Analyzing microbial consortia for biotechnological processes design

Navarrete-Bolaños, J.L*., Serrato-Joya, O., Botello-Alvarez, E. Jimenez-Islas, H., Cardenas-Manriquez, M., Conde-Barajas, E., Rico-Martínez, R.

Departamento de Ingeniería Química-Bioquímica, Instituto Tecnológico de Celaya, Av. Tecnológico s/n, C.P. 38010. Celaya, Gto., México. Tel: (+52) 461 61 1 75 75, Fax: (+52) 461 61 1 79 79.

For decades industrial fermentation processes have been used in the manufacture of a wide range of products including enzymes, foods, beverages, chemical feedstocks, fuels and pharmaceuticals, and as the basis for clean technologies employed for waste treatment and pollution control. As the first stage in any fermentation process, inoculum consistency is clearly important. In general, in fermentation processes two types of inoculum are used: (i) endogenous microorganisms isolated from natural sources and (ii) pure cultured microorganisms obtained from collections, but that also were isolated from natural sources. In nature, microorganisms do not live isolated in a certain space and time; they coexist with many different microorganisms establishing relationships that have an effect in the biological adequacy of all interacting species. As a result, different species (consortia) inside an ecosystem propagate with very different dynamics, which depend on their genetic potentiality as well as their capacity of adjustment to the micro environmental conditions. Although historically fermentation research has focused on characterizing pure cultures, such efforts have almost always a common origin: microbial consortia studies from which a pure culture involved in a natural phenomenon is isolated, and selected. Focusing on single cultures, however, may neglect important interactions between microorganisms isolated from microsystems were high levels of nonlinear interactions are present. Such interactions can range from synergism to antagonism among microorganisms, and may depend on, besides the microbial groups and species involved, the substrate type and availability. Even if these interactions are important, studying sets of isolated strains retains paramount importance as the basis for constructing microbial consortia and define starter inoculum. Whether an inoculum should contain a single microorganism, or a mixture, in order to design and develop a successful process, requires understanding the precise role and overall contribution of each microorganisms to the fermentation process; however, a procedure to reach such understanding has not been, up to now, established. In this work a methodology to analyze microbial consortia from natural spontaneous fermentation is presented. An experimental strategy to reach an improved understanding of microorganisms interactions and define a microbial system to be used on fermentations in order to design and optimize processes is presented.

Introduction

Microorganisms, living or dead, genetically modified, otherwise selected or wild, and material derived from these organisms (e.g., enzymes, antibodies, cellular organelles) are used commonly as biological systems on industrial fermentations (Crueger and Crueger, 1990). These systems, or biocatalysts, carry out a vast array of biochemical reactions which produce useful chemicals (e.g., foods, pharmaceuticals, fuels, pesticides, fine chemicals, fertilizers, plastics, and industrial catalysts to name a few), modify existing ones (e.g., biotransformation in control pollution), and degrade unwanted ones (e.g., waste treatment). Upgrading fermentation processes requiring microbial cultivation often depends on maintenance in medium culture of the single species with which the fermenter was initially inoculated. On the other hand, biological treatment of liquid, solid and gaseous effluents for environmental pollution control or abatement use mixed microbial populations as the biocatalytic entities responsible for degradation of waste (Moo-Young and Chisti, 1994). For several thousand years, natural or spontaneous fermentation has been used to obtain many types of foods, fermented feeds and beverages that comprise a significant proportion of our diet. These processes are used to prepare and preserve products in order

* Corresponding author: e-mail: jlnb@itc.mx., Phone: (+52) 461 61 1 75 75, Fax: (+52) 461 61 1 79 79.

to: increase nutritive value, improve flavour, and others qualities associated with edibility (Achi, 2005). However, in general the traditional fermented methods render low yields, and poor and variable quality as a result of two main factors: suboptimal processing conditions associated to the absence of appropriate technology, and partial knowledge of the endogenous microorganisms responsible of substrate transformation (Achi, 2005). In strict term, the design of a bioprocess and the engineering of the process equipment require of previous knowledge of the biological system that the inoculum conforms in order to specify their supply (e.g. culture medium, temperature, pH, shear forces, oxygen supply, CO_2 removal, antifoaming characteristics, etc) for fermentation, and of the microorganisms growth and the product formation in order to define a controllable, predicable, and efficient process (Moo-Young and Chisti, 1994; Moo-Young and Blanch, 1981).

In spite of significant contributions in the fields of microbiology and biochemistry that have lead to improved understanding of natural fermentation phenomena, microorganism selection remains largely based on pure culture evaluation and relies in the use of single strains. If the microorganism interactions are evaluated and included in such selection process, a microbial consortium may outperform the results achieved by pure cultures in almost every case. The aim of this work is to present a strategy to analyze microbial consortia from natural spontaneous fermentation. The strategy stresses the importance of microbiological tools to isolate, analyze the performance and identify microorganisms. A special emphasis is given to tools to analyze consortia based on experimental mixture design and surface response methodology. These tools allow the study of interactions, and lead to the definition of a starter inoculum, that can be formed by single strain or a microbial consortium, as the main requirement for designing and optimizing a fermentation process.

