A systems approach to elucidate heterosis of protein abundances in yeast

Melisande Blein-Nicolas, Warren Albertin, Telma da Silva, Benoît Valot,
Thierry T. Balliau, Isabelle Masneuf-Pomarede, Marina Bely, Philippe
Marullo, Delphine Sicard, Christine Dillmann, et al.

HAL Id: hal-02636704
https://hal.inrae.fr/hal-02636704

Submitted on 8 Jun 2021

A Systems Approach to Elucidate Heterosis of Protein Abundances in Yeast*⒮

Mélisande Blein-Nicolas[A], **Warren Albertin**[B,C], **Telma da Silva**[A,D], **Benoît Valot**[E], **Thierry Balliau**[A], **Isabelle Masneuf-Pomarède**[C,F], **Marina Bely**[C], **Philippe Marullo**[C,G], **Delphine Sicard**[H,I], **Christine Dillmann**[H], **Dominique de Vienne**[H], and **Michel Zivy**[J,K]

Heterosis is a universal phenomenon that has major implications in evolution and is of tremendous agro-economic value. To study the molecular manifestations of heterosis and to find factors that maximize its strength, we implemented a large-scale proteomic experiment in yeast. We analyzed the inheritance of 1,396 proteins in 55 inter- and intraspecific hybrids obtained from *Saccharomyces cerevisiae* and *S. uvarum* that were grown in grape juice at two temperatures. We showed that the proportion of heterotic proteins was highly variable depending on the parental strain and on the temperature considered. For intraspecific hybrids, this proportion was higher at non-optimal temperature. Unexpectedly, heterosis for protein abundance was strongly biased toward positive values in interspecific hybrids but not in intraspecific hybrids. Computer modeling showed that this observation could be accounted for by assuming concave relationships between protein abundances and their controlling factors, in line with the metabolic model of heterosis. These results point to nonlinear processes that could play a central role in heterosis. *Molecular & Cellular Proteomics 14: 10.1074/mcp.M115.048058, 2056–2071, 2015.*

Nonadditive inheritance in hybrids, whereby the phenotype of offspring is not the average of the parental phenotypes, is commonly observed in all species. For monogenic traits, the departure from additivity is called dominance (1) or overdominance if the phenotypic value of the hybrid is outside the range defined by the parental values (2). For polygenic traits, it is called heterosis. Heterosis is commonly associated to macroscopic traits, but it also applies to less integrated traits such as metabolite abundances (3, 4), fluxes and enzyme activities (5–7), mRNA, and protein amounts (8).

The concept of heterosis is not universally shared and depends on the scientific communities. Strictly speaking, heterosis is defined as the superiority of the hybrid over the mean parental value (mid-parent heterosis, MPH[1]) or over its parent exhibiting the highest value (best-parent heterosis, BPH). This definition and the associated terminology are historical and come from the fact that heterosis was commonly associated to traits such as growth rate, biomass, size, yield, or fertility, for which higher values are beneficial. However, lower values can also occur (*e.g.* (9)). The definition of heterosis has therefore been broadened to include also the cases where the hybrid is below the mean parental value (negative MPH) or below its parent exhibiting the lowest value (worst-parent heterosis, WPH). In this paper, we will adopt this second definition, motivated by its lack of presumption about whether the changes observed in hybrids are beneficial or detrimental.

Heterosis has fascinated scientists and breeders for more than 100 years. It has major implications in evolution and domestication of crop plants (10, 11), and it has been exploited since the 1930s in plant breeding to produce hybrids of high agronomic value (12). In this context, heterosis has proven to efficiently accelerate the process of selection for various crops (reviewed in (13)). Heterosis is opposite to inbreeding depression, that is supposed to be predominantly caused by the homozygosity of deleterious recessive alleles (14). Heterosis can provide a heterozygote advantage by buffering against these alleles and confers genetic plasticity to adapt to environmental changes (11).

[1] The abbreviations used are: ANOVA: analysis of variance; MPH, mid-parent heterosis; BPH, best-parent heterosis; CFU, colony-forming unit; LC-MS/MS, liquid chromatography-tandem mass spectrometry; RPM, round per minute; TF, transcription factor; WPH, worst-parent heterosis; YPD, yeast extract peptone dextrose; ZALS, zwitterionic acid labile surfactant.

Given the importance of heterosis for agriculture and because it is an intriguing phenomenon, many studies have focused on the understanding of its genetic and molecular bases (11, 15–25). Three nonexclusive hypotheses based on genetic effects are classically put forward to explain heterosis. First, the dominance hypothesis attributes heterosis to complementation: In the hybrid, the recessive alleles are masked by dominant and generally favorable alleles (26, 27). Second, the overdominance hypothesis assumes that heterosis is inherent to the heterozygous state (2, 28). Third, the epistasis hypothesis proposes that heterosis is due to intergenic interactions created in the hybrid (29, 30). Scientists have long sought a unifying theory to account for heterosis, but it is now commonly admitted that this phenomenon likely arises from the combination of several genetic mechanisms, the relative effects of which vary according to the trait, the cross, or the species (23, 25).

These genetic effects are consistent with the factors known to maximize the occurrence of heterosis. When compiling the results obtained so far across numerous studies, it appears that heterosis is of greatest magnitude for highly integrated, and hence polygenic, traits such as crop yield (23, 31); it is larger in allogamous than in autogamous species (31); it requires genetic divergence between parents; and interspecific crosses commonly produce higher levels of heterosis (21, 24, 32). However, these general trends are not sufficient for a reliable prediction of heterosis, which is a major challenge for plant and animal breeding and for biotechnology. Future strategies for heterosis prediction will have to rely both on an accurate description of its manifestations and on the detailed knowledge of the factors that maximize its strength. To address these issues, we performed a large-scale study of heterosis by analyzing the inheritance of the abundance of a high number of proteins in an unprecedented number of yeast hybrids grown in two conditions. The proteomic level is particularly relevant to the large-scale study of heterosis because protein abundances are polygenic molecular traits (33) that can be measured by high-throughput quantitative proteomics (34, 35).

Yeast has only rarely been used to study heterosis (9, 36–41). Yet, it is amenable to large-scale laboratory experiments, and it is of major industrial interest for wine making. Hybrids with exceptional performances were reported in *Saccharomyces cerevisiae* (36, 38, 42, 43), and several observations indicate that interspecific hybridization could be used in breeding to produce improved strains for wine making. For instance, many interspecific hybrids between *S. cerevisiae* and *S. uvarum* Beijerinck or *S. kudriavzevii* isolated in wine and natural environments showed important biotechnological potential, such as a better robustness than their parents (44–47). In addition, several wine strains empirically selected for their biotechnological properties proved to be interspecific hybrids (48, 49). For all these reasons, we chose to study heterosis for protein abundance in yeast strains from *S.*

cerevisiae and *S. uvarum*, which are the two main species associated with grape juice fermentation (50).

By using shotgun proteomics, we analyzed more than 1,300 proteins in an experimental design, including 11 parental strains of *S. cerevisiae* and *S. uvarum* and their 55 intra- and interspecific hybrids, which were grown at two temperatures to take into account adaptation differences between parental species (18 °C and 26 °C optimal for *S. uvarum* and *S. cerevisiae*, respectively (44, 51, 52)). We showed that heterosis for protein abundance was strongly biased toward positive values in interspecific hybrids but not in intraspecific hybrids, which, to our knowledge, has never been reported. We also showed that our experimental results were consistent with results obtained from modeling approaches assuming nonlinear relationships between protein abundances and their controlling factors.

EXPERIMENTAL PROCEDURES

Yeast Strains—Four diploid *S. uvarum* strains, seven diploid *S. cerevisiae* strains, and their 55 hybrids produced from a half diallel design (53) were analyzed in this study. Parental strains were derived from strains isolated from different geographical locations and from either natural or food-processing origins (Table I): the *S. cerevisiae* strains were isolated from diverse media (distillery, enology, oak exudate) to maximize the genetic diversity within this species (54); the *S. uvarum* strains, originating from grape must or cider fermentation, were chosen to cover a wide part of the genetic diversity of the *S. uvarum* species (Masneuf-Pomarède, I., personal communication). For each original strain, one meiospore was isolated with a micromanipulator (Singer MSM Manual, Singer Instrument, Somerset, UK). All the original strains but Alcotec 24 were homothallic (*HO/HO*); therefore, fully homozygous diploid strains were spontaneously obtained by fusion of opposite mating type cells. For A24 (*ho/ho*), one isolated haploid meiospore was diploidized via transient expression of the HO endonuclease (55). All strains were grown at 30 °C in YPD medium containing 1% yeast extract (Difco Laboratories, Detroit, MI), 1% bactopeptone (Difco), and 2% glucose, supplemented or not with 2% agar. When necessary, antibiotics were added at the following concentrations: 100 μg/ml for G418 (Sigma, L'Isle d'Abeau, France), and nourseothricin (Werner bioagent, Jena, Germany) and 300 μg/ml for hygromycin B (Sigma).

Construction of the Half Diallel Design—Hybrid construction was performed as described in Albertin *et al.* (53). Briefly, the 11 diploid parental strains were transformed with a cassette containing the *HO* allele disrupted by a gene of resistance to either G418 (*ho::KanR*), hygromycin B (*ho::HygR*), or nourseothricin (*ho::NatR*). Strain transformation allowed conversion to heterothallism for the homothallic strains. Then the mating-type (*MATa* or *MATalpha*) of antibiotic-resistant monosporic clones were determined using testers of well-known mating type. For each cross, parental strains of opposite mating type were put in contact 2 to 6 h in YPD medium at room temperature and plated on YPD-agar containing the appropriate antibiotics. Ten independent hybrids per cross were recovered. After recurrent cultures on YPD-agar corresponding to ∼ 80 generations, the nuclear chromosomal stability of the hybrids was controlled by pulsed field electrophoresis (CHEF-DRIII, Bio-Rad, Marnes-La-Coquette, France) as well as homoplasmy (only one parental mitochondrial genome). One hybrid per cross was finally retained for further experiments.

Yeast Strain Characterization—Two polymorphic microsatellites specific to *S. cerevisiae* (Sc-YFR038 and Sc-YML091 (56)) and two

TABLE I
Origin of parental strains

Species	Parental strains	Monosporic derivate	Collection/Supplier[a]	Isolation origin	Area of origin	Reference
S. uvarum	PM12	U1	ISVV	Grape must fermentation	Jurançon, France	(98)
S. uvarum	PJP3	U2	ISVV	Grape must fermentation	Sancerre, France	(98)
S. uvarum	Br6.2	U3	ADRIA NORMANDIE	Cider fermentation	Normandie, France	(95)
S. uvarum	RC4–15	U4	ISVV	Grape must fermentation	Alsace, France	(57)
S. cerevisiae	CLIB-294	D1	CIRM-Levures	Distillery	Cognac, France	(98)
S. cerevisiae	Alcotec 24	D2	Hambleton Bard	Distillery	UK	(100)
S. cerevisiae	CLIB-328	E1	CIRM-Levures	Enology	UK	(100)
S. cerevisiae	BO213	E2	LAFFORT Oenologie	Enology	France	(101)
S. cerevisiae	F10	E4	LAFFORT Oenologie	Enology	Bordeaux, France	(101)
S. cerevisiae	VL3	E5	LAFFORT Oenologie	Enology	Bordeaux, France	(102)
S. cerevisiae	YPS128	W1	SGRP	Forest, oak exudate	Pennsylvania, USA	(54)

[a] ISVV, http://www.oenologie.u-bordeaux2.fr/.
ADRIA NORMANDIE, http://www.adria-normandie.com.
SGRP, http://www.sanger.ac.uk/research/projects/genomeinformatics/sgrp.html.
CIRM-Levures, http://www.inra.fr/internet/Produits/cirmlevures.
Hambleton Bard, http://www.hambletonbard.com.
LAFFORT Œnologie, http://www.laffort.com.

specific to *S. uvarum* (locus 4 and 9 (57)) were used to discriminate rapidly the hybrids from the parental strains. These four markers were amplified in a multiplex PCR reaction (95 °C for 5 min for initial denaturation step; 95 °C for 30 s, 55 °C for 90 s, and 72 °C for 60 s repeated 35 times; a final elongation step of 30 min at 60 °C). The PCR products were analyzed on an ABI3730 apparatus (Applied Biosystem, Villebon-sur-Yvette, France), and microsatellite lengths were analyzed using the Peak Scanner tool (Applied Biosystem).

