MAP-BASED IDENTIFICATION AND POSITIONAL CLONING OF *XYLELLA FASTIDIOSA* RESISTANCE GENES FROM KNOWN SOURCES OF PIERCE'S DISEASE RESISTANCE IN GRAPES.

Principal Investigator:
Andrew Walker
Dept. of Viticulture and Enology
University of California
Davis, CA 95616
awalker@ucdavis.edu

Cooperating Staff:
Summaira Riaz
Dept. of Viticulture and Enology
University of California
Davis, CA 95616

Reporting Period: The results reported here are from work conducted September 2009 to October 2010.

ABSTRACT
This report presents updated results on the refined genetic and physical mapping of the Pierce's disease (PD) resistance locus, *PdR1*, which originates from *Vitis arizonica/candicans* b43-17 and is flanked by the SSR markers, VVCh14-78 and VVCh14-81, within a 1cM distance. We have two BAC libraries for b43-17, each with one restriction enzyme (*Hind* III and *Mbo* I), and the screening of the *Hind* III BAC library with flanking markers was completed. The Pinot noir genome sequence was used to develop SSR markers to screen the BAC library, and these markers were used to reduce *PdR1*'s physical distance. Two screenings of the libraries identified 24 (with markers VVCh14-56 and VVCh14-10) and 17 positive BAC clones (with marker VVCh14-58). Five clones were positive with VVCh14-56 and VVCh14-58. Clone 'H69J14' (which is bigger than 200Kb) was selected for sequencing and clone spans scaffold 68 and 171 of the Pinot noir genome sequence. A total of 42,000 sequences were generated, however assembly was complicated by a large number of transposable elements in the resistance region. A Fosmid library of H69J14 is now being generated to obtain larger (35kb inserts) to help resolve assembly problems based on the repetitive regions. Clones in the region from the Pinot noir genome possess four tandem repeats of serine threonine protein kinase with leonine rich repeat domains, genes that are involved in microbial recognition and plant defense reactions. Genetic mapping is also underway in three other populations to enable a better understanding of PD resistance and of *PdR1*: the 07744 (resistance from *V. arizonica* b40-14); the 04191 (resistance from F8909-17 *PdR1a*); and the 05347 (resistance from *V. arizonica/ girdiana* b42-26). A total of 152 markers were completed for 07744 to develop the framework map. Greenhouse screening of the 07744 population is complete. Preliminary results with the 07744 indicated that PD resistance (*PdR1c*) resides on chromosome 14, in the same region where *PdR1a* (resistance from F8909-17) and *PdR1b* (resistance from F8909-08) mapped from the b43-17 background. However, b40-14's SSR alleles for resistance are very different from those of b43-17. Genetic mapping of the quantitative resistance from *V. arizonica/girdiana* b42-26 continues in the 05347 (*V. vinifera* F2-35 x b42-26) population. About 70% of the population has been greenhouse screened, the remaining 75 seedlings will be tested early next year. b42-26 is surprisingly homozygous with the SSR markers we have available, which prompted the development of 71 new markers and the acquisition of about 200 others so that genetic mapping can be improved.

LAYPERSON SUMMARY
Genetic mapping from two different forms of *Vitis arizonica* have identified a region on chromosome 14 that is responsible for Pierce's disease (PD) resistance, which we termed *PdR1*. We have mapped two forms of *PdR1* from *V. arizonica / candicans* b43-17, and have mapped a third form, *PdR1c*, that originated from *V. arizonica* b40-14. These forms are both single dominant genes for PD resistance. We are also examining another source of resistance that is controlled by multiple genes that originated from *V. arizonica / girdiana* b42-26 and have begun the fine-scale mapping necessary to determine if markers are tightly enough linked to these multiple resistance genes to be used for marker-assisted selection. We plan to combine these multiple resistance sources in our breeding program to ensure broad and durable resistance to PD. Genetic markers to these forms of resistance will make this possible and allow the confirmation and tracking of interbred progeny. These mapping efforts are also essential to physically locating and characterizing PD resistance genes. At present, the chromosome region where *PdR1* exists has been sequenced and the pieces of sequence were arranged and compared to the Pinot noir genome sequence. This comparative analysis indicates that the susceptible Pinot region carries four genes. The Pinot noir region was compared to the sequences we have from the resistant b43-17 and we identified four candidate resistant genes. We are in the process of characterizing their function and determining which are likely to be involved in PD resistance.