Microorganisms isolation

The isolation of microorganisms (such as bacteria, fungi or yeast) from complex mixed cultures, and their cultivation in a pure culture, is an essential prerequisite for their precise identification and characterization. Several different processes are used for this purpose, including the classic plate-casting process, as well as dilution methods or selective enrichment based on nutritional requirements and sensitivities to various chemicals. In all cases, it is necessary to consider the characteristics of the product and of the process taking place. The first will indicate where to take microorganisms samples, while the second will dictate how to design the isolation media for the microorganisms present in the samples. A complete sampling effort should contain the following steps: (1st) list of groups of microorganisms that are going to be isolated (the medium and culture techniques are different for bacteria and fungi); (2nd) microsystem or habitat description in which they samples are going to be taken, (3rd) sample grouping based on the material's origin (soil, water, vegetable material, etc.); (4th) description of important environmental parameters (pH, redox potential, temperature, oxygen supplies, etc.); (5th) listing of natural substrata available for autochthonous microorganisms in their natural environment, (6th) design of techniques for isolation using the information from the previous points, (7th) techniques for evaluation of isolation, and (8th) enrichment procedures for microorganisms of interest (Ellwood, 1972; Stanier, et al.,1986; Prescott, et al., 1999).

The optimal recovery of microorganisms requires inoculation on an appropriate medium. There are three types of media commonly used for bacterial isolation: Selective medium that allows the growth of certain microorganisms, while inhibiting others; Differential medium which, because of its chemical composition, permits the differentiation of organisms by their reaction to specific components in the medium; and enriched medium which contains special nutrients. A medium can have more that one property. For example, MacConkeys agar supports preferentially the growth of Gram negative organisms and inhibits the growth of most Gram positive organisms due to the presence of violet crystals and bile salt, in addition, it is used to isolate microorganisms with affinity for lactose; therefore, this medium is both selective and differential. Once the plates have been inoculated using the streak method, the samples must be incubated. After incubation, microorganisms isolation can proceed based on appearance on the agar plates (size, shape, texture, colour, etc.). The representative isolated colonies are then processed by heat-fixed smearing, stained, and examined microscopically. The prospective isolated

©FORMATEX 2007

colony is subcultured to a second agar plate, and once again streaked for isolation and incubated. These steps are repeated until well-isolated colonies lead to a pure culture. A pure culture must meet macroscopic and microscopic criteria. At the macroscopic level, all colonies should be alike on subculture, except for minor variations, and they should resemble the parent type from which the inoculum was picked. While at the microscopic level, all the colonies should have the same morphology and stain reaction. Once pure cultures were isolated using classical microbiology tools, is suitable to realize biochemical test that are based on substrate utilization and biochemical reaction in order to increase the reliability that the isolated cultures are pure (Stanier, et al.,1986; Prescott, et al., 1999; Korneel, et al., 2004).

Microbial consortia analysis

After the extensive isolation process, the screening of microorganisms must be carried out in order to select, of a group of microorganisms, those involved in the fermentation of interest. In general, the results of consortia studies that include combinations of pure cultures, as a mixed inoculum, face experimental difficulties related to both the syntrophic association and interpretation of the phenomenon. As a consequence, researchers often prefer to use only one of the candidate strains. The use of pure cultures in fermentation processes has had great impact on all facets of human civilization. However, in order to design new fermentation process, or optimize the existing ones, consortia studies must be considered in order to take advantage of the high levels of non-linear interactions among consortia members. Up to date, this approach has not been extensively described in the literature. Among the described alternatives for consortia studies, the use of genetic algorithms, as well as those efforts based on experimental design and response surface methodology, are the most promising (Vandecasteele, 2003; Navarrete et al., 2003).

Genetic algorithms

Genetic algorithms mimic the process of mutation and selection fundamental to evolutionary process, and are based on the principle of survival of the fittest. This method has been used for constructing microbial consortia for dry biomass production and optimization (Vandecasteele, 2003). Seemingly the genetic algorithm method applied quickly eliminated certain strains with negative influence, while promoting those with a dominant overall positive influence on the consortia, but the strategy has been not demonstrated.