Alcoholic Fermentation in Grape Must—All the 66 strains (11 parents and 55 hybrids) were grown in the same batch of white grape must obtained from Sauvignon grapes harvested in vineyards in the Bordeaux area (2009 vintage). Tartaric acid precipitation was stabilized, and turbidity was adjusted to 100 NTU (nephelometric turbidity unit) before storage at -20 °C. The sugar concentration was 189 g.l^{-1}, the nitrogen content was 242 mg.l^{-1} and the pH was 3.3. The indigenous yeast population, estimated by YPD-plate counting after must thawing, was less than 20 CFU (colony-forming unit) *per* ml. Precultures of each strain were run in half-diluted must filtered through a 0.45 μm nitrate-cellulose membrane (24 °C, 150 rpm (rounds per minute)) during 24 h, after what one million cells per ml were sampled and added to a final volume of 125 ml of Sauvignon must. Then, fermentations were run into 125 ml glass reactors at two different temperatures (18 °C and 26 °C, 300 rpm) and repeated three times independently. In total, 396 alcoholic fermentations were performed (66 strains $\times$ 2 temperatures $\times$ 3 replicates) following a randomized experimental design. Of them, 31 failed due to the poor fermenting abilities of some strains (Table S1). The amount of CO_2 released was regularly determined by measurement of glass-reactor weight loss.

Protein Extraction and Digestion—Samples were harvested at 40% of CO_2 release to perform proteomic analyses. At this time, all strains had reached their maximum population size and performed alcoholic fermentation without growing. Only strain $\times$ temperature combinations with at least two successful fermentations were kept for further mass-spectrometry analysis (Table S1). Five milliliters of fermentative media were sampled and centrifuged (5 min, 2,750 g). The pellets were rinsed two times with 5 ml of water, frozen in liquid nitrogen, and stored at -80 °C until protein extraction. Total protein extracts were isolated via acetone precipitation as described in Blein-Nicolas *et al.* (58). Dried protein pellets were solubilized in 300 μl of a solution containing 6 M of urea, 2 M of thiurea; 10 mM of dithiothreitol (DTT); 30

mM of TrisHCl, pH 8.8; and 0,1% of zwitterionic acid labile surfactant (ZALS, Proteabio, Morgantown, WV, USA) and centrifuged for 10 min at 14,000 rpm. Protein concentration was determined using PlusOne 2-D Quant Kit (GE Healthcare, Velizy-Villacoublay, France) and adjusted to 4 μg.μl^{-1}. After a 10-times dilution in 50 mM of ammonium bicarbonate, proteins were reduced 1 h in 100 mM DTT, alkylated 1 h in 40 mM iodoacetamide, and digested overnight at 37 °C with 1/50 (w/w) trypsin (Promega, Charbonnière, France). Digestion was stopped by adding 0.4% of trifluoroacetic acid (TFA). Peptides were purified on solid phase extraction using polymeric C18 column (Phenomenex, Le Pecq, France) with a washing solution containing 0.06% acetic acid and 3% acetonitrile (ACN). After elution with 0.06% acetic acid and 70% ACN, peptides were speedvac-dried and suspended in 2% ACN and 0.08% TFA.

LC-MS/MS Analysis—LC-MS/MS analyses were performed using a NanoLC-Ultra System (nano2DUltra, Eksigent, Les Ulis, France) connected to a Q-Exactive mass spectrometer (Thermo Electron, Waltham, MA, USA). A 700 ng of protein digest were loaded onto a PepMap C18 precolumn (0.3 $\times$ 5 mm, 100 Å, 5 μm; NanoSeparations, Nieuwkoop, Netherlands) at 7.5 μl·min^{-1} and desalted with 0.1% formic acid and 2% ACN. After 3 min, the precolumn was connected to a PepMap C18 nanocolumn (0.075 $\times$ 150 mm, 100 Å, 3 μm). Buffers were 0.1% formic acid in water (A) and 0.1% formic acid and 100% ACN (B). Peptides were separated using a linear gradient from 5 to 35% buffer B for 40 min at 300 nl·min^{-1}. One run took 60 min, including the regeneration step at 100% buffer B and the equilibration step at 100% buffer A.

Ionization was performed with a 1.3-kV spray voltage applied to an uncoated capillary probe (10 μm tip inner diameter; New Objective, Woburn, MA, USA). Peptide ions were analyzed using Xcalibur 2.2 (Thermo Electron) with the following data-dependent acquisition steps: (1) MS scan (mass-to-charge ratio (*m/z*) 400 to 1,400, 70 000 resolution, profile mode), (2) MS/MS (17,500 resolution, collision energy = 30%, profile mode). Step 2 was repeated for the eight major ions detected in step 1. Dynamic exclusion was set to 40 s. Xcalibur raw data files were transformed to mzXML open source format using msconvert software in the ProteoWizard 3.0.3706 package (59). During conversion, MS and MS/MS data were centroided.

MS Data Availability—The raw MS output files were deposited online using PROTICdb database (60–62) at the following URL: http://moulon.inra.fr/protic/heterosyeast2.

Protein Identification—Protein identification was performed using the custom database described in Blein-Nicolas *et al.* (58). This database, containing 10,851 entries, was constructed from the translations of all systematically named ORFs of *S. cerevisiae* and *S. uvarum* downloaded from the Saccharomyces Genome Database (SGD project, http://www.yeastgenome.org/, versions dated October 5, 2010 and December 15, 2003, respectively). The proteins of *S. cerevisiae* and of *S. uvarum* encoded by orthologous genes were attributed unique labels. A contaminant database containing the sequences of standard contaminants and the sequences of 16 proteins of *Vitis vinifera* previously identified in extracts of yeast grown in grape juice was also interrogated. The decoy database comprised the reverse protein sequences of the custom database. Database search was performed with X!Tandem (version 2011.12.01.1; http://www.thegpm. org/TANDEM/) with the following settings. Enzymatic cleavage was declared as a trypsin digestion with one possible misscleavage. Carboxyamidomethylation of cysteine residues and oxidation of methionine residues were set to static and possible modifications, respectively. Precursor mass precision was set to 10 ppm. Fragment mass tolerance was 0.02 Th. A refinement search was added with the same settings, except that protein N-ter acetylations were also searched. Only peptides with an E-value smaller than 0.05 were reported.

Identified proteins were filtered and sorted by using X!Tandem-Pipeline (version 3.3.0, http://pappso.inra.fr/bioinfo/xtandempipe-line/). Criteria used for protein identification were (i) at least two different peptides identified with an E-value smaller than 0.05 and (ii) a protein E-value (product of unique peptide E-values) smaller than 10^{-4}. These criteria led to a false discovery rate estimated by using the decoy database of 0.12% and 1.15% for peptide and protein identification, respectively.

Peptide Quantification and Processing Intensity Data—Peptides were quantified based on extracted ion chromatograms using Mass-ChroQ software version 1.2.2 (63) with the parameters given in File S1. The detection threshold on min and max were set at 30,000 and 50,000, respectively. Due to progressive fouling of the quadrupole, sensitivity losses were observed over time, leading to a global decrease of measured intensities, particularly for hydrophobic peptides. To take these sensitivity losses into account, samples were classified according to their running order and divided into five blocks representing homogeneous global intensities. For each peptide, the block effect was retrieved and subtracted from intensity measures by using an analysis of variance (ANOVA). Then, normalization was performed to take into account possible global quantitative variations between LC-MS runs. For each LC-MS run, the ratio of all peptide values to their value in the chosen reference LC-MS run was computed. Normalization was performed by dividing peptide values by the median value of peptide ratios.

Raw data (containing intensity measures of 25,060 peptides) were then filtered to remove (i) dubious peptides for which standard deviation of retention time was superior to 60 s, (ii) peptide × strain × temperature combinations quantified in only one replicate, and (iii) peptides shared by several proteins, representing less than 5% of all the quantified peptides. To avoid bias on the estimation of total protein abundances in hybrids, we removed parent-specific peptides by using peptides presenting presence/absence variation among parental strains as a proxy. However, parent-specific peptides were confounded with species-specific peptides, which represented nearly 65% of the valid peptides. To exploit as far as possible the data available for intraspecies crosses, we thus split the dataset into three subsets: one contained *S. cerevisiae* triplets (hybrid and its parents), another contained *S. uvarum* triplets, and the last one contained interspecific triplets. Parent-specific peptides were removed separately in the three subsets. To finish, in order to estimate the peptide

effect properly, peptides quantified in less than four strains × temperature combinations in a given subset of data were removed.

Detection of Protein Abundance Changes—

Protein abundances were estimated independently in the three subsets of data by using

$$\log(l_{istr}) = \theta_{kst} + D_i + B_r + C_{tr} + \epsilon_{istr}$$

where l_{istr} is the normalized intensity value for peptide *i* in strain *s*, temperature *t,* and replicate *r,*

θ_{kst} is the natural logarithm of the abundance of protein *k* in strain *s* and temperature *t,*

$B_r \sim N\,(0,\sigma_B^2)$ is an error due to the biological variation of replicate,

$C_{str} \sim N\,(0,\sigma_C^2)$ is an error due to the technical variation of sample *str,*

$D_i \sim N\,(0,\sigma_D^2)$ is an error due to the LC-MS response of peptide *i,* and

$\epsilon_{istr} \sim N\,(0,\sigma_\epsilon^2)$ is the residual error.

Estimation of the parameters of the model was performed as described in Blein-Nicolas *et al.* (64). Protein abundance changes were detected by multiple test procedure across four different contrasts: (i) hybrid-mean of parents, (ii) hybrid–parent$_1$, (iii) hybrid–parent$_2$, (iv) parent$_1$–parent$_2$. Since several couples of strains × temperature combinations and several proteins were tested, p-values were adjusted for multiple testing by a Benjamini–Hochberg procedure (65). Of note, the statistical power was reduced in the subset of data containing interspecific hybrids compared with the two other subsets since intensity data were more drastically filtered (on average, there were 6.2 peptides per protein in the subset containing interspecific hybrids against 8.9 and 8.2 in the subsets containing *S. cerevisiae* hybrids and *S. uvarum* hybrids, respectively).

Data Analysis—Protein abundances estimated in different subsets of data were not directly comparable. To overcome this drawback, the subset of data containing interspecific hybrids (further named *B* for between) was taken as a reference, and the following linear regression was performed for each protein in the subsets of data containing intraspecific hybrids (referred to as *W* for within):

$$\theta_{pt}^W = a + b\theta_{pt}^B + \epsilon_{pt}$$

where θ_{pt}^W and θ_{pt}^B are the abundances estimated in parental strain *p* at temperature *t* in the subsets of data *W* and *B*, respectively,

a and b are the parameters of intercept and slope, respectively, and ϵ_{pt} is the residual error.

The median of the coefficient of determination R^2 was 0.83, indicating that the protein abundances estimated separately in different subsets of data were globally well correlated. For proteins with *b* significantly different from 0 (adjusted $p < .05$), estimators of *a* and *b* were used to correct the abundances estimated for intraspecific hybrids:

$$\omega_{ht}^W = (\theta_{ht}^W - \hat{a})/\hat{b}$$

where θ_{ht}^W is the abundance estimated in hybrid *h* at temperature *t* in the subset *W*. Then, protein abundances in the subset *B* were gathered with the ω_{ht}^W computed in the subset *W*.

A total of 615 proteins quantified in more than 122 strains × temperature combinations were kept for data representation as heat map and principal component analysis. Missing data were imputed from a uniform distribution with minimum = 0 and maximum = 10^6 under the hypothesis that they corresponded to low abundance values.

All data analyses and graphical representations were performed using R version 3.0.2 (66). Appropriate statistical tests were used for each kind of data: χ^2 tests were used to compare distributions; Student and Mann–Whitney tests were used to compare means in the

case of normally distributed and nonparametric data, respectively; Pearson and Kendall correlation tests were used to analyze associations between normally distributed data and counting data, respectively; and analyses of variance were performed by using a linear model for normally distributed data or generalized linear model with Poisson distribution for counting data. Residuals were examined for normality and independence.

In Silico Simulation of Heterosis—We wrote an R program to simulate heterosis in the framework of a nonlinear genotype–phenotype relationship (File S2). We assumed that the protein abundances were controlled by 10 factors ($i = 1, ..., 10$), the inheritance of which was either additive of nonadditive. We varied the number of polymorphic factors from 1 to 10. Each factor was defined by its value E_i (concentration or activity of factor i) and its contribution to the abundance of the controlled protein, a_i. The E_i values were drawn in a gamma distribution (mean = 6, coefficient of variation = 0.3; $\sim \Gamma(11.11, 0.54)$) and the a_i in a uniform distribution $\sim U(1,10)$. These parameters were chosen to get distributions of protein abundances similar to the distributions observed for the most abundant proteins, which are right skewed and have coefficients of variation around 0.2–0.3. For a given protein, the a_i's were the same for the parents and their hybrid. Homozygous parents were created by randomly attributing an E_i value (allelic value) to each factor. The Euclidean distance between two parents j and j' was computed as follows:

$$D_{jj'} = \sqrt{\sum_{i=1}^{10} (E_{ij} - E_{ij'})^2}$$

Protein abundances were computed assuming a concave relationship between the factors and the abundances. To this end, we used a simple hyperbolic function derived from that of the metabolic control theory (67):

$$A_j = \frac{X}{\sum_{i=1}^{10} \left(\frac{1}{a_i E_{ij}} \right)},$$

where A_j is the abundance of the protein in parent j and X is a constant.