INTRODUCTION
Genetic mapping to identify genomic regions that carry disease resistance genes can greatly facilitate breeding and lead to the map-based positional cloning of the resistant genes. In this project, we initiated mapping of Pierce's disease (PD) resistance in different forms of *V. arizonica* (Riaz et al. 2007). These efforts are closely coupled to a breeding program focused on developing PD resistant winegrapes (see companion report). The breeding program produces and greenhouse screens the seedling populations upon which the genetic mapping program depends. While the tightly linked genetic markers generated in these mapping efforts are used to optimize and greatly accelerate the PD breeding program (Riaz et al. 2009). These markers are essential to the successful introgression of resistance from multiple sources, and thus for the production of broader and more durably resistant grapevines (Riaz et al. 2008a). Genetic maps associate DNA markers with phenotypic

traits, and allow the linking of these traits with markers positioned relative to each other on chromosomes. Fine scale mapping of given regions and careful screening of recombinant progeny (those with a given genetic marker but without resistance, or vice versa, because of a recombination event) is critical to the identification of relatively short genetic regions that can then be sequenced so the genes responsible for PD resistance can be characterized and their function studied (Riaz et al. 2008b).

OBJECTIVES

1. Completely characterize and refine the PD resistance locus on chromosome 14 by genetically mapping in four populations that derive resistance from *V. arizonica/candicans* b43-17 and its *V. rupestris* x b43-17 progeny F89090-08 (*PdR1b*) and F8909-17 (*PdR1a*): 04190 (*V. vinifera* F2-7 x F8909-08), 9621 (D8909-15 x F8909-17), 04191 (F2-7 x F8909-17), and 04373 (*V. vinifera* F2-35 x *V. arizonica/candicans* b43-17).
2. Genetically map PD resistance from other forms of *V. arizonica*: b42-26 (*V. arizonica/girdiana*) and b40-14 (*V. arizonica*).
3. Develop a BAC (bacterial artificial chromosome) library for the homozygous resistant genotype b43-17 (parent of F8909-08, and F8909-17) and screen the library with closely linked markers.
4. Complete the physical mapping of *PdR1a* and *PdR1b* and initiate the sequencing of BAC clones that carry *PdR1a* gene candidates.

RESULTS AND DISCUSSION

Objective 1. We have refined the genetic position of the *PdR1a* and *PdR1b* resistance loci between marker VVCh14-56 and VVCh14-77 in the maps of two populations 9621 and 04190. These SSR markers have unique allele sizes for resistant parents that are not found in susceptible *vinifera* parents. This characteristic makes these markers very valuable for marker-assisted selection (MAS) in our ongoing wine grape-breeding program (see companion project report). The physical distance between these markers is 200-230 Kb, and they have been used to test additional plants and identify key recombinants critical for fine-scale mapping. Currently there are three key recombinants in the 9621 population from a tested set of more than 900 plants. F8909-08 possesses the *PdR1b* resistance locus, which is being mapped in the 04190 population. The greenhouse screen was repeated for key recombinants, which also helped to refine the data. In addition, marker analysis identified 14 more recombinants from 15 different crosses (totally about 1,000 plants) based on resistance from F8909-08. An additional 11 recombinants were found from the *PdR1a* source, and greenhouse screening was completed on the 35 recombinants. This screen identified four key recombinants. In two plants the recombination event occurred between *PdR1b* and VVCh14-02, and in one plant the recombination event occurred between *PdR1b* and VVCh14-70. The greenhouse screen is being repeated for four other recombinants that had inconclusive first test results. In the most recent map, we have placed the *PdR1b* locus between markers VVCh14-81 and VVCh14-77. Both of these markers are roughly 200Kb apart based on the Pinot noir genome sequence.