Experimental mixture design

A mixed design is necessary when the response of interest is a function of the relative proportions of the factors (components) X_1, X_2, X_3,…X_p that constitute a mixture, and therefore allowed value for every component is limited to [0,1] ($0 \le X_i \le 1$), and the sum of all components present in the mixture must be equal to one or 100% ($X_1 + X_2 + X_3 + \ldots X_p = 1$ or 100%). Therefore, the component values are not independent (Montgomery, 2005). This restriction is illustrated in figure 1-a, where each variable is represented as a spatial coordinate. However, in assay formulation, the constraints are not always realistic because one or more of the components might not take values of cero or one. The reason is simple, for example if the hardness of cement depend of the relative amounts of cement (X_1), water (X_2) and aggregate (X_3), a formulation of 0% of cement, 100% water, and 0% aggregate is not acceptable. In these cases, the space value for every variable must be among w_i and one ($w_i \le X_i \le 1$), where w_i is the lower limit for variable *i* (figure 1-b).

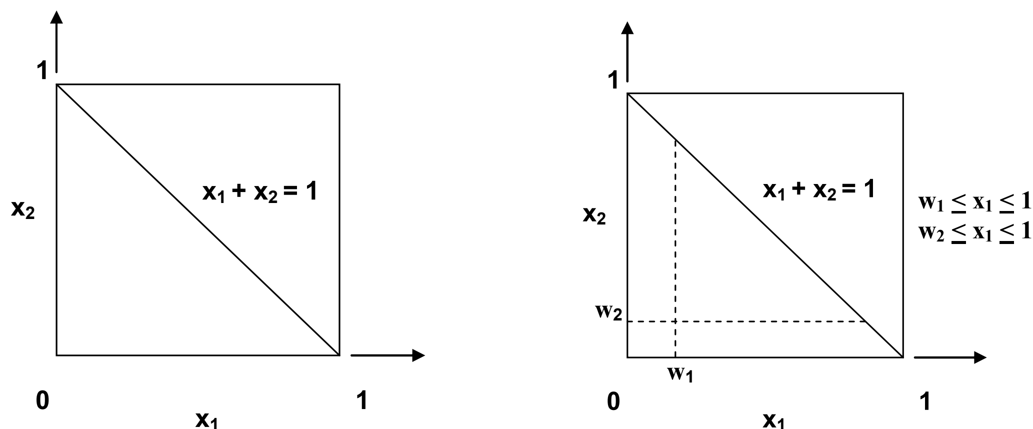

Figure 1. Factor space for two-component mixture: (a) constrained factor space, (b) lower-bound constraints.

In the evaluation of isolated microorganisms from consortia, a cero value for a variable means that a microorganism is not present in the mixture. While a value of one means that a pure culture is being evaluated. In general, the constrained experimental region is conveniently represented by triangles with vertices corresponding to formulations consisting of pure strains. When there are three components of a mixture, the constrained region can be represented on trilinear coordinates as shown in figure 2. Each of the three sides of the graphic represents a mixture composed by only two of the three components (the components labelled on the opposite vertex). When the mixture contains more than three components, for example four components, the graphic takes the form of a pyramid with triangular base. Once again, the vertices are for pure components, the edges for binary mixtures, the faces for ternary mixtures, and the inner area for quaternary mixtures. Independently of the number of variables or components to be evaluated in the mixture design, three different approaches can be used; simplex design, simplex lattice design, and simplex centroid design (Montgomery, 2005). Of them, simplex centroid design are the most efficient.
The simplex centroid design for p components consist of points defined by the p permutation of (1,0,0,0,….0), the $\binom{p}{2}$ permutation of ($\frac{1}{2}$, $\frac{1}{2}$,0,…0), the $\binom{p}{3}$ permutations of ($\frac{1}{3}$, $\frac{1}{3}$, $\frac{1}{3}$,0,….0), etc, and the overall centroid ($\frac{1}{p}$, $\frac{1}{p}$, $\frac{1}{p}$,….,$\frac{1}{p}$). In general, the number of points in these design is given by N = 2^p -1. As an example for p = 3, the experimental scheme is presented in table 1.