To compute protein abundance in the hybrids, we took into account an index of inheritance $x_{ijj'}$ drawn in a normal distribution N (0.5, 0.15). If $x_{ijj'} = 0.5$, the factor was additively inherited in the hybrid between parents j and j', otherwise there was positive or negative deviation from additivity. If $x_{ijj'} = 0$ (respectively $x_{ijj'} = 1$), there was strict dominance of parent j (respectively j') over parent j' (respectively j). Therefore, the abundance of a protein in a hybrid is written:

$$A_{jj'} = \frac{X}{\sum_{i=1}^{10} \left(\frac{1}{(x_{ijj'}[n_{ijj'} - 1] + 1)a_i E_{ij}} \right)},$$

where $A_{jj'}$ is the abundance of the protein in the hybrid between parents j and j', and $n_{ijj'} = E_{ij'}/E_{ij}$.

The simulations were performed with 20,000 proteins (Figs. 8C and 8E and Fig. S6) or 400,000 proteins (Figs. 8B and 8D)

RESULTS

Protein Quantification by LC-MS/MS—A total of 396 alcoholic fermentations (66 strains × 2 temperatures × 3 replicates) were performed, of which 31 failed due to the poor fermenting abilities of some strains (Table S1). Yeast samples taken from the 365 successful fermentations were analyzed by shotgun label-free quantitative proteomics. Detailed information on all the peptides and proteins identified in all LC-MS/MS runs are shown in Table S2 and S3, respectively. Peptides were quantified by integrating precursor ion peak areas. The quantification measurements obtained for each peptide are shown in Table S4.

In total, 1,583 proteins were quantified in at least one strain × temperature combination (Table S5). Of them, 1,396 proteins were quantified both in a hybrid and its parents at the same temperature. These 1,396 proteins belonged to 16 functional categories following the MIPS Functional Catalogue Database (68) (Fig. S1, Table S6). Metabolism was the most represented category, with 534 proteins (31.1% coverage; Fig. S1).

Representation of protein abundances as a heat map showed that the strain × temperature combinations were separated in three main clusters corresponding globally to *S. uvarum* strains, interspecific hybrids, and *S. cerevisiae* strains (Fig. 1, clusters A, B, and C, respectively). Interspecific hybrids differed from all the other strains by a cluster of proteins that were globally more abundant than in the other strains (Fig. 1, cluster II). *S. uvarum* strains and *S. cerevisiae* strains differed by two clusters of proteins: one containing proteins that were more abundant in *S. cerevisiae* (Fig. 1, cluster I) and one containing proteins that were more abundant in *S. uvarum* (Fig. 1, cluster III). Except for a particular group containing the parental strain D2 and all its descendants including interspecific hybrids (Fig. 1, cluster D), the strains × temperature combinations within the clusters A, B, and C were grouped by temperature.

Protein Inheritance Patterns—To analyze the inheritance of protein abundances at a given temperature, we considered the triplets (formed by one hybrid and its parents) where at least two successful fermentations were obtained for each member. This was the case for 53 triplets at 18 °C and for 44 triplets at 26 °C (Table S1). For each protein × hybrid × temperature combination, we computed the deviation from additivity (*d*) as the difference between hybrid abundance and mid-parental abundance. A protein was considered as heterotic whenever *d* was significantly different from zero (Wald test, adjusted $p < .05$, Table S7). A total of 97,360 protein × hybrid × temperature combinations were examined. For 65.2% (63,469) of them, no significant abundance variation was detected neither between a hybrid and its parent nor between parents (invariant proteins). The remaining 33,891 protein × hybrid × temperature combinations were classified depending on their inheritance pattern (Tables S7 and S8, Fig. 2): 66.8% (22,634) displayed additivity; 11.7% (3,965) displayed negative or positive MPH, meaning that the protein abundance in the hybrid was within the parental range; 11.0% (3,746) displayed BPH or WPH, meaning that the protein abundance in the hybrid fell outside the parental range; and 10.5% (3,546) corresponded to cases of unresolved heterosis
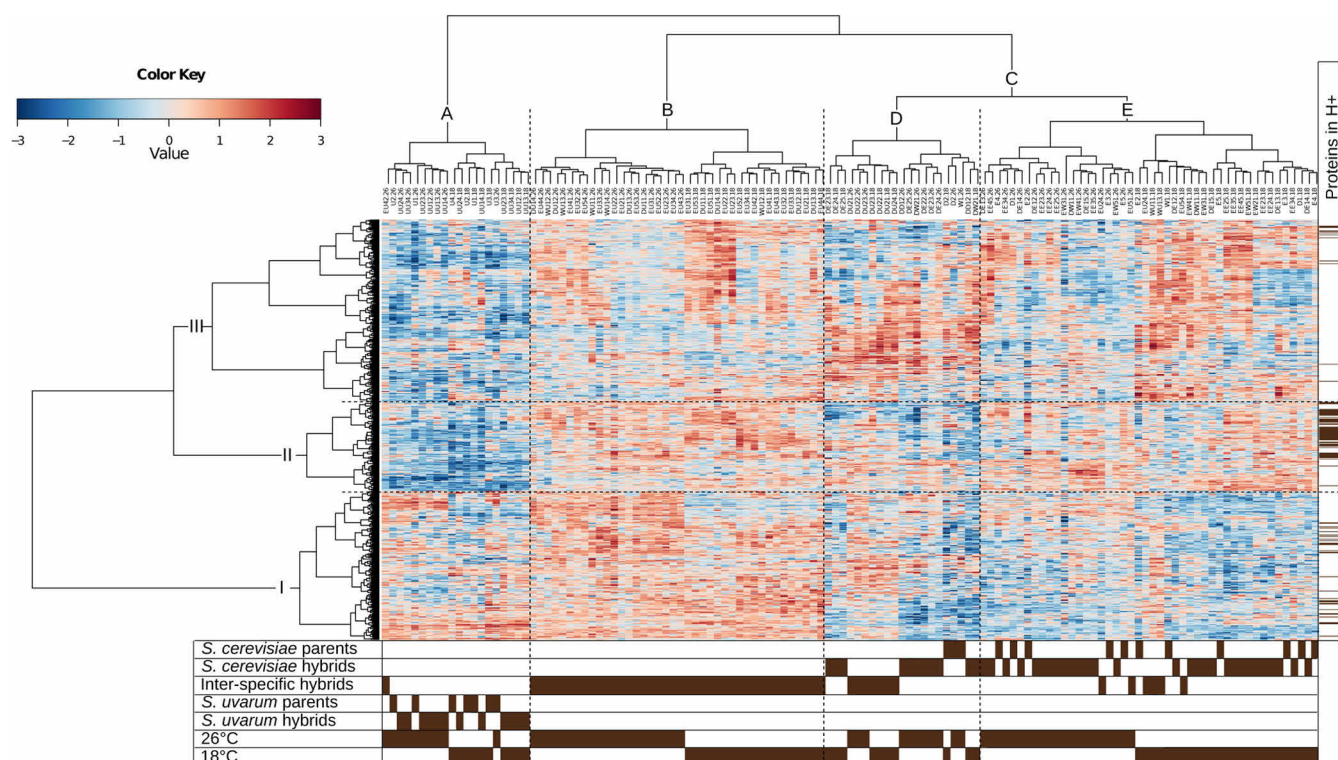
FIG. 1. **Heat map representation of the estimated protein abundances.** Each line corresponds to a protein and each column to a strain × temperature combination. A total of 615 proteins that were quantified in at least 122 strains × temperature combinations are presented (see experimental procedures for details). For each protein, abundance values were scaled and represented by a color code as indicated by the color-key bar: blue for low abundances and red for high abundances. Hierarchical clusterings of the strains (*top*) and of the proteins (*left*) were built by using Euclidean distances and Ward aggregation method. Letters on the top indicate clusters of strain × temperature combinations presenting similar proteomes. Roman numerals on the left indicate clusters of proteins exhibiting similar abundance patterns. Membership of a protein to the set H+ (see results and Fig. 6) is shown in brown on the right. The type of strain and the growth temperature is indicated in brown at the bottom.
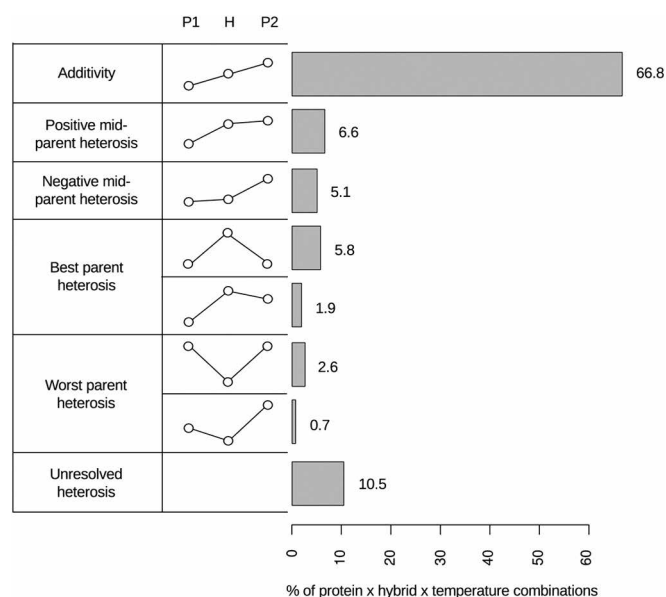

FIG. 2. **Inheritance pattern of the proteins exhibiting abundance variation between hybrid and parental strains.** P1, parent 1; H, hybrid; P2, parent 2.

because statistical tests did not allow us to distinguish between mid-parent and best/worst-parent heterosis.

The proportion of heterotic proteins per hybrid × temperature combination (invariant proteins omitted) was highly variable, ranging from 8.4 to 61.2% with a median at 31.4% (Table II). Globally, hybrids having at least one *S. cerevisiae* strain as a parent showed more heterotic proteins at 18 °C than at 26 °C (Fig. 3). On the contrary, *S. uvarum* intraspecific hybrids showed slightly more heterotic proteins at 26 °C than at 18 °C (Fig. 3).

Interspecific Hybrids Exhibit Specific Characteristics Regarding Protein Abundance Inheritance—We further analyzed heterosis for protein abundance in inter- *versus* intraspecific hybrids. By examining the distribution of relative additivity deviation (computed as d/m, where m is the parental mean), we showed that d/m was globally higher in inter- than in intraspecific hybrids (Fig. 4*A*, Fig. S2). In addition, the proportion of heterotic proteins with positive d values was, on average, much higher in inter- than intraspecific hybrids (78.8%, 52.3%, and 42.6% in interspecific, *S. cerevisiae* and *S. uvarum* hybrids, respectively; Fig. 4*B*). This indicates a strong bias toward positive heterosis in interspecific hybrids.