The 04191 population (*V. vinifera* F2-35 x F8909-17) has 153 progeny plants and has *PdR1a*-based resistance. This population allows *PdR1* to be examined without possible confounding effects from D8909-15, the other parent of the 9621 population, since D8909-15 has a multigenic resistance from b42-26. The resistance locus *PdR1a* is mapped in the 9621 (D8909-15 x F8909-17) population, and the 04190 population mentioned above, and refined mapping focused only on chromosome 14. The 04191 population is critical for the identification of any minor genes that might contribute to PD resistance. Therefore, we expanded the framework genetic map to all 19 chromosomes. A total of 143 SSR markers representing all 19 chromosomes were added to the set of 153 seedlings, of which 141 were greenhouse screened for resistance. The greenhouse screen results of seven of the plants did not match with the marker data. These plants are being rescreened with results expected later in Fall 2010. ELISA results from remaining the 134 plants matched with the marker results. A total of 75 seedlings have low bacterial titer values and carried resistant alleles with the tightly linked flanking markers and fifty-nine seedlings had high bacterial numbers and they inherited alleles linked to the susceptibility to PD, which confirmed a single dominant gene 1:1 segregation. We are in the process of evaluating the greenhouse results, genetic mapping with 138 markers and QTL(Qualitative Trait Loci) analysis.

Objective 2. In response to recommendations from the CDFA-PD board and reviewer recommendations to broaden resistance, we are characterizing resistance from two additional sources and making good progress. The main purpose is to identify additional resistance sources, genetically map them and use tightly linked molecular markers to pyramid resistance from different backgrounds into single line. We are pursuing two other resistant *V. arizonica* forms: b42-26 *V. arizonica/girdiana* from Loreto, Baja California; and b40-14 *V. arizonica* from Chihuahua, Mexico. Although they are morphologically different than b43-17, they both posses strong resistance to PD and greatly suppress *Xylella. fastidiosa* levels in stem tissue after greenhouse screening.

Greenhouse screening data indicate that resistance from b42-26 is quantitative. A small breeding F1 population 05347 (*V. vinifera* F2-35 x b42-26) was produced in 2005, and a subset of 48 genotypes was greenhouse screened and found 35 were resistant and 13 were susceptible. A total of 337 markers were tested on small a parental data set. Results found a high level of homozygosity for b42-26 (only 113 markers were polymorphic); 184 markers were homozygous for the male parent b42-26, and 40 markers did not amplify. In the Spring of 2008 and 2009, crosses were made to increase the population size.

Currently, 239 seedlings exist in the field and greenhouse screening is complete on 63 seedlings, underway on 111, and the final set of 75 will be screened in 2011. In addition to increasing the population size, we also need more markers that are polymorphic for b42-26, a very homozygous selection. We developed 71one new SSR markers from clone sequences generated from the Vitis Microsatellite Consortium. These clones had microsatellite repeats in the beginning or end of the sequence, which left no room for primer design. A total of 238 of these clone sequences were compared to the nearly homozygous 12X Pinot noir genome in order to obtain additional flanking sequences. There were good matches for 71 of the clones and primers were designed for them; 69 of the newly designed primers amplified and 67 of them generated a clean banding pattern with *V. vinifera* DNA samples (results form this study were submitted as a research note in AJEV). We also acquired primer sequences of an additional 200 markers that have not been tested with b42-26. Marker testing on small set of parents and progeny is underway. We are now adding markers to develop a framework map of the entire population set; 50 markers have been completed and more are being added.

