THE MIAME/PLANT
(based on the MIAME/PLANT protocol: http://www.mged.org/Workgroups/MIAME/MIAME-plant_Dec2005.pdf)

Experiment Design:
The goal of the experiment: To identify differentially expressed genes in the two grapevine genotypes, *Vitis vinifera* 'Cabernet sauvignon' and *Vitis aestivalis* 'Norton', in response to the powdery mildew (PM) fungus, *Erysiphe necator* (synonym *Uncinula necator* (Schw.) Burr.).

I. Array Design Description
Affymetrix GeneChip *Vitis vinifera* Genome Array is a 100-format, 11-micron array design. Each transcript is represented by 16 probes with oligonucleotide length of 25-mer per sequence. The array contains 16,437 probe sets that were derived from 14,509 ESTs of *V. vinifera* and 1,922 ESTs of non-vinifera *Vitis* species. It is manufactured by Affymetrix, Inc. (3380 Gentral Expressway, Santa Clara, CA 95051 USA). Catalog number for the 6-array is 900510. Catalog number for GeneChip one-cycle target labeling and control reagents is 900493.

II. Experiment Description
1. Plant experimental design
1) Pooling of samples
One leaf was harvested from each plant and ten leaves were pooled as one sample at each time point. Powdery mildew conidia-inoculated and mock-inoculated samples were ground separately in liquid nitrogen.

2) Experimental design

Treatment	Conidia inoculation						Mock inoculation					
Time Point	0 hr	4 hr	8 hr	12 hr	24 hr	48 hr	0 hr	4 hr	8 hr	12 hr	24 hr	48 hr
Repeat 1	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)
Repeat 2	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)
Repeat 3	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)
Repeat 4	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)
Repeat 5	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)	1 chip (10-plant sample)

Sixty plants of the identical developmental stage were grown separately in two identical growth chambers that were labeled as PM-Inoculation and Mock-inoculation, respectively. Vines of the same size were labeled, one with an "I" prefix for "PM-inoculated", the other with an "M" prefix "mock-inoculated". The "I" and "M" vines of a pair received the same ID number and were then placed in corresponding locations on the bench of the two growth chambers. The locations of the vines were determined by randomizing their ID numbers. Thus, the "I" and "M" groups were randomized according to their size and to their locations within the growth chamber. All vines were positioned at the same height from the light source. The locations of each individual plant within the growth chamber were recorded. At each time point, 10 randomly selected pairs of plants were sampled. (For example, at 8 hours post inoculation, vines I-3, I-17, I -23, …etc., and vines M-3, M-17, M-23, ...etc. were sampled from the "I" and the "M" growth chambers, respectively). The 10 corresponding leaves were combined into a single sample. At the next time

point, another set of randomly selected 10 pairs were sampled from the remaining plants, and so on. This random selection was independent from the random selection implemented to determine the locations of the vines in the growth chambers. The entire experiment was repeated five times. Samples of three replicates were processed for analysis.

2. Plant Samples used, RNA extraction and labeling

1) Biosource properties

Plant Strain or line Genotype

Vitis vinifera 'Cabernet sauvignon' and *Vitis aestivalis* 'Norton' were used. Hardwood cuttings of *V. vinifera* 'Cabernet sauvignon' Clone No. 4, which was certified to be free of viruses, were donated by the Sunridge Nurseries (Sunridge Nurseries, 441 Vineland Rd. Bakersfield, CA 93307). Hardwood cuttings of *V. aestivalis* 'Norton' were collected from the 7 year-old Norton vines that are grown in the Missouri State Fruit Experiment Station (9740 Red Spring Road, Mountain Grove, MO 65711).

Starting material

Hardwood cuttings with three buds were propagated. To promote rooting, the bottom ends of cuttings were cut and dipped in 1% indole-butyric acid solution and the cuttings were placed in a rooting bed (soil temperature at roots: 26°C; ambient temperature at buds: 4°C). After the hard-wood cuttings were rooted, they were planted in pots individually (1 plant in each pot). The vines were allowed to grow in the greenhouse where natural PM infection was prevented by vaporized sulfur. Vines were watered with commercial tap water every other day.

Developmental stage

One single shoot was allowed to grow 8 to 10 inches long for each vine.

Organism part

Fully-expanded leaves at the third or fourth position from the tip of each shoot were used for inoculation.