TABLE II
Counting of quantified proteins, invariant proteins and heterotic proteins in each hybrid x temperature combination at 18 °C, 26 °C, and at both temperatures

Hybrid	18 °C			26 °C			Both 18 and 26 °C		
	Total nb of proteins	% of invariant proteins (a)	% of heterotic proteins (b)	Total nb of proteins	% of invariant proteins (a)	% of heterotic proteins (b)	Total nb of proteins	% of invariant proteins (a)	% of heterotic proteins (b)
DD12	1,189	64.5	29.6	1,210	62.8	41.1	1,174	48.5	10.1
DE12	1,217	68.4	39.5	1,178	65.9	30.6	1,164	51.6	11.4
DE13	1,212	85.2	23.5	1,184	80.8	33.9	1,174	72.1	0.9
DE14	1,222	81.8	17.9	1,210	77.0	37.8	1,202	68.3	2.1
DE15	1,190	73.6	25.8	1,221	72.3	37.0	1,182	59.8	7.8
DE22	NA	NA	NA	1,127	65.7	30.2	NA	NA	NA
DE23	1,178	56.9	51.4	1,208	70.1	30.5	1,163	47.7	9.5
DE24	1,187	63.5	48.0	1,216	62.7	17.0	1,177	47.6	5.7
DE25	1,187	60.7	37.9	1,209	58.6	43.1	1,170	43.1	10.4
DW11	1,214	69.9	20.2	1,216	71.2	36.9	1,198	57.0	7.2
DW21	1,175	64.2	31.4	1,212	61.5	46.3	1,162	47.3	12.4
EE23	1,203	71.0	39.8	1,178	69.9	29.1	1,156	54.2	4.2
EE24	1,211	69.9	46.8	1,168	65.8	20.1	1,153	51.4	2.9
EE25	1,205	59.6	61.2	1,176	70.3	25.8	1,157	47.2	7.4
EE34	1,207	83.8	41.5	1,213	77.4	25.9	1,193	70.7	6.9
EE35	1179	73.8	46.3	1,215	80.1	38.0	1,169	63.8	8.7
EE45	1,223	76.6	36.7	1,219	68.1	33.7	1,209	59.5	7.8
EW21	1,200	61.7	55.0	NA	NA	NA	NA	NA	NA
EW31	1,197	70.7	23.1	891	67.0	33.3	884	54.2	5.2
EW41	1,211	67.2	28.5	1,214	60.1	25.0	1,194	47.1	6.3
EW51	1,215	69.4	44.9	1,217	66.0	21.7	1,200	50.8	4.7
DU11	860	60.3	35.5	870	56.4	24.0	841	42.7	6.2
DU12	883	58.9	25.9	869	59.4	34.3	844	42.5	9.5
DU13	886	52.8	38.0	878	58.4	20.5	857	41.1	7.1
DU14	820	58.8	49.7	NA	NA	NA	NA	NA	NA
DU21	813	59.0	30.9	837	59.1	21.3	790	43.4	5.1
DU22	821	56.3	40.4	851	54.2	39.7	795	39.0	11.1
DU23	804	55.0	24.0	854	55.9	23.3	794	39.7	6.1
DU24	779	61.6	34.8	NA	NA	NA	NA	NA	NA
EU21	833	61.3	32.9	806	57.7	15.5	778	42.4	4.5
EU22	827	54.5	39.9	807	58.9	27.1	772	38.0	3.8
EU23	836	52.3	46.4	813	59.7	20.7	778	39.2	5.9
EU24	781	71.3	19.6	NA	NA	NA	NA	NA	NA
EU31	841	61.1	37.3	865	63.7	31.2	822	49.0	8.8
EU32	833	61.1	27.8	874	66.2	44.7	812	49.9	9.3
EU33	841	64.4	21.7	856	62.1	42.9	812	48.8	7.0
EU34	795	64.3	25.4	NA	NA	NA	NA	NA	NA
EU41	877	58.3	37.4	820	57.3	16.9	811	41.3	5.0
EU42	878	56.0	44.0	834	65.9	24.6	822	45.3	8.9
EU43	880	56.5	32.9	868	56.0	22.8	845	41.4	6.9
EU44	822	57.5	35.5	NA	NA	NA	NA	NA	NA
EU51	870	60.3	46.7	843	66.0	8.4	826	48.9	0.9
EU52	871	53.5	44.2	874	64.6	21.7	844	43.1	5.4
EU53	871	50.4	36.1	869	60.8	25.2	844	39.1	6.4
EU54	816	65.0	36.4	NA	NA	NA	NA	NA	NA
WU11	804	63.8	32.3	803	58.5	28.5	783	44.6	4.8
WU12	803	60.6	33.9	803	59.9	28.9	780	43.3	6.1
WU13	809	56.4	39.9	800	54.5	28.3	782	42.7	13.2
UU12	1,050	77.1	23.3	1,038	67.3	27.1	1,031	59.6	3.8
UU13	1,052	72.3	27.1	1,044	66.1	39.5	1,036	56.9	8.9
UU14	1,047	72.0	49.8	NA	NA	NA	NA	NA	NA
UU23	1,038	62.3	30.7	1,036	72.2	30.6	1,023	51.9	4.7
UU24	1,053	76.3	28.4	NA	NA	NA	NA	NA	NA
UU34	1,050	66.6	24.2	NA	NA	NA	NA	NA	NA

(a) proteins whose abundance did not vary neither between a hybrid and its parent nor between parents.
(b) invariant proteins omitted.

We next looked whether the temperature affected protein inheritance similarly in interspecific hybrids compared with intraspecific hybrids. For the majority of the protein × hybrid combinations (82.5%), the protein was heterotic at only one temperature. This indicates that heterosis for protein abundance is generally dependent on the temperature. For the remaining 17.5%, four scenarios were possible depending on the sign of d: positive at both 18 °C and 26 °C ($+/+$), negative at both 18 °C and 26 °C ($-/-$), positive at 18 °C and negative at 26 °C ($\pm$), negative at 18 °C and positive at 26 °C ($-/+$).

FIG. 3. **Relationships between the proportion of heterotic proteins, the parental strains and the temperature.** (*A*) Distribution of the proportion of heterotic proteins according to parental strains and temperature. (*B*) Distributions of the proportion of heterotic proteins among *S. cerevisiae* hybrids, interspecific hybrids and *S. uvarum* hybrids at the two temperatures.
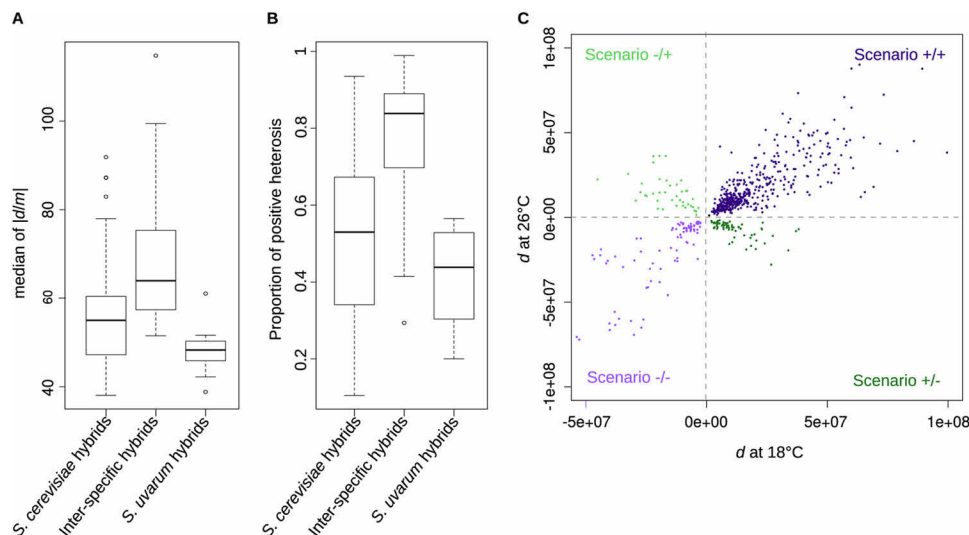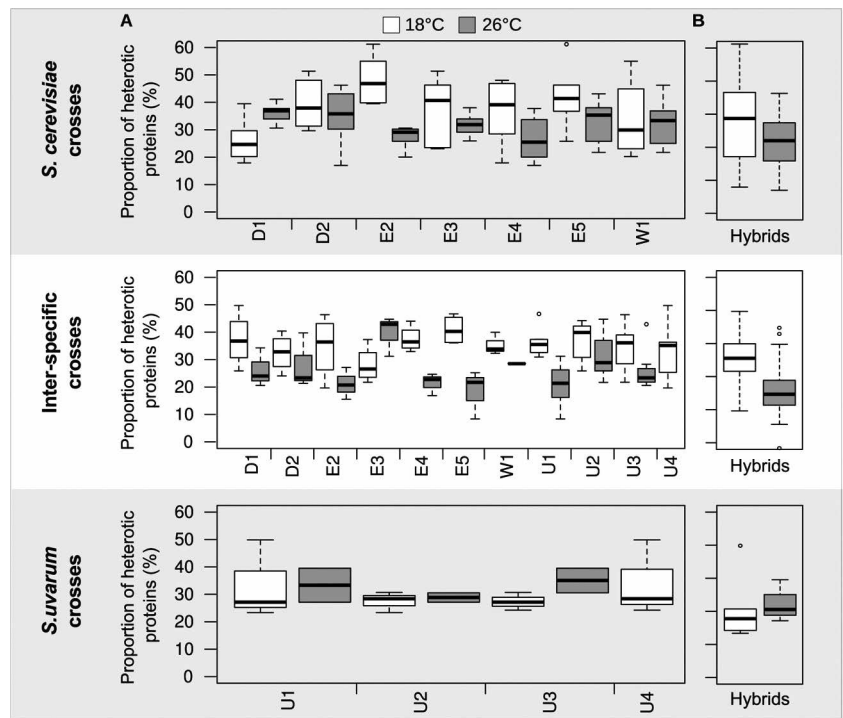

FIG. 4. **Comparing additivity deviations in inter- and intra-specific hybrids.** (*A*) Distributions of the medians of absolute values of relative additivity deviation (*d/m*) computed from all the proteins analyzed in the hybrid × temperature combinations. (*B*) Distribution of the proportions of proteins showing positive heterosis in the hybrid ×x temperature combinations. (*C*) Relationships between additivity deviation (*d*) at 18 °C and at 26 °C for the proteins exhibiting heterosis at the two temperatures in interspecific hybrids. The same representation for intra-specific hybrids is shown in Fig. S4.

Globally, interspecific hybrids presented an excess of +/+ scenarios (451 over 656, χ^2 test, $p = 7.6 \times 10^{-22}$; Fig. 4*C*). This result holds true for nearly all interspecific hybrids (Fig. S3*A*). Regarding intraspecific hybrids, *S. cerevisiae* hybrids presented an excess of $\pm$ scenarios (252 over 719, χ^2 test, $p = 1.1 \times 10^{-15}$; Fig. S4*A*), while *S. uvarum* hybrids lacked $\pm$ scenarios (6 over 79, χ^2 test, $p = 4.8 \times 10^{-14}$; Fig. S4*B*). However, this result largely depended on the hybrid considered (Fig. S3*B*).

The Remodeling of the Proteome of Interspecific Hybrids Predominantly Affects Particular Categories of Proteins—Principal component analysis based on the estimated protein abundances was performed in order to visualize the effects of the strains and of the temperature on the proteome (Fig. 5). The first axis (PC1, 15% of the total variance) separated the parental and hybrid strains of *S. cerevisiae* from those of *S. uvarum,* with interspecific hybrids located between the two species. Interestingly, within each type of hybrid (*S. cerevi-*
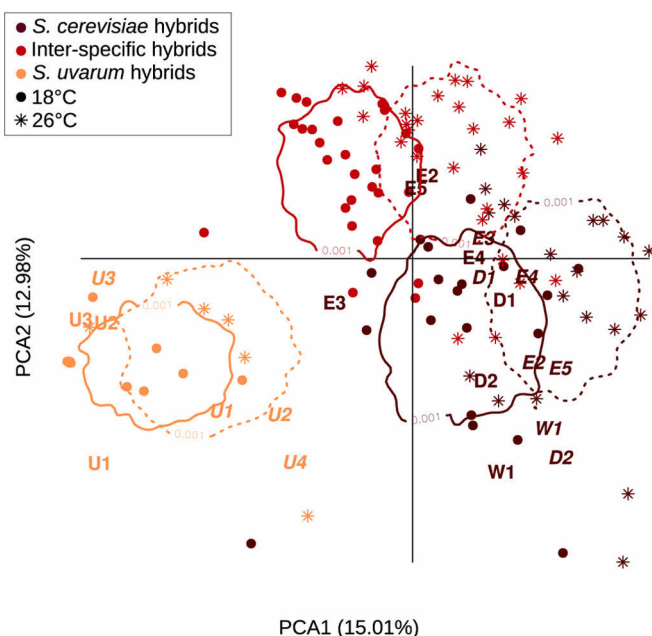
FIG. 5. **Principal component analysis based on the estimated abundances of 615 proteins quantified in at least 122 strains x temperature combinations (see experimental procedures for details).** Parental strains are written in upright (18 °C) or italics (26 °C) characters. Plain and dotted lines represent the limits that contain 99.9% of the distribution of the PC1 and PC2 coordinates of strain × temperature combinations for each group. They were obtained by simulating the kernel densities from group's means and variances assuming bivariate normal distributions and using the R package MASS.