Resistance in *V. arizonica* b40-14 seems to be homozygous and is controlled by a single dominant gene. We mentioned in previous reports that all F1 progeny from a cross of *V. rupestris* x b40-14 (the R series) were resistant to PD except three genotypes with intermediate results. Two resistant siblings of this population were used to develop the 07388 (R8918-02 x *V. vinifera*) and 07744 (R8918-05 x *V. vinifera*) populations. A summary of the genetic mapping and QTL analysis is presented below: 227 markers were polymorphic for one of the parents; 152 were analyzed on the entire set of 122 plants; a framework map of R8918-05 was produced with MAP QTL (4.0) and the Kruskal-Wallis approach was used to complete the preliminary analysis. PD resistance mapped only on chromosome 14 – the same chromosome where *PdR1a* and *PdR1b* mapped. PD resistance from b40-14 (which we have named *PdR1c*) also maps in the same general region between flanking markers VVCh14-77 and VVIN64 and within 1.5 cM. The LOD threshold for the presence of this QTL was 33 and 82% of the phenotypic variation was explained (**Figure 1**). In 2009, crosses were made with F1 resistant selections from 07744 population.

Objective 3 and 4. Two BAC libraries were created from the homozygous resistant b43-17. Screening of the library with markers VVCh14-10, VVCh14-56 and VVCh-58 identified 41 positive clones– four of the clones were positive with VVCh14-10, VVCh14-56 markers (H23-P13, H34-B5 and H64-M16 and H45-J22) and five of them were positive with the VVCh14-56 and VVCh14-58 marker (**Figure 1a**). Two clones were selected for sequencing with an overlap of 60Kb and a spanned region of 340Kb. A shotgun library of BAC clone H64M16 was Sanger sequenced. Clone H69J14 was selected for 454 sequencing. A total of 42,000 sequences were generated and two different programs were used to assemble the sequence. However, the sequenced region was highly enriched with repetitive elements, which complicated the assembly. Newbler software as well as Lasergene program SEQMAN do not work well with sequences containing many repeated regions. In order to generate longer sequence fragments, a shot gun library was constructed for clone H69J14; 384 sequences were generated in both directions to develop paired ends in order to fill the gaps between the contigs from the 454 sequence data. We then masked the repetitive region from all the sequences (both H69J14 and H64M16 clones) to carry on the assembly with MIRA assembler program. This improved the assembly, but the contig number was still very high and not suitable for primer walking. Moreover, all the major contigs had masked repetitive regions on both ends indicating that the primer design effort would not generate sequence specific results capable of bridging the gaps. We are now in the process of developing a Fosmid library with an insert size of 35-40Kb, and the resulting 384 sequences in both directions will allow us to tag smaller contigs from the 454 and shotgun reads data. Because the fosmid clones are 35Kb inserts, it will help resolve assembly problems based on the repetitive regions.

Recently, the 12X assembly of Pinot noir (PN400204) sequence became available. It is an improved version of the 8X assembly we used previously. Detailed analyses identified four tandem repeats of serine threonine protein kinase with a leucine-rich repeat domain gene family in the resistance region (**Table 1**). All four genes have large introns indicating that they may carry transposable element like sequences. We utilized CENSOR software to screen query sequences against a reference collection of repeats to generate a report classifying all detected repeats (**Table 1**). All four genes carry DNA transposons as well as LTR(Long Terminal Repeat) retrotransposons documenting the complexity of the region. In this situation, a direct comparison of the H69J14 clone sequence to the PN40024 sequence is not advisable because it is possible that the arrangement of repetitive elements is significantly different between the resistance region of the two genomes. A comparison of a larger sequence of the region, without the resistance genes, shows up to 98% homology with the susceptible PN40024. However, the b43-17 resistance region sequences that overlap with scaffold 68 of Pinot noir matched to multiple sites and the level of similarity was reduced. These results suggest that the b43-17 genomic region with the PD resistance gene(s) is divergent from PN40024 and that transposable elements may play a major role in these sequence differences. Our main emphasis is assembling the complete region in order to make comparisons to the susceptible sequence. This will help us understand the causes of sequence divergence and the evolution of the PD resistant gene family. Previous studies have indicated that the sequence of chromosome 14 is known to carry members of 13 different families of DNA transposons and retrotransposon (Moisy et al. 2008). We identified four potential candidate genes from the partial assembly of the H69J14 clone sequence (**Table 2**). Three of the candidate genes are 94 to 98% identical to each other and they are also 50 to 70 % identical to four PN40024 genes (**Figures 2 and 3**). We have initiated cloning work with candidate gene C4000-1. In later stages, as the fosmid sequence data becomes available, more detailed analysis will be carried out.

