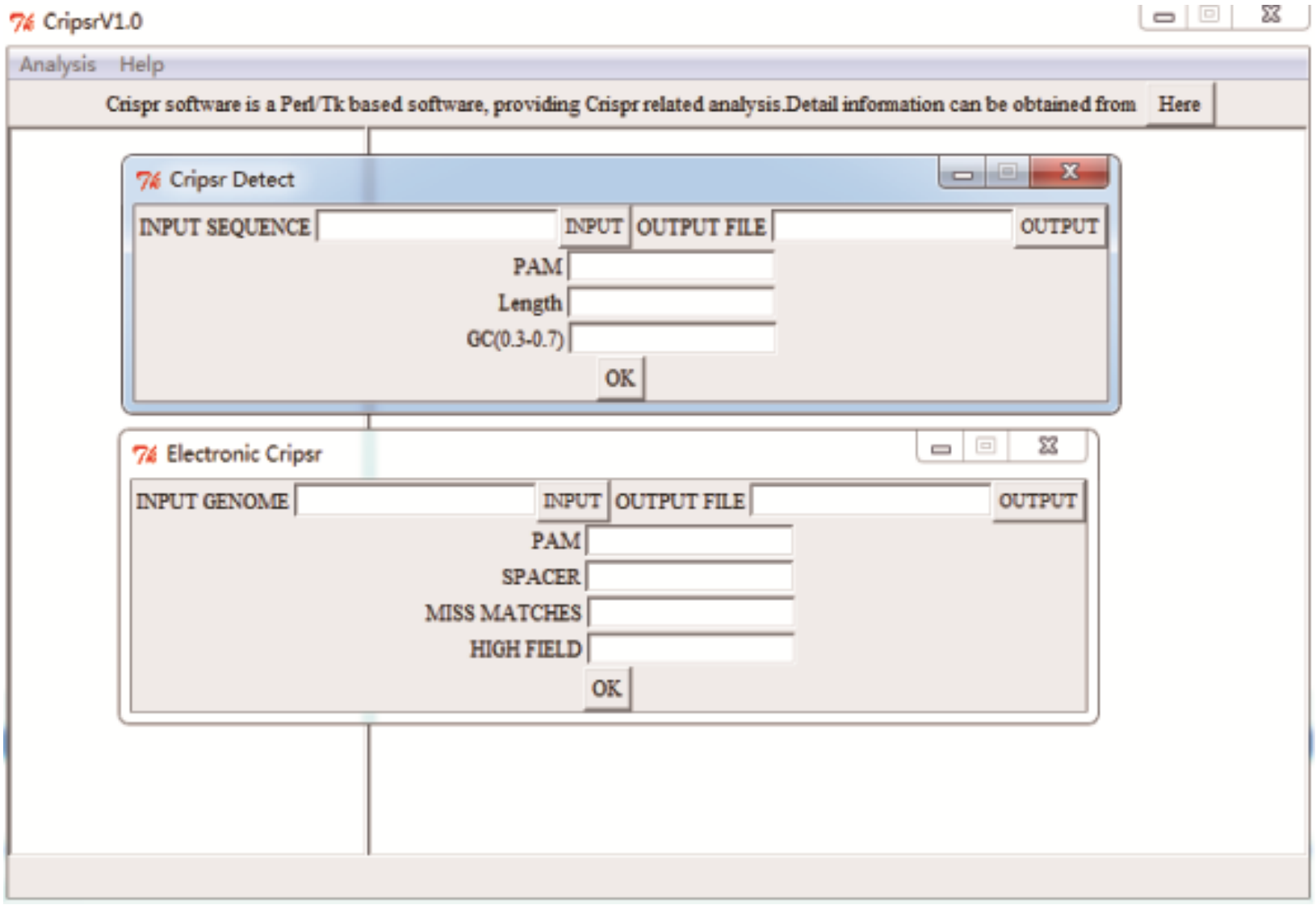


The results of qPCR showed that the CRISPR-Cas9 system which had the higher editing efficient expressed higher level of *SpCas9.*

Data from grape genome browser （GRAPE-CRISPR）

- http://biodb.sdau.edu.cn/gc/index.html

PLANT-CRISPR (Desktop software)



Current work

CRISPR/Cas9-mediated genome editing

- Grape berry development and trait
  Sugar content;
  tartaric acid;
  aroma

- Plants with increased biotic resistance
  cold and freeze tolerance;
  drought tolerance

CRISPR/Cpf1-mediated genome editing

Acknowledgement

Shaohua Li, Institute of Botany, Chinese Academy of Sciences
Peige Fan, Institute of Botany, Chinese Academy of Sciences
Yi Wang, Institute of Botany, Chinese Academy of Sciences
Cuixia Liu, Institute of Botany, Chinese Academy of Sciences
Chong Ren, Institute of Botany, Chinese Academy of Sciences
Fengrui Ren, Institute of Botany, Chinese Academy of Sciences
Meilong Xu, Institute of Botany, Chinese Academy of Sciences
Zhanwu Dai, Institute of Botany, Chinese Academy of Sciences
Serge Delrot, University of Bordeaux, ISVV, INRA
David Lecourieux, University of Bordeaux, ISVV, INRA
Fatma Lecourieux, University of Bordeaux, ISVV, INRA
Jianfu Jiang, Zhengzhou Fruits Research Institute, Chinese Agriculture Academy of Sciences
Perte Nick, Karlsruhe Institute of Technology

Thanks for your attention!