FIG. 6. **Characteristics of the proteins in the set H+.** The proteins correlated with $r > 0.5$ to the second axis of the PCA shown Fig. 5 (Pearson correlation test, set H+) were compared with those that were not correlated (set NH) for: (A) the mean abundance in parental strains (Student test); (B) the proportion of proteins encoded by essential genes (χ^2 test); (C) the proportion of proteins involved in protein metabolism (χ^2 test). Symbols: . $6.10^{-2} > p \geq 5.10^{-2}$; * $5.10^{-2} > p \geq 5.10^{-3}$; ** $5.10^{-3} > p \geq 5.10^{-4}$; *** $5.10^{-4} > P$.

siae, *S. uvarum,* and interspecific), PC1 also separated hybrid × temperature combinations according to the temperature. Globally, the effect of temperature on the proteome was similar for all the genotypes: *S. uvarum*, *S. cerevisiae,* and interspecific strains grown at 26 °C were shifted to the right of PC1. Consequently, *S. uvarum* strains moved along PC1 toward *S. cerevisiae* when temperature changed from 18 °C to 26 °C, and reciprocally, *S. cerevisiae* strains moved along PC1 toward *S. uvarum* when temperature changed from 26 °C to 18 °C. This result shows that, when a species is grown at nonoptimal temperature, its proteome tends to resemble that of the other species for which the temperature is optimal.

The second axis (PC2, 13% of the total variance) separated interspecific hybrids from the other strains. PC2 contributed nearly as much as PC1 to the total variance, indicating that interspecific hybridization has extensively remodeled the proteome. To characterize the proteins involved in the differentiation of interspecific hybrids, we analyzed the proteins significantly correlated to PC2 (Pearson correlation test, adjusted $p < 0.01$) with $r > 0.5$ (set H, 104 proteins; Table S9). For all of them but one, r was positive, which indicates that these proteins contributed positively to a greater abundance in interspecific hybrids, regarding the part of variation represented by PC2. This is in agreement with Fig. 1, showing that the
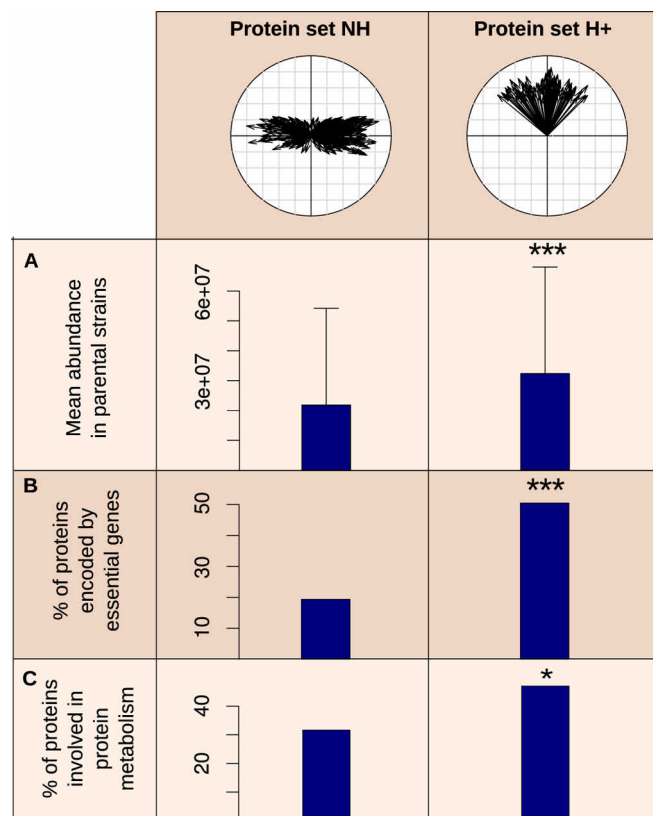
majority of the proteins in the set H+ (*i.e.* the set H without the protein negatively correlated to PC2) were included in cluster II. These proteins contributed poorly to PC1, which is consistent with the fact that they displayed little abundance variation between the parents of interspecific hybrids and between temperatures (Fig. S5). Compared with the proteins that were not correlated to PC2 (set NH, 253 proteins; Table S9), these proteins exhibited other specific characteristics: They were more abundant than other proteins (average abundance in parental strains: 2.2×10^7 in NH *versus* 3.1×10^7 in H+; Student test, $p = 2.5 \times 10^{-64}$, Fig. 6A); they were significantly enriched in proteins encoded by essential genes, *i.e.* genes that are required for viability of *S. cerevisiae* under standard laboratory conditions (69, 70) (22,6% in NH *versus* 55.3% in H+; χ^2 test, $p = 1.6 \times 10^{-7}$, Fig. 6B); and they were slightly enriched in proteins involved in protein metabolism (protein synthesis and protein fate; 32.0% in NH *versus* 49.0% in H+; χ^2 test, adjusted $p = .029$; Fig. 6C).

Altogether, these results show that interspecific hybridization caused BPH for a defined portion of the proteome that
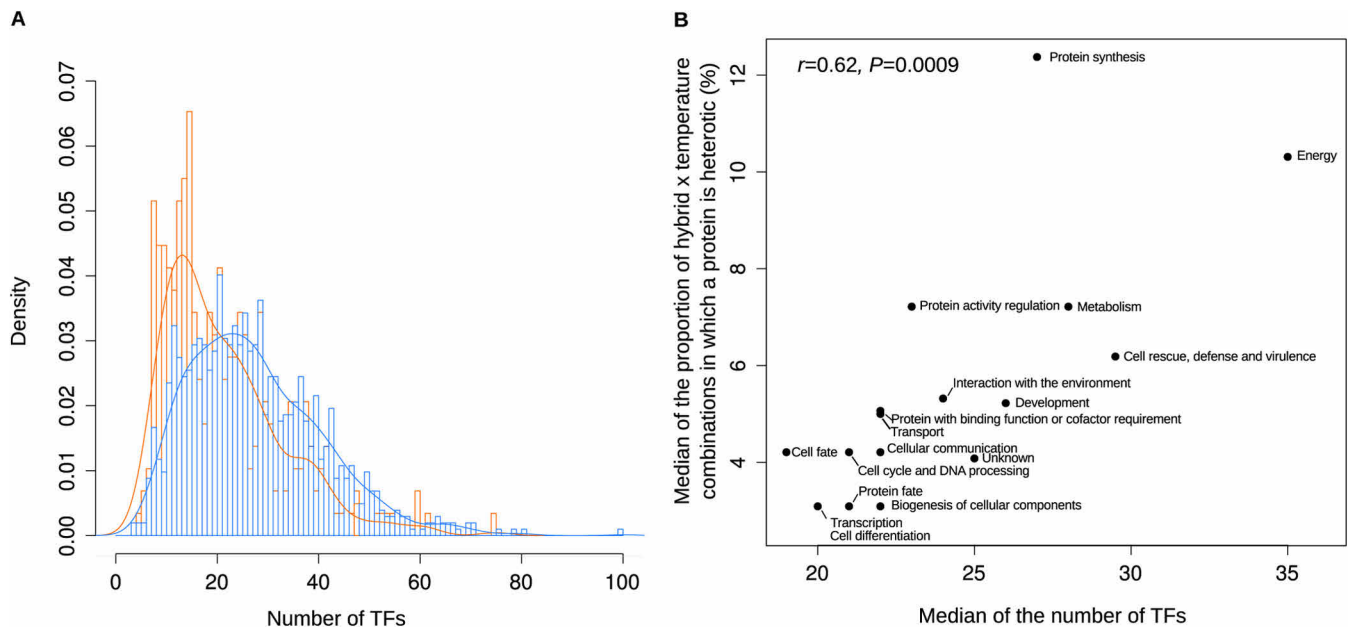
FIG. 7. **Relationships between the frequency at which a protein is heterotic and the number of putative transcription factors (TFs) of a gene.** (*A*) Distributions of the number of putative TFs of a gene for proteins that were heterotic in at least one hybrid × temperature combination (blue) and proteins never observed as heterotic (orange). (*B*) Relationship between the proportion of strain × temperature combinations in which proteins were heterotic and the number of putative TFs of the encoding genes for data organized by functional category.

contains proteins characterized by the stability of their abundances toward genetic and environmental changes, by their high abundances, and by their importance for the cell viability.

Heterosis for Protein Abundance Is Partly Related to the Complexity of Transcriptional Regulation—To determine the extent to which the factors controlling protein abundances could be involved in heterosis for protein abundance, we focused on the transcription factors (TFs) possibly involved in the regulation of the genes encoding the proteins quantified in our study. A total of 162 TFs with a consensus DNA-binding sequence were retrieved from the Yeastract database (www.yeastract.com; 71–74). On average, the genes encoding proteins that were heterotic in at least one hybrid × temperature combination were putative targets of a higher number of TFs than the genes encoding non-heterotic proteins (27.7 *versus* 21.4; Mann–Whitney test $p = 1.96 \times 10^{-15}$; Fig. 7*A*). In addition, a significant correlation was found between the number of putative TFs of a gene and the proportion of hybrids × temperature combinations in which the encoded protein was heterotic (Kendall correlation test, $r = 0.18$, $p < 2.2 \times 10^{-16}$, Fig. S6).

The number of putative TFs of a gene depended significantly on the functional category of the gene (generalized linear model, ANOVA $p < 2.2 \times 10^{-16}$). As a consequence, the frequency at which a protein was heterotic was also dependent on its functional category. For example, the genes involved in metabolism, energy and cell rescue, defense, and virulence had, on average, more putative TFs, and their proteins were more frequently heterotic than those involved in cell differentiation (Fig. 7*B*). Note that the protein synthesis

category appeared as an outlier, containing proteins that were heterotic in a high proportion of hybrids but not presenting a very high number of putative TFs.

Altogether, these results suggest that the number of factors involved in transcriptional regulation may have an influence on heterosis for protein abundance, which may also explain why some functional categories are more prone to heterosis than others.

Predicting Protein Inheritance According to a Nonlinear Model—A general property of metabolic systems is the nonlinear response of the fluxes to genetic variations of enzyme concentrations and/or activity parameters (75). This relationship allowed Kacser and Burns (76) to propose a metabolic basis for dominance. In addition, Fiévet *et al.* (7) showed that when two or more enzymes are variable, the concave relationship between a flux and its parameters necessarily results in positive MPH or in BPH for the flux when the enzyme parameters, *i.e.* any genetic parameter that determines the enzymatic activity, are additively inherited (Fig. 8*A*).