CONCLUSIONS:

Genetic mapping efforts have identified valuable genetic markers for marker-assisted selection and enabled rapid progress towards PD resistant winegrapes (see companion report). These mapping efforts have now identified three alleles of *PdR1*: *PdR1a* and *PdR1b* derived from *V. arizonica / candi*cans b43-17; and *PdR1c* derived from *V. arizonica* b40-14. These alleles were found to map within the same general region, but suggest that although PdR1 seems to be a single gene trait, the region may be composed of a number of tightly linked genes. BAC library sequence analysis of b43-17 is resulting in candidate genes suggestions for *PdR1* and these are being compared to the PN40024 genome sequence and to similar regions in other plants. The genomic characterization of this region will help us determine how this form of PD resistance functions and which genes control it. Genetic mapping of the multigenic source of resistance, b42-26, is progressing and tightly linked markers will greatly expedite the interbreeding of this resistance with that from b43-17 to increase the durability of PD resistance.

Table 1. Details for the transposable elements, sizes of exons and introns in four genes that are present in the Pinot noir (PN40024) 12X genome sequence in the region correlated to the *PdR1* region of b43-17. The analysis of the sequence was carried out with Censor program, which compares the sequence to known repeat regions from *Vitis*, Maize, Wheat and Pine.

Gene ID (12X genome of PN40024)	Gene size (bp)	No. of exons and introns	Size without introns (bp)	Repeat class categories	Fragments	Length
GSVIVT01033116001	7,729	6, 5	2,496	DNA transposon	1	294
GSVIVT01001802001	33,894	10, 9	3,360	Transposable Element	31	13969
				DNA transposon	18	2747
				EnSpm	3	485
				Harbinger	3	305
				Helitron	3	474
				MuDR	5	640
				hAT	1	408
				LTR Retrotransposon	12	11164
				Copia	3	406
				Gypsy	8	10613
				Non-LTR Retrotransposon		
				L1	1	58
GSVIVT01001803001	11,310	10, 9	3,309	Transposable Element	8	2453
				DNA transposon	7	2182
				Harbinger	1	377
				MuDR	4	1121
				hAT	2	684
				LTR Retrotransposon	1	271
				Gypsy	1	271
GSVIVT01001804001	12,165	11, 10	2,691	Transposable Element	9	2778
				DNA transposon	8	2520
				Harbinger	1	383
				MuDR	5	1450
				hAT	2	687
				LTR Retrotransposon	1	258
				Gypsy	1	258

Table 2. Information for four candidate genes from a partial assembly of the resistance region of b43-17.

Contig ID	size (Kb)	current gene ID	Size (Kb)
Contig 4000	47.3	Contig 4000-1	3.07
Contig 4002	25.1	Contig 4002-1	2.12
Contig 3995	7.34	Contig 3995-1	2.97
Contig 3974	22.2	Contig 3974-1	2.94

REFERENCES CITED

Moisy C., Garrison K.E., Meredith C.P., Pelsy F. 2008. Characterization of ten novel TyI/copia-like retrotransposon families of the grapevine genome. BMC Genomics 9:469

Riaz, S., S. Vezzulli, E.S. Harbertson, and M.A. Walker. 2007. Use of molecular markers to correct grape breeding errors and determine the identity of novel sources of resistance to *Xiphinema index* and Pierce's disease. Amer. J. Enol. Viticult. 58:494-498.

Riaz, S, A.C. Tenscher, B.P. Smith, D.A. Ng and M.A. Walker. 2008. Use of SSR markers to assess identity, pedigree, and diversity of cultivated muscadine grapes. J. Amer. Soc. Hort. Sci. 133:559-568

Riaz, S., A.C. Tenscher, J. Rubin, R. Graziani, S.S. Pao and M.A, Walker. 2008. Fine-scale genetic mapping of two Pierce's disease resistance loci and a major segregation distortion region on chromosome 14 of grape. Theor. Appl. Genet. 117:671-681

Riaz, S., A.C. Tenscher, R. Graziani, A.F. Krivanek and M.A. Walker. 2009. Using marker-assisted selection to breed Pierce's disease resistant grapes. Amer. J. Enol. Viticult. 60:199-207.