2) Biomaterial manipulations

Growth substrates

Promix BX horticultural medium (Premier Horticulture, Inc. 127 South 5th Street, Suite 300. Quakertown, PA 18951) in 3.785 L pots.

Growth environment

Conviron growth chambers

Environmental conditions

Duration: 14 hours day / 10 hours night. Light intensity: 500 μmol $m^{-2}s^{-1}$. Light source: high pressure sodium and metal halide. Temperature: 25°C. Relative humidity: 85%.

Treatment type

Biotic factors

Fungus: *Erysiphe necator*. A single colony of the PM fungus was isolated originally from a naturally infected leaf of *V. vinifera* 'Cabernet sauvignon'. PM colonies were maintained on

detached *V. vinifera* 'Cabernet sauvignon' leaves that were placed on 1% water agar plates inside Petri dishes and transferred to the next set of sterilized leaves every 10 to 12 days. Ten day-old sporulating colonies of PM were used as the inoculum. Powdery mildew and mock inoculations were performed simultaneously at one time for all 60 pairs of plants. For powdery mildew inoculation, conidia were deposited onto the third or fourth fully expanded leaf of an "I" plant by gently touching a PM-infected leaf onto the target leaf. Mock inoculation was done similarly with a non-infected leaf.

3) Extraction method

Total RNA was extracted from each sample for each replicate. The leaves were ground in liquid nitrogen and homogenized at 1g per 15 ml of extraction buffer (2% hexadecyltrimethyl ammonium bromide, 1% sodium dodecyl sulfate, 2.5M NaCl, 0.5M Tris, 50mM EDTA, 5% β-marcaptoethanol, and 3% polyvinyl poly-pyrrolidone). Following 30-min incubation at -80°C, the samples were thawed and centrifuged. The supernatant was supplemented with 1/30 volume of 3 M sodium acetate (pH5.2) and 1/10 volume of ethanol, incubated on ice and centrifuged three times at 7000g, 30 min 4 ºC, each time transferring the supernatant into a clean tube. The supernatant was supplemented with 1/9 volume of 3 M sodium acetate (pH 5.2) and iso-propanol to a final concentration of 33%. Following 60-min incubation at -20°C, the RNA was collected by centrifugation at 12,000g, 30 min, 4 ºC. The RNA was then supplemented with 1/3 volume of 8 M LiCl and incubated overnight at 4 °C. RNA was then collected by centrifugation at 12,000g, 30 min, 4 ºC, washed with 75% ethanol, and resuspended in 40 µl TE buffer. Total RNA were treated with DNase I in TURBO DNA-free™ reagents (Ambion, Inc. Austin, TX) and purified using RNeasy MinElute Cleanup Kit (Qiagen, Valencia, CA).

Array hybridizations, data processing, and data quality assessment

The following procedures were completed at the Penn Microarray Facility, University of Pennsylvania, following standard protocol as recommended by Affymetrix manual in the website (http://www.affymetrix.com/support/technical/manual/expression_manual.affx). Quality of total RNA, cDNA, Complementary RNA (cRNA) and fragmented cRNA was verified on an Agilent 2100 bioanalyzer. Total RNA samples (4 µg) were used as the template for cDNA synthesis and linear RNA amplification. The cDNA synthesis and antisense cRNA amplification/biotin labeling was performed using the One-Cycle Target Labeling Kit (Catalog number: 900493, Affymetrix, Santa Clara, California). The cRNA was cleaned up by the Affymetrix sample cleanup module which is included in the One-Cycle Target Labeling kit. Fifteen micrograms of cRNA at concentration of 1.301 µg/µl were used for hybridization. Prior to hybridization, cRNA was heated at 94ºC for 35 min, and was fragmented to 35 to 200 nucleotides. After adding hybridization buffer to the fragmented samples, hybridization cocktails were boiled for 5 min, incubated for 5 min at 45ºC, then centrifuged for 5 minutes prior to loading per standard one-cycle protocol. Hybridization took place in an Affy hybridization oven model 640, for 16 h at 45ºC. All washing took place on the fluidics station (as per the recommended fluidics protocol) with non-stringent wash buffer (6X SSPE and 0.01% Tween-20) and Stringent wash buffer (100mM MES, 0.1M NaCl, 0.01% Tween-20). Fluorescence was amplified by first adding streptavidin-phycoerythrin (SAPE) stain, then adding a biotinylated antibody (anti-streptavidin) solution, followed by another SAPE staining (as per standard one-cycle protocol). The fluidics protocol Midi_Euk2v3 was used (as per *V. vinifera* microarray package insert) on fluidics station

model 450. Affymetrix Scanner 3000 7G was used to collect signals at resolution of 1.56 μm with emission filter at a wavelength of 570 nm and excitation wavelength of 532 nm.