Interestingly, the protein synthesis rates seem also to be a concave function of various factors, such as mRNA amount (77), translation factor abundance (78), ribosomal initiation rate, and elongation rate (79). Therefore, the basis of heterosis put forward for metabolic fluxes could apply for protein abundances, even though the relationship is mathematically different. In order to test this hypothesis and interpret our results, we used a simple nonlinear function for modeling and simulating the consequences of concavity on protein heterosis.
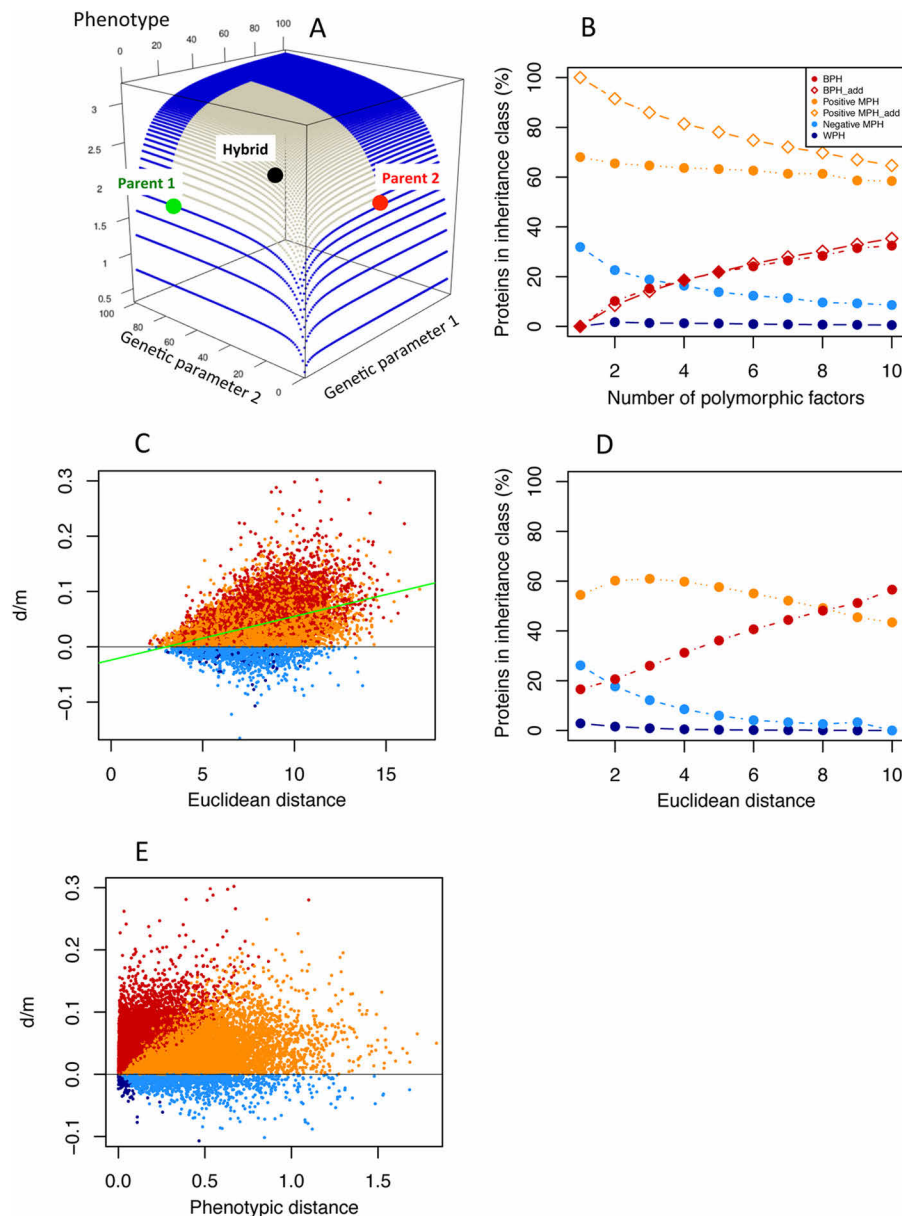
FIG. 8. *In silico* simulation of heterosis. (*A*) Phenotypic response with respect to the variation of two genetic parameters in the framework of a concave genotype-phenotype relationship. Parent 1 (green filled circle) and parent 2 (red filled circle) have close phenotypic values. In case of additivity of the values of both genetic parameters, there is BPH (black filled circle). Otherwise the hybrid value can occupy any point of the gray sub-surface, depending on the values of the inheritance parameters (see experimental procedures). (*B*) Effect of the number of polymorphic factors on the percentages of proteins in the different classes of inheritance. Empty diamonds: additivity of the values of the factors. Filled circles: general case, with inheritance parameters of the factors drawn in normal distributions. In case of additivity of the factors, only BPH and positive MPH can be observed. If there is only one polymorphic factor, BPH and WPH cannot be observed, whatever the inheritance of the factors. (*C*) Relationship between d/m and Euclidean distance between parents, for 10 polymorphic factors ($R^2 = 0.20$). (*D*) Relationship between the Euclidean distance between parents and the percentages of proteins in the different classes of inheritance for 10 polymorphic factors. (*E*) Relationship between the phenotypic distance between parents and d/m, for 10 polymorphic factors.

We analyzed the relationships between heterosis for protein abundance and (i) the type of inheritance of the genetic factors controlling abundance; (ii) the number of polymorphic factors controlling a protein, which is an indicator of the complexity of its genetic control; (iii) the Euclidean distance between parents computed from the values of the factors

controlling protein abundances; and (iv) the phenotypic distance, *i.e.* the difference in protein abundance between parents.

When there is only one polymorphic factor, only positive MPH can be observed if there is additivity of the factor and positive and negative MPH in case of nonadditivity (Fig. 8*B*).

When two or more factors are polymorphic, positive MPH *and* BPH are possible if there is additivity of the factor, and if there is nonadditivity, the four types of heterosis are possible. When the number of polymorphic factors increases, BPH proportion increases at the expense of the other types of heterosis (Fig. 8*B*).

The Euclidean distance between parents was positively correlated with *d/m* (Fig. 8*C*), which is consistent with the well-known relationship between genetic distance and heterosis. More interestingly, the proportions of the different types of heterosis depended on the distance. For the smallest distances, all the types of heterosis were observed and all *d/m* values were small, while for the largest distances almost exclusively BPH and positive MPH were observed and their *d/m* values were high (Figs. 8*C* and 8*D*). This observation was valid whatever the number of polymorphic factors (Fig. S7). As expected from the concavity of the function, the distribution of the four types of heterosis tightly depended also on the phenotypic distance between parents. For the closest parents, there was a majority of BPH cases and few WPH, while the hybrids between distant parents displayed mainly positive MPH and to a lesser extent negative MPH (Figs. 8*A* and 8*E* and Fig. S7).

DISCUSSION

We used label-free quantitative proteomics in yeast to perform a large-scale study of heterosis for protein abundance. In agreement with previous results (58), we confirmed that the proteomes of *S. cerevisiae* and *S. uvarum* were highly differentiated. Interestingly, this differentiation is partly related to the adaptation of species to their optimal temperatures (18 °C for *S. uvarum* and 26 °C for *S. cerevisiae*), as evidenced by the fact that lower temperatures drive *S. cerevisiae*'s proteome close to that of *S. uvarum*, while the higher temperature drives *S. uvarum*'s proteome close to that of *S. cerevisiae*.

*Heterosis for Protein Abundance Is Subject to Genotype ×
Environment Interactions*—Heterotic proteins were detected in every hybrid × temperature combinations analyzed. This is in line with previous results showing that heterosis for gene expression and protein abundance is a common occurrence, regardless the species or genotypes considered (reviewed in (11)). The proportion of heterotic proteins varied from 8.4 to 61.2% depending on the hybrid × temperature combination considered. Comparatively, Khan *et al.* (80) found 85.9% of heterotic proteins (342 out of 398) in one *S. cerevisiae* × *S. uvarum* cross. However, these authors used an arbitrary threshold without statistical test to decide on the inheritance of the proteins, which may explain the discrepancy with our results. In any case, our study is much more representative of both the proteome and the genetic diversity of *S. cerevisiae* and *S. uvarum* since we examined 1,396 proteins quantified in 55 crosses and at two temperatures. This allowed us to show that there were genotype × environment interactions for heterosis since the temperature did not affect protein inheritance

similarly in the different types of hybrid examined. Indeed, the proportion of heterotic proteins was higher at 18 °C for *S. cerevisiae* and interspecific hybrids and at 26 °C for *S. uvarum* hybrids. Note that in the case of intraspecific hybrids, these temperatures were nonoptimal, suggesting that there may be a relationship between the proportion of heterotic proteins and stressful growth conditions. In addition, the sign of *d* was little affected by temperature in interspecific hybrids, which was not the case in intraspecific hybrids.

Heterosis for Protein Abundance Primarily Affects Highly Regulated Proteins—Our results suggest that the number of putative TFs of a gene is related to the heterosis for the abundance of the encoded protein. Regulation of transcription is complex, involving a combination of several TFs individually acting as activator and/or repressor (81). Previous studies have shown that genetic polymorphism in *cis* and *trans* regulators can influence the inheritance pattern of gene expression level, polymorphism of trans regulators being preferentially associated to heterotic patterns (39, 82–84). If the number of polymorphic TFs increases with the number of TFs, the relationship between the number of putative TFs of a gene and the frequency at which the encoded protein is heterotic is consistent. Conceptually, our results are similar to what has been observed for agronomic traits in plants. Indeed the results obtained from previous studies show that highly complex, polygenic traits such as yield are more prone to heterosis (23, 31). By analyzing a very high number of traits, we show here that the relationship between genetic complexity and heterosis is robust, even for less-integrated traits such as protein abundance.

The number of putative TFs of a gene depended significantly on the functional category of the gene, explaining why the proteins from some functional categories showed more heterosis than others. Among the functional categories containing genes putatively regulated by a high number of TFs and showing frequently heterotic proteins, we found energy, metabolism, and cell rescue, defense, and virulence. This result is consistent with many studies in plants that showed that these categories were involved in heterosis for gene expression (reviewed in (22, 24)). In addition, since these categories are generally involved in response to environmental changes (85), they were expected to be highly regulated.

The proteins involved in protein synthesis appeared as outliers regarding the relationship between the number of putative TFs of a gene and heterosis for protein abundance, presenting frequencies of heterosis higher than expected based on the number of putative TFs of their encoding genes. To explain the peculiar behavior of these proteins, we assume that factors other than TFs are involved in heterosis for protein abundance as, for example, posttranslational modifications, that were recently shown to be related to the variations of phenotypic traits (86).

Best-Parent Heterosis Is Related to Proteins Under Evolutionary Constraints—BPH in interspecific hybrids was more

particularly related to a particular group of proteins (set H+) that were highly abundant and exhibited little abundance variations between temperatures and between *S. cerevisiae* and *S. uvarum*, yet two distantly related species (87). Observation of interspecific heterosis for these proteins necessarily implies that the two species are genetically contrasted at the loci controlling protein abundances (Protein Quantity Loci (33)). This is in agreement with dominance hypothesis, which attributes heterosis to allele complementation. In addition to these results, we also showed that the set H+ was enriched in proteins encoded by essential genes and in proteins involved in protein metabolism. Essential genes are thought to be under strong purifying selection since they are highly conserved across large evolutionary distances in yeasts and mammals (88, 89). Moreover, protein metabolism includes proteins of ribosomes and proteasome that are structurally and functionally conserved (90, 91). This suggests that the proteins of the set H+ are under evolutionary constraint. However, we have currently no hypothesis to establish a relationship between evolutionary constraint and heterosis.

Heterosis for Protein Abundance Is Consistent with a Model of Nonlinear Genotype–Phenotype Relationship—It has been observed from numerous experiments that heterosis is generally biased toward positive values (for example, (3, 9)). This bias is accounted for in the dominance hypothesis, where recessive deleterious alleles are complemented by dominant superior alleles (26, 27). In the context of metabolic systems, dominance of the high over the low allele is explained by the hyperbolic response of fluxes toward the variations of enzyme parameters (*e.g.* activity, concentration): Due to the concavity of the curve, the flux value in a hybrid obtained from a cross between two parents presenting contrasted enzyme parameters is systematically biased toward the highest parent, provided the value of the enzyme parameter is additively inherited (76, 92). Generalized to networks with several variable enzymes, this hyperbolic relationship generates heterosis for the metabolic flux (7, 93).

In this study, we analyzed a high number of traits in a large number of hybrids, which allowed us to examine the extent to which the bias toward positive heterosis was robust. Unexpectedly, we showed that heterosis for protein abundance was strongly biased toward positive values in interspecific hybrids but not in interspecific hybrids, where positive and negative heterosis were relatively well balanced. This result was difficult to explain from the current knowledge on heterosis, since, as far as we know, there is no model for negative heterosis. To interpret this result, we relied on previous observations showing that (i) concave genotype–phenotype relationships exist at various levels of cell organization (76, 94–97) and in particular for the protein synthesis rate (77–79) and (ii) nonadditivity can occurs at every level of cell organization, from transcript abundance to more integrated traits (24, 25).

By simulating heterosis for protein abundance using a nonlinear model of genotype–phenotype relationship, we obtained *in silico* results in agreement with those obtained from the experiments. First, we showed that negative heterosis can occur when there is nonadditive inheritance of the genetic factors, which is biologically realistic. Second, we showed that for small genetic distances positive and negative heterosis can be observed, while for large distances there is much more positive than negative heterosis. This is consistent with the bias we observed between intra- and interspecific hybrids. Third, we showed that the proportion of BPH was maximal for hybrids obtained from distant parents *and* for proteins displaying similar abundances in the parents. This is consistent with the frequent BPH observed in interspecific hybrids for the proteins of set H+. Finally, we showed that heterosis was related to the number of polymorphic factors controlling a protein. This is consistent with the observation that the proteins regulated by a high number of TFs were more prone to heterosis.

To conclude, we performed a large-scale study of heterosis, which allowed us to obtain original results: (i) heterosis was strongly biased toward positive values in interspecific hybrids but not in intraspecific hybrids and (ii) BPH in interspecific hybrids occurred preferentially for a special group of proteins assumed to be under evolutionary constraint. These results shed new light on heterosis by supporting a model where protein abundances would be related to transcriptional and translational parameters by concave relationships. In agreement with this hypothesis, we also showed that the complexity of transcriptional regulation, estimated through the number of putative TFs of a gene, is related to heterosis for protein abundance, which supports a general relationship between heterosis and trait complexity. Taken together, our results show the interest of high-throughput technologies to provide a more comprehensive view of complex biological phenomena such as heterosis.

K To whom correspondence should be addressed: PAPPSO, Plateforme d'Analyse Protéomique de Paris Sud-Ouest, UMR de Génétique Végétale, INRA-Univ Paris Sud-CNRS, Ferme du Moulon, F-91190 Gif-sur-Yvette, France. Fax: +33 1 69 33 23 40; E-mail: zivy@moulon.inra.fr.