FUNDING AGENCIES

Funding for this project was provided by the CDFA Pierce's Disease and Glassy-winged Sharpshooter Board. Additional support from the Louis P. Martini Endowed Chair in Viticulture is also gratefully acknowledged.

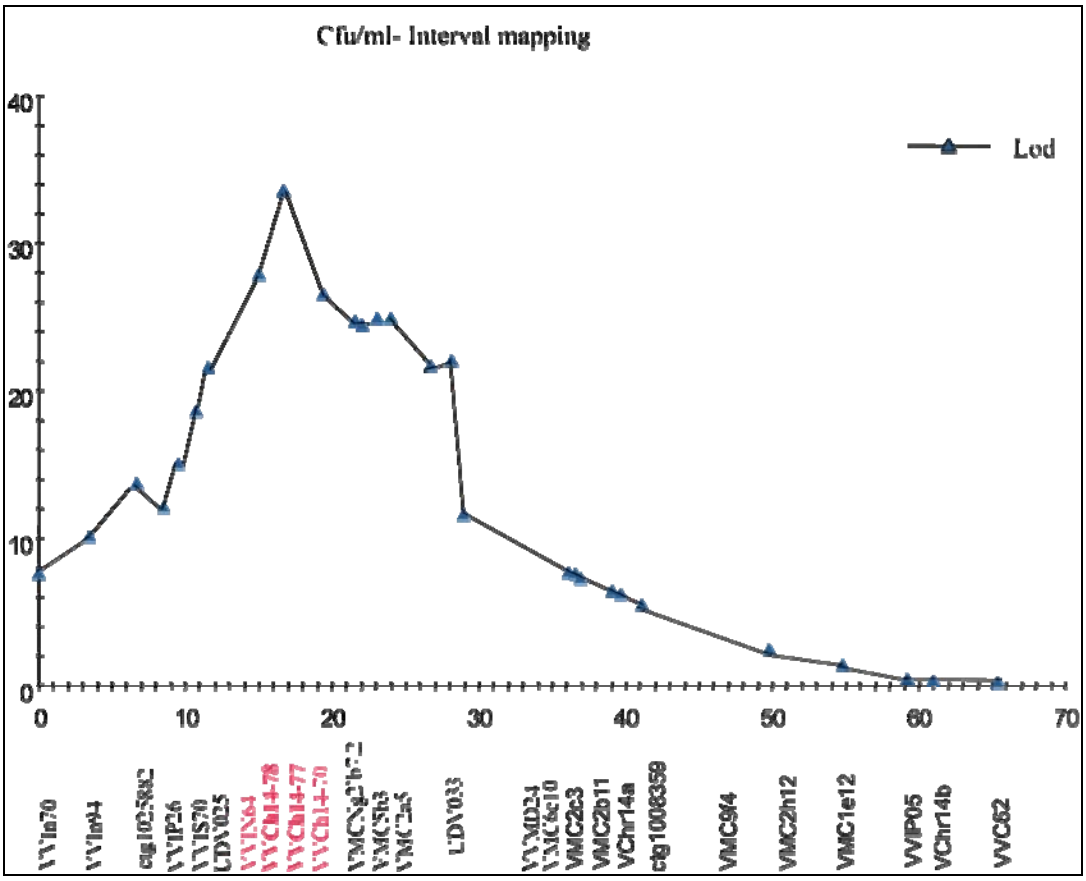

Figure 1. Interval mapping of *PdR1* indicating a peak at LDD 34.0 with the 95% confidence interval. The X-axis indicates the position of the markers; LOD values are plotted on the Y-axis.