Data processing: Raw intensity data were processed using the GeneChip Operating Software Version 1.2 (GCOS 1.2; Affymetrix, 2001). Background correction and expression value calculation were performed as described by Affymetrix (2002). Normalization was done by global scaling, with a target intensity value of 150 (Affymetrix, 2001). Probe sets with missing values were deleted across all replicates and genotypes to allow comparisons among datasets.

Normalization: The median signal value for each chip was computed using all probes except for the controls. All observations from a given chip were divided by the chip median, and then log-transformed using the natural logarithm. If all 36 samples for a particular feature were deemed 'absent', that feature was removed from further consideration. Of 16,437 non-control features, 1858 were always 'absent' and removed, the remaining features (14,579) were analyzed in the 36 GeneChips for *Vitis vinifera*. Of 16,437 non-control features, 1992 were always absent and removed, the remaining features (14,445) were analyzed in the 36 GeneChips for *Vitis aestivalis.*

Individual statistical analysis of each *V. vinifera* 'Cabernet sauvignon' and *V. aestivalis* 'Norton' hybridization data to identify the features that are significantly changed in response to the PM inoculation

ANOVA

For each feature separately the linear model $Y_{ijk} = \mu + A_i + B_j + (AB)_{ij} + \gamma_k + \varepsilon_{ijk}$ was fit. In each ANOVA, Y_{ijk} is the normalized transcript level (normalization described above) for the i^{th} inoculation treatment at time j in replicate k; μ is the mean expression for the feature over all time points and treatments; A_i represents the effect of the i^{th} inoculation treatment (Mock-Inoculated, and PM-Inoculated); B_i represents the effect of the j^{th} time point (j = 0, 4, 8, 12, 24, and 48); $(AB)_{ij}$ represents the interaction between the inoculation treatment and time, γ_k is the effect of the replicate; and ε_{ijk} is a random error term for the ijk^{th} observation. Replicate effects were considered random while the other effects were considered fixed. The initial design was a split plot with treatment as the whole plot factor. However, the whole plot error, that is block by treatment interaction, showed no evidence for differences from zero. Thus, in the absence of compelling evidence for block interaction with treatment, we fit only the block as random.

An F test of the effect of PM-inoculation by time point interaction for each gene was conducted and the P-value for the test of the null hypothesis $(AB)_{ij} = 0$ for all ij (i.e., the mean expression profiles for the inoculated and mock inoculated treatments are parallel) was calculated. None of the 14,579 tests were significant even if we use an FDR of 10% in 36 GeneChips for *V. vinifera*. We use the False Discovery Rate (FDR) to control type I error (Benjamini and Hochberg, 1995) with the FDR correction suggested in this publication. None of the 14,445 tests were significant using an FDR of 10% or a nominal alpha of 0.01 in 36 GeneChips for *V. aestivalis*.

To test the effect of treatment, a second F test for the effect of the PM-inoculation across all time points (i.e. mean expression levels of the PM-inoculated and mock-inoculated treatments are the same) was constructed. An FDR significance level (Benjamini and Hochberg 1995) was used as an initial criterion for rejecting the null hypotheses of significant PM-inoculation effects (FDR = 5%). If the test of the null hypothesis that the mean expression levels of the PM-

inoculated and mock-inoculated treatments are the same for all time points was rejected, we declared the gene differentially expressed (transcribed) across PM-inoculation treatments. A total of 626 transcripts were declared to be differentially expressed at FDR 0.05 in GeneChips for *V. vinifera* that were hybridized with cRNAs that were made from PM-inoculated samples. Four transcripts were declared to be differentially expressed at FDR 0.05 in GeneChips for *V. aestivalis* that were hybridized with cRNAs that were derived from PM-inoculated samples. For these features, we examined additional contrasts comparing the effect of the PM-inoculation treatments at each time point. If the P-value for this test of differences between treatments at a given time point was greater than 0.05, the expression levels for that time point was classified as "SAME". If the P-value for this test of differences between treatments at a given time point was equal to or less than 0.05, and the PM-inoculation treatment had the highest expression level, the comparison was classified as "UP". If the P-value for this test of differences between treatments at a given time point was equal to or less than 0.05, and the PM-inoculation treatment had the lowest expression level, the comparison was classified as "DOWN".