REFERENCES

1. Mendel, G. (1866) Versuche über Pflanzenhybriden. *Verh Naturforsch Ver Brünn* **4,** 3–47
2. Hull, F. H. (1946) Overdominance and corn breeding where hybrid seed is not feasible. *Agronom. J.* **38,** 1100–1103
3. Lisec, J., Römisch-Margl, L., Nikoloski, Z., Piepho, H., Giavalisco, P., Selbig, J., Gierl, A., and Willmitzer, L. (2011) Corn hybrids display lower metabolite variability and complex metabolite inheritance patterns. *Plant J.* **68,** 326–336
4. Salinas, F., Cubillos, F. A., Soto, D., Garcia, V., Bergström, A., Warringer, J., Ganga, M. A., Louis, E. J., Liti, G., and Martinez, C. (2012) The

genetic basis of natural variation in oenological traits in *Saccharomyces cerevisiae*. *PLoS ONE* **7**, e49640

5. Warner, R. L., Hageman, R. H., Dudley, J. W., and Lambert, R. J. (1969) Inheritance of nitrate reductase activity in *Zea mays* L. *Proc. Natl. Acad. Sci. U.S.A.* **62**, 785–792

6. Causse, M., Rocher, J.-P., Pelleschi, S., Barrière, Y., de Vienne, D., and Prioul, J.-L. (1995) Sucrose phosphate synthase: An enzyme with heterotic activity correlated with maize growth. *Crop Sci.* **35**, 995–1001

7. Fiévet, J. B., Dillmann, C., and de Vienne, D. (2010) Systemic properties of metabolic networks lead to an epistasis-based model for heterosis. *Theor. Appl. Genet.* **120**, 463–473

8. Mohayeji, M., Capriotti, A. L., Cavaliere, C., Piovesana, S., Samperi, R., Stampachiacchiere, S., Toorchi, M., and Lagana, A. (2014) Heterosis profile of sunflower leaves: A label free proteomics approach. *J. Proteomics* **99**, 101–110

9. Zörgö, E., Gjuvsland, A., Cubillos, F. A., Louis, E. J., Liti, G., Blomberg, A., Omholt, S. W., and Warringer, J. (2012) Life history shapes trait heredity by promoting accumulation of loss-of-function alleles in yeast. *Mol. Biol. Evol.* **29**, 1781–1789

10. Rieseberg, L. H., Baird, S. J., and Gardner, K. A. (2000) Hybridization, introgression, and linkage evolution. *Plant Mol. Biol.* **42**, 205–224

11. Lippman, Z. B., and Zamir, D. (2007) Heterosis: Revisiting the magic. *Trends Genet.* **23**, 60–66

12. Crow, J. F. (1998) 90 years ago: The beginning of hybrid maize. *Genetics* **148**, 923–928

13. Duvick, D. N. (1999) Heterosis: Feeding People and Protecting Natural Resources. In *The Genetics and Exploitation of Heterosis in Crops,* pp. 19–29, American Society of Agronomy, Crop Science Society of America, Madison, WI

14. Charlesworth, D., and Willis, J. H. (2009) The genetics of inbreeding depression. *Nat. Rev. Genet.* **10**, 783–96

15. Sprague, G. F. (1983) Heterosis in maize: Theory and practice. In *Heterosis*, pp. 47–70. Springer, Berlin Heidelberg

16. Tsaftaris, S. A. (1995) Molecular aspects of heterosis in plants. *Physiol. Plant.* **94**, 362–370

17. Tsaftaris, A. S., and Kafka, M. (1997) Mechanisms of heterosis in crop plants. *J. Crop Prod.* **1**, 95–111

18. Birchler, J. A., Auger, D. L., and Riddle, N. C. (2003) In search of the molecular basis of heterosis. *Plant Cell* **15**, 2236–2239

19. Hochholdinger, F., and Hoecker, N. (2007) Towards the molecular basis of heterosis. *Trends Plant Sci.* **12**, 427–432

20. Birchler, J. A., Yao, H., Chudalayandi, S., Vaiman, D., and Veitia, R. A. (2010) Heterosis. *Plant Cell* **22**, 2105–2112

21. Chen, Z. J. (2010) Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci.* **15**, 57–71

22. Baranwal, V. K., Mikkilineni, V., Zehr, U. B., Tyagi, A. K., and Kapoor, S. (2012) Heterosis: Emerging ideas about hybrid vigour. *J. Exp. Bot.* **63**, 6309–6314

23. Kaeppler, S. (2012) Heterosis: Many genes, many mechanisms—End the search for an undiscovered unifying theory. *ISRN Bot.* **2012**, 1–12

24. Chen, Z. J. (2013) Genomic and epigenetic insights into the molecular bases of heterosis. *Nat. Rev. Genet.* **14**, 471–482

25. Schnable, P. S., and Springer, N. M. (2013) Progress toward understanding heterosis in crop plants. *Annu. Rev. Plant Biol.* **64**, 71–88

26. Davenport, C. B. (1908) Degeneration, albinism and inbreeding. *Science* **28**, 454–455

27. Bruce, A. B. (1910) The Mendelian theory of heredity and the augmentation of vigor. *Science* **32**, 627–628

28. Crow, J. F. (1948) Alternative hypotheses of hybrid vigor. *Genetics* **33**, 477–487

29. Richey, F. D. (1942) Mock-dominance and hybrid vigor. *Science* **96**, 280–281

30. Powers, L. (1944) An expansion of Jones's theory for the explanation of heterosis. *Am Nat.* **78**, 275–280

31. Becker, H. C. (1993) *Pflanzenzüchtung,* pp 1–368, UTB, Eugen Ulmer Verlag, Stuttgart, Germany.

32. East, E. M. (1936) Heterosis. *Genetics* **21**, 375–397

33. Damerval, C., Maurice, A., Josse, J. M., and de Vienne, D. (1994) Quantitative trait loci underlying gene product variation: A novel perspective for analyzing regulation of genome expression. *Genetics* **137**, 289–301

34. Cox, J., and Mann, M. (2011) Quantitative, high-resolution proteomics for

35. Bantscheff, M., Lemeer, S., Savitski, M. M., and Kuster, B. (2012) Quantitative mass spectrometry in proteomics: Critical review update from 2007 to the present. *Anal. Bioanal. Chem.* **404**, 939–965

36. Timberlake, W. E., Frizzell, M. A., Richards, K. D., and Gardner, R. C. (2011) A new yeast genetic resource for analysis and breeding. *Yeast* **28**, 63–80

37. Lindegren, C. C., Braham, J. E., and Calle, J. D. (1953) Heterosis in *Saccharomyces*. *Nature* **172**, 800–802

38. Steinmetz, L. M., Sinha, H., Richards, D. R., Spiegelman, J. I., Oefner, P. J., McCusker, J. H., and Davis, R. W. (2002) Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416**, 326–330

39. Schaefke, B., Emerson, J. J., Wang, T.-Y., Lu, M.-Y. J., Hsieh, L.-C., and Li, W.-H. (2013) Inheritance of gene expression level and selective constraints on trans- and cis-regulatory changes in yeast. *Mol. Biol. Evol.* **30**, 2121–2133

40. Plech, M., de Visser, J. A., and Korona, R. (2014) Heterosis is prevalent among domesticated but not wild strains of *Saccharomyces cerevisiae*. *Genes Genom. Genet.* **4**, 315–323

41. Shapira, R., Levy, T., Shaked, S., Fridman, E., and David, L. (2014) Extensive heterosis in growth of yeast hybrids is explained by a combination of genetic models. *Heredity*, **113**, 316–326

42. Romano, P., Soli, M. G., Suzzi, G., Grazia, L., and Zambonelli, C. (1985) Improvement of a wine *Saccharomyces cerevisiae* strain by a breeding program. *Appl. Environ. Microbiol.* **50**, 1064–1067

43. Jolly, N. P., Janse, B. J. H., Rooyen, T. J. V., and Louw, J. H. (1993) Hybridization and typing of yeasts used in sparkling wine fermentations. *Am. J. Enol. Vitic.* **44**, 217–226

44. Belloch, C., Orlic, S., Barrio, E., and Querol, A. (2008) Fermentative stress adaptation of hybrids within the *Saccharomyces sensu stricto* complex. *Int. J. Food Microbiol.* **122**, 188–195

45. Arroyo-López, F. N., Orlić, S., Querol, A., and Barrio, E. (2009) Effects of temperature, pH and sugar concentration on the growth parameters of *Saccharomyces cerevisiae*, *S. kudriavzevii* and their interspecific hybrid. *Int. J. Food Microbiol.* **131**, 120–127

46. Tronchoni, J., Gamero, A., Arroyo-López, F. N., Barrio, E., and Querol, A. (2009) Differences in the glucose and fructose consumption profiles in diverse *Saccharomyces* wine species and their hybrids during grape juice fermentation. *Int. J. Food Microbiol.* **134**, 237–243

47. Gamero, A., Tronchoni, J., Querol, A., and Belloch, C. (2013) Production of aroma compounds by cryotolerant *Saccharomyces* species and hybrids at low and moderate fermentation temperatures. *J. Appl. Microbiol.* **114**, 1405–1414

48. Borneman, A. R., Desany, B. A., Riches, D., Affourtit, J. P., Forgan, A. H., Pretorius, I. S., Egholm, M., and Chambers, P. J. (2012) The genome sequence of the wine yeast VIN7 reveals an allotriploid hybrid genome with *Saccharomyces cerevisiae* and *Saccharomyces kudriavzevii* origins. *FEMS Yeast Res.* **12**, 88–96

49. Erny, C., Raoult, P., Alais, A., Butterlin, G., Delobel, P., Matei-Radoi, F., Casaregola, S., and Legras, J. L. (2012) Ecological success of a group of *Saccharomyces cerevisiae/Saccharomyces kudriavzevii* hybrids in the Northern European wine making environment. *Appl. Environ. Microbiol.* **78**, 3256–3265

50. Masneuf-Pomarède, I., Bely, M., Marullo, P., Lonvaud-Funel, A., and Dubourdieu, D. (2010) Reassessment of phenotypic traits for *Saccharomyces bayanus* var. *uvarum* wine yeast strains. *Int. J. Food Microbiol.* **139**, 79–86

51. Kishimoto, M., and Goto, S. (1995) Growth temperatures and electrophoretic karyotyping as tools for practical discrimination of *Saccharomyces bayanus* and *Saccharomyces cerevisiae*. *J. Gen. Appl. Microbiol.* **41**, 239–247

52. Naumov, G. (1996) Genetic identification of biological species in the *Saccharomyces sensu stricto* complex. *J. Ind. Microbiol. Biotechnol.* **17**, 295–302

53. Albertin, W., da Silva, T., Rigoulet, M., Salin, B., Masneuf-Pomarede, I., de Vienne, D., Sicard, D., Bely, M., and Marullo, P. (2013) The mitochondrial genome impacts respiration but not fermentation in interspecific *Saccharomyces* hybrids. *PLoS ONE* **8**, e75121

54. Liti, G., Carter, D. M., Moses, A. M., Warringer, J., Parts, L., James, S. A., Davey, R. P., Roberts, I. N., Burt, A., Koufopanou, V., Tsai, I. J., Bergman, C. M., Bensasson, D., O'Kelly, M. J., van Oudenaarden, A., Barton,

D. B., Bailes, E., Nguyen, A. N., Jones, M., Quail, M. A., Goodhead, I., Sims, S., Smith, F., Blomberg, A., Durbin, R., and Louis, E. J. (2009) Population genomics of domestic and wild yeasts. *Nature* **458,** 337–341

55. Albertin, W., Marullo, P., Aigle, M., Bourgais, A., Bely, M., Dillmann, C., De Vienne, D., and Sicard, D. (2009) Evidence for autotetraploidy associated with reproductive isolation in *Saccharomyces cerevisiae*: Towards a new domesticated species. *J. Evol. Biol.* **22,** 2157–2170

56. Richards, K. D., Goddard, M. R., and Gardner, R. C. (2009) A database of microsatellite genotypes for *Saccharomyces cerevisiae. Antonie Van Leeuwenhoek* **96,** 355–359

57. Masneuf-Pomarède, I., Le Jeune, C., Durrens, P., Lollier, M., Aigle, M., and Dubourdieu, D. (2007) Molecular typing of wine yeast strains *Saccharomyces bayanus* var. *uvarum* using microsatellite markers. *Syst. Appl. Microbiol.* **30,** 75–82