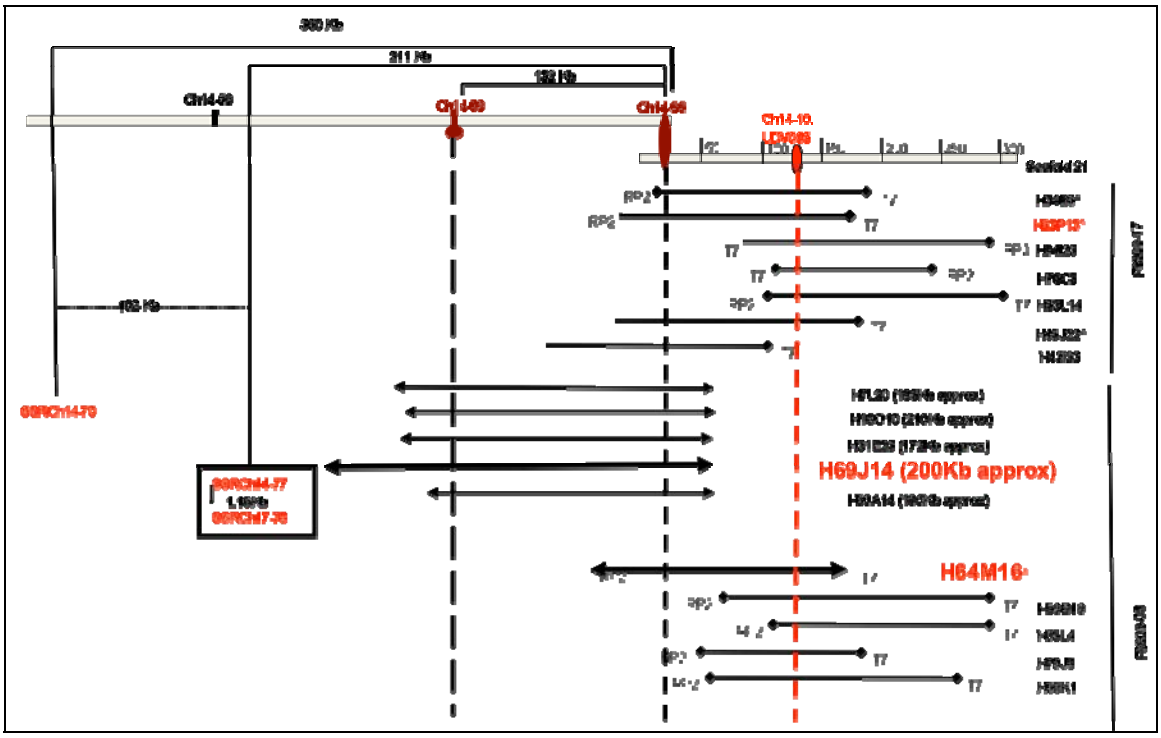

Figure 2. The arrangement of BAC clones and relative positions in comparison to the Pinot noir (PN40024) sequence. Three markers in red were used to screen the BAC library. H69J14 clone has been sequenced via 454 and sanger shotgun method. H64M16 clone has been sequenced by sanger method only. Both clones overlap for 60Kb approximately. These two clones represent the haplotype *PdR1b* region of b43-17.

Percent Identity

	1	2	3	4	5	6	7	8		
1		38.6	94.9	95.3	55.1	54.2	65.2	77.8	1	contig 4000-1.seq
2	131.9		39.3	39.3	36.9	35.8	35.4	37.4	2	contig4002-1.seq
3	5.2	128.1		98.7	57.5	56.2	66.8	78.9	3	contig3995-1.seq
4	4.9	128.1	1.1		57.9	56.5	67.3	79.2	4	contig3974-1.seq
5	69.3	143.3	63.2	62.6		66.9	39.9	60.6	5	PN-2001.seq
6	71.3	149.3	66.1	65.5	44.3		39.8	56.3	6	PN-3001.seq
7	47.0	152.9	43.9	43.1	126.0	124.3		51.3	7	PN-4001.seq
8	26.4	139.7	24.6	24.5	56.8	66.3	80.1		8	PN-6001.seq
	1	2	3	4	5	6	7	8		

Figure 3. The sequence comparisons of genes identified from the Pinot noir (PN40024) and from the partial assembly of the BAC clone of the resistant genotype b43-17