We examined the model for conformation to the assumption of normality of the residuals by testing the null hypothesis that the residuals for each gene were normally distributed using the Shapiro-Wilkes Test and a nominal alpha of 0.05. These models showed no overall evidence for concern, with only about 6% of the features showing any evidence for violation of the normality assumption, well within what would be expected by chance alone.

Combined statistical analysis of both *V. vinifera* 'Cabernet sauvignon' and *V. aestivalis* 'Norton'

ANOVA

For each feature separately the linear model $Y_{ijkm} = \mu + V_i + T_j + (VT)_{ij} + \gamma_{k(i)} + t_m + (Vt)_{im} + (Tt)_{jm} + (VTt)_{ijm} + \varepsilon_{ijkm}$ was fit. In each ANOVA, Y_{ijkm} is a normalized transcript level for the i^{th} genotypes in inoculation treatment j at time m in replicate k; μ is the mean expression for the feature over all varieties, time points and treatments; V_i represents the effect of the i^{th} genotypes (*V. vinifera* 'Cabernet sauvignon' and *V. aestivalis* 'Norton'), T_j represents the effect of the j^{th} inoculation treatment (Mock-Inoculated and PM-Inoculated); t_m represents the effect of the m^{th} time point (j = 0, 4, 8, 12, 24, and 48 hours); $\gamma_{k(i)}$ is the effect of the k^{th} replicate within variety i; and ε_{ijkm} is a random error term for the ijkmth observation. Replicate effects were considered random while the other effects were considered fixed.

There were some issues with heteroscedasticity of error variances across varieties, although neither genotypes was consistently more variable. Therefore, the distribution of the residuals ε_{ijkm} was assumed normal with error variance σ^2_N for residuals associated with *V. aestivalis* observations and σ^2_C for residuals associated with *V. vinifera* observations.

The same F tests built for the separate analyses were constructed using the combined model. The results were identical to those obtained in the separate ANOVAs (no significant interactions treatment by time in any variety, 4 significant treatment effects in *V. aestivalis*, and 626 significant treatment effects in *V. vinifera*). To further determine whether these features were statistically different in their PM response between the genotypes, we tested the null hypothesis that $(VT)_{ij}=0$ for these features. We used an FDR level of 0.2 for this test as tests for interaction have lower power than tests for main effects. We found that 533 of them showed evidence for differential response to PM treatment between the genotypes.

To investigate whether expression levels were constitutively higher in *V. aestivalis* than in *V. vinifera*, an F test was conducted to compare the expression levels of the two grapevine

genotypes under the mock inoculation condition. This tests the null hypothesis that the mean expression levels of the mock-inoculated treatments in the two grapevine genotypes are the same across time points. If this test of the null hypothesis was rejected at an FDR of 5%, we examined additional contrasts comparing the effect of the two genotypes at each time point individually. If the nominal P-value for this test of differences between the two genotypes at a given time point was greater than 0.01, the expression levels for that time point was classified as "SAME" between the two genotypes. If the P-value for this test of differences between the two genotypes at a given time point was less or equal to than 0.01, and *V. aestivalis* had the highest expression level, the comparison was classified as "UP". If the P-value for this test of differences between the two genotypes at a given time point was less or equal to than 0.01, and *V. aestivalis* had the lowest expression level, the comparison was classified as "DOWN". Least squared means from the mixed model were used as estimates for the quantity in each condition and as the original data was transformed using the natural logarithm, the fold was calculated as the antilog of the difference between the lsmeans from the two conditions.

We examined the model for conformation to the assumption of normality of the residuals by testing the null hypothesis that the residuals for each gene were normally distributed using the Shapiro-Wilk's Test and a nominal alpha of 0.05. These models showed no overall evidence for concern, with only about 6% of the features showing any evidence for violation of the normality assumption in *V. aestivalis* and about 9% of the features showing any evidence for violation of the normality assumption in *V. vinifera*.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical approach to multiple testing. Journal of the Royal Statistical Society Series B-Statistical Methodology **57:** 289-300