58. Blein-Nicolas, M., Albertin, W., Valot, B., Marullo, P., Sicard, D., Giraud, C., Huet, S., Bourgais, A., Dillmann, C., de Vienne, D., and Zivy, M. (2013) Yeast proteome variations reveal different adaptive responses to grape must fermentation. *Mol. Biol. Evol.* **30,** 1368–1383

59. Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. (2008) ProteoWizard: Open source software for rapid proteomics tools development. *Bioinformatics* **24,** 2534–2536

60. Ferry-Dumazet, H., Houel, G., Montalent, P., Moreau, L., Langella, O., Negroni, L., Vincent, D., Lalanne, C., de Daruvar, A., Plomion, C., Zivy, M., and Joets, J. (2005) PROTICdb: A web-based application to store, track, query, and compare plant proteome data. *Proteomics* **5,** 2069–2081

61. Langella, O., Zivy, M., and Joets, J. (2007) The PROTICdb database for 2-DE proteomics. *Methods Mol. Biol.* **355,** 279–303

62. Langella, O., Valot, B., Jacob, D., Balliau, T., Flores, R., Hoogland, C., Joets, J., and Zivy, M. (2013) Management and dissemination of MS proteomic data with PROTICdb: Example of a quantitative comparison between methods of protein extraction. *Proteomics* **13,** 1457–1466

63. Valot, B., Langella, O., Nano, E., and Zivy, M. (2011) MassChroQ : A versatile tool for mass spectrometry quantification. *Proteomics* **11,** 3572–3577

64. Blein-Nicolas, M., Xu, H., de Vienne, D., Giraud, C., Huet, S., and Zivy, M. (2012) Including shared peptides for estimating protein abundances: A significant improvement for quantitative proteomics. *Proteomics* **12,** 2797–2801

65. Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Methodol.* **57,** 289–300

66. R Core Team (2013) *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, VA.

67. Fell, D. A. (1997) Understanding the Control of Metabolism, pp. 1–300, Portland Press, Oxford Brookes University, London, UK

68. Ruepp, A., Zollner, A., Maier, D., Albermann, K., Hani, J., Mokrejs, M., Tetko, I., Güldener, U., Mannhaupt, G., Münsterkötter, M., and Mewes, H. W. (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res.* **32,** 5539–5545

69. Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., André, B., Arkin, A. P., Astromoff, A., El-Bakkoury, M., Bangham, R., Benito, R., Brachat, S., Campanaro, S., Curtiss, M., Davis, K., Deutschbauer, A., Entian, K.-D., Flaherty, P., Foury, F., Garfinkel, D. J., Gerstein, M., Gotte, D., Güldener, U., Hegemann, J. H., Hempel, S., Herman, Z., Jaramillo, D. F., Kelly, D. E., Kelly, S. L., Kötter, P., LaBonte, D., Lamb, D. C., Lan, N., Liang, H., Liao, H., Liu, L., Luo, C., Lussier, M., Mao, R., Menard, P., Ooi, S. L., Revuelta, J. L., Roberts, C. J., Rose, M., Ross-MacDonald, P., Scherens, B., Schimmack, G., Shafer, B., Shoemaker, D. D., Sookhai-Mahadeo, S., Storms, R. K., Strathern, J. N., Valle, G., Voet, M., Volckaert, G., Wang, C., Ward, T. R., Wilhelmy, J., Winzeler, E. A., Yang, Y., Yen, G., Youngman, E., Yu, K., Bussey, H., Boeke, J. D., Snyder, M., Philippsen, P., Davis, R. W., and Johnston, M. (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418,** 387–391

70. Deutschbauer, A. M., and Davis, R. W. (2005) Quantitative trait loci mapped to single-nucleotide resolution in yeast. *Nat. Genet.* **37,** 1333–1340

71. Teixeira, M. C., Monteiro, P., Jain, P., Tenreiro, S., Fernandes, A. R., Mira, N. P., Alenquer, M., Freitas, A. T., Oliveira, A. L., and Sá-Correia, I.

(2006) The YEASTRACT database: A tool for the analysis of transcription regulatory associations in *Saccharomyces cerevisiae. Nucleic Acids Res.* **34,** D446–D451

72. Monteiro, P. T., Mendes, N. D., Teixeira, M. C., d'Orey, S., Tenreiro, S., Mira, N. P., Pais, H., Francisco, A. P., Carvalho, A. M., Lourenço, A. B., Sá-Correia, I., Oliveira, A. L., and Freitas, A. T. (2008) YEASTRACT-DISCOVERER: New tools to improve the analysis of transcriptional regulatory associations in *Saccharomyces cerevisiae. Nucleic Acids Res.* **36,** D132–D136

73. Abdulrehman, D., Monteiro, P. T., Teixeira, M. C., Mira, N. P., Lourenço, A. B., dos Santos, S. C., Cabrito, T. R., Francisco, A. P., Madeira, S. C., Aires, R. S., Oliveira, A. L., Sá-Correia, I., and Freitas, A. T. (2011) YEASTRACT: Providing a programmatic access to curated transcriptional regulatory associations in Saccharomyces cerevisiae through a web services interface. *Nucleic Acids Res.* **39,** D136–D140

74. Teixeira, M. C., Monteiro, P. T., Guerreiro, J. F., Gonçalves, J. P., Mira, N. P., dos Santos, S. C., Cabrito, T. R., Palma, M., Costa, C., Francisco, A. P., Madeira, S. C., Oliveira, A. L., Freitas, A. T., and Sá-Correia, I. (2014) The YEASTRACT database: An upgraded information system for the analysis of gene and genomic transcription regulation in Saccharomyces cerevisiae. *Nucleic Acids Res.* **42,** D161–D166

75. Fell, D. A. (1997) *Understanding the Control of Metabolism*, pp. 1–300, Portland Press, Oxford Brookes University, London, UK

76. Kacser, H., and Burns, J. A. (1981) The molecular basis of dominance. *Genetics* **97,** 639–666

77. Rossignol, R., Faustin, B., Rocher, C., Malgat, M., Mazat, J.-P., and Letellier, T. (2003) Mitochondrial threshold effects. *Biochem. J.* **370,** 751–762

78. Firczuk, H., Kannambath, S., Pahle, J., Claydon, A., Beynon, R., Duncan, J., Westerhoff, H., Mendes, P., and McCarthy, J. E. (2013) An in vivo control map for the eukaryotic mRNA translation machinery. *Mol. Syst. Biol.* **9,** 635

79. Poker, G., Zarai, Y., Margaliot, M., and Tuller, T. (2014) Maximizing protein translation rate in the non-homogeneous ribosome flow model: a convex optimization approach. *J. R. Soc. Interface R. Soc.* **11,** 20140713

80. Khan, Z., Bloom, J. S., Amini, S., Singh, M., Perlman, D. H., Caudy, A. A., and Kruglyak, L. (2012) Quantitative measurement of allele-specific protein expression in a diploid yeast hybrid by LC-MS. *Mol. Syst. Biol.* **8,** 602

81. Hahn, S., and Young, E. T. (2011) Transcriptional regulation in *Saccharomyces cerevisiae*: Transcription factor regulation and function, mechanisms of initiation, and roles of activators and coactivators. *Genetics* **189,** 705–736

82. Lemos, B., Araripe, L. O., Fontanillas, P., and Hartl, D. L. (2008) Dominance and the evolutionary accumulation of cis- and trans-effects on gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 14471–14476

83. McManus, C. J., Coolon, J. D., Duff, M. O., Eipper-Mains, J., Graveley, B. R., and Wittkopp, P. J. (2010) Regulatory divergence in *Drosophila* revealed by mRNA-seq. *Genome Res.* **20,** 816–825

84. Tirosh, I., Reikhav, S., Levy, A. A., and Barkai, N. (2009) A yeast hybrid provides insight into the evolution of gene expression regulation. *Science* **324,** 659–662

85. Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., Botstein, D., and Brown, P. O. (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell* **11,** 4241–4257

86. Albertin, W., Marullo, P., Bely, M., Aigle, M., Bourgais, A., Langella, O., Balliau, T., Chevret, D., Valot, B., da Silva, T., Dillmann, C., de Vienne, D., and Sicard, D. (2013) Linking post-translational modifications and variation of phenotypic traits. *Mol. Cell. Proteomics* **12,** 720–735

87. Kellis, M., Patterson, N., Endrizzi, M., Birren, B., and Lander, E. S. (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423,** 241–254

88. Kim, D.-U., Hayles, J., Kim, D., Wood, V., Park, H.-O., Won, M., Yoo, H.-S., Duhig, T., Nam, M., Palmer, G., Han, S., Jeffery, L., Baek, S.-T., Lee, H., Shim, Y. S., Lee, M., Kim, L., Heo, K.-S., Noh, E. J., Lee, A.-R., Jang, Y.-J., Chung, K.-S., Choi, S.-J., Park, J.-Y., Park, Y., Kim, H. M., Park, S.-K., Park, H.-J., Kang, E.-J., Kim, H. B., Kang, H.-S., Park, H.-M., Kim, K., Song, K., Song, K. B., Nurse, P., and Hoe, K.-L. (2010) Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe. Nat. Biotechnol.* **28,** 617–623

89. Georgi, B., Voight, B. F., and Bućan, M. (2013) From mouse to human: Evolutionary genomics analysis of human orthologs of essential genes. *PLoS Genet.* **9,** e1003484

90. Wool, I. G., Chan, Y.-L., and Glück, A. (1995) Structure and evolution of mammalian ribosomal proteins. *Biochem. Cell Biol.* **73,** 933–947

91. Fu, H., Doelling, J. H., Rubin, D. M., and Vierstra, R. D. (1999) Structural and functional analysis of the six regulatory particle triple-A ATPase subunits from the *Arabidopsis* 26S proteasome. *Plant J.* **18,** 529–539

92. Wright, S. (1934) Molecular and evolutionary theories of dominance. *Am. Nat.* **68,** 24–53

93. de Vienne, D., Bost, B., Fiévet, J., Zivy, M., and Dillmann, C. (2001) Genetic variability of proteome expression and metabolic control. *Plant Physiol. Biochem.* **39,** 271–283

94. Mazat, J.-P., Rossignol, R., Malgat, M., Rocher, C., Faustin, B., and Letellier, T. (2001) What do mitochondrial diseases teach us about normal mitochondrial functions that we already knew: Threshold expression of mitochondrial defects. *Biochim. Biophys. Acta* **1504,** 20–30

95. Gout, J.-F., Duret, L., and Kahn, D. (2009) Differential retention of metabolic genes following whole-genome duplication. *Mol. Biol. Evol.* **26,** 1067–1072

96. Giorgetti, L., Siggers, T., Tiana, G., Caprara, G., Notarbartolo, S., Corona, T., Pasparakis, M., Milani, P., Bulyk, M. L., and Natoli, G. (2010) Non-cooperative interactions between transcription factors and clustered DNA binding sites enable graded transcriptional responses to environmental inputs. *Mol. Cell* **37,** 418–428

97. MacLean, R. C. (2010) Predicting epistasis: An experimental test of metabolic control theory with bacterial transcription and translation. *J. Evol. Biol.* **23,** 488–493

98. Naumov, G. I., James, S. A., Naumova, E. S., Louis, E. J., and Roberts, I. N. (2000) Three new species in the *Saccharomyces sensu stricto* complex: *Saccharomyces cariocanus, Saccharomyces kudriavzevii* and *Saccharomyces mikatae. Int. J. Syst. Evol. Microbiol.* **50,** 1931–1942

99. Demuyter, C., Lollier, M., Legras, J.-L, and Le Jeune, C. (2004) Predominance of *Saccharomyces uvarum* during spontaneous alcoholic fermentation, for three consecutive years, in an Alsatian winery. *J. Appl. Microbiol.* **97,** 1140–1148

100. Albertin, W., Marullo, P., Aigle, M., Dillmann, C., de Vienne, D., Bely, M., and Sicard, D. (2011) Population size drives industrial *Saccharomyces cerevisiae* alcoholic fermentation and is under genetic control. *Appl. Environ. Microbiol.* **77,** 2772–2784

101. Marullo, P., Bely, M., Masneuf-Pomarède, I., Pons, M., Aigle, M., and Dubourdieu, D. (2006) Breeding strategies for combining fermentative qualities and reducing off-flavor production in a wine yeast model. *FEMS Yeast Res.* **6,** 268–279

102. Marullo, P., Mansour, C., Dufour, M., Albertin, W., Sicard, D., Bely, M., and Dubourdieu, D. (2009) Genetic improvement of thermo-tolerance in wine *Saccharomyces cerevisiae* strains by a backcross approach. *FEMS Yeast Res.* **9,** 1148–1160