 ©FORMATEX 2007

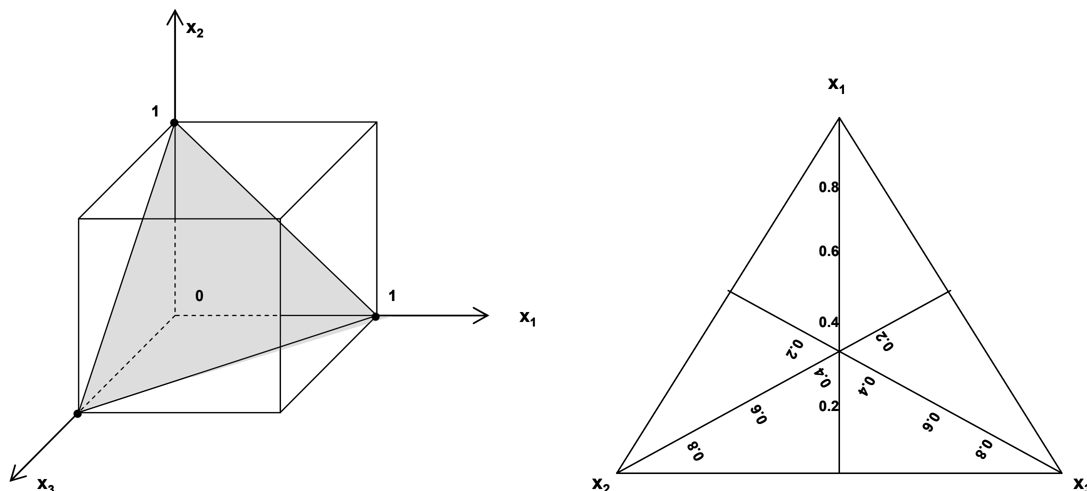

Figure 2. Constrained factor space for three-component mixture. (a) Triangle with vertices for pure component, (b) representation on trilinear coordinates.

Table 1. Simplex centroid design for p = 3.

Description	Assay	x_1	x_2	x_3	y
Pure components	1	1	0	0	y_1
	2	0	1	0	y_2
	3	0	0	1	y_3
Binary mixtures	4	$\frac{1}{2}$	$\frac{1}{2}$	0	y_4
	5	$\frac{1}{2}$	0	$\frac{1}{2}$	y_5
	6	0	$\frac{1}{2}$	$\frac{1}{2}$	y_6
Centroid	7	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	y_7

Mixture design analysis

A drawback of the simplex lattice design is that most of the experimental assays occur on the boundary of the experimental region and few in the interior region spanned by the p components. A similar drawback is presented by the simplex centroid design despite the fact that only includes half the number of assays in the boundary as compared with the simplex lattice design. Additionally, the simplex design, which includes assays in the interior, also can be questioned because it does not include assays with three or more components with exception of the assay that includes all components, defined as centroid and considered optional (Montgomery, 2005). In general, the three described strategies, specially the simplex centroid design and simplex design, have the same resolution to study up to five or fewer components. However, these design strategies present serious problems when six or more components are under study; firstly, the increasing number of required experimental essays (Table 2), and secondly, the difficulty in describing high order nonlinear interactions. Therefore, for studies that include six o more components, a better strategy involves the use of experimental designs that lead to the selection of significant variables with few experimental assays, such as fractional designs.

Table 2. Required assays and interaction between variables as a function of variable numbers

Variables	Assays	Pure components	Interactions				
			2 Comp.	3 Comp.	4 Comp.	5 Comp.	6 Comp.

2	3	2	1				
3	7	3	3	1			
4	15	4	6	4	1		
5	31	5	10	10	5	1	
6	63	6	15	20	15	6	1
7	127	7	21	35	35	21	7

Fractional mixture design

Fractional mixture designs have not been previously described as tools for consortia discrimination. In previous efforts (Navarrete et al., 2003) assays with mixtures including every possible combination of microorganisms were included as part of a fractional mixture design strategy, but a logical sequence for experimental scheme construction was lacking. Here, the construction of an experimental fractional scheme is described. Such scheme allows the analysis of two-variable interactions, includes combinations of all microorganisms, and it is based on balanced block partitions.

The logic behind the construction of the scheme for a fractional mixture design is as follows: the scheme must contain so many blocks as components to screen; each block, except the last (centroid), must contain so many assays as components to screen; in every block, except the last, the first column must begin with the combination of the components described in the block; the second column must begin with the value of the last variable of the first column, followed by all values contained in the first column; the third column must begin with the value of the last variable of the second column, followed by all values contained in the second column, and continue up to completing the definition of variable combination for all columns contained in the block (table 3). The last block must contain only the mixture that includes all components. In other words, every block; except the last, are constructed realizing cyclical permutations on the columns up to the formation of all the columns contained in the block. As a result, an incomplete balanced block design is obtained. In general the number of assays in a fractional mixture design is given for N = p (p-1) + 1.

Table 3. Fractional simplex centroid design for p components.

Description	Assay	x_1	x_2	x_3	....	x_p
Pure components	1	1	0	0	....	0
	2	0	1	0	....	0
	3	0	0	1	....	0
	:	:	:	:	....	:
	p	0	0	0	....	1
Binary mixtures	1	$\frac{1}{2}$	$\frac{1}{2}$	0	....	0
	2	0	$\frac{1}{2}$	$\frac{1}{2}$	....	0
	3	0	0	$\frac{1}{2}$	....	0
	:	:	:	:	....	:
	p	$\frac{1}{2}$	0	0	....	$\frac{1}{2}$
Ternary mixtures	1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	....	0
	2	0	$\frac{1}{3}$	$\frac{1}{3}$	....	0
	3	0	0	$\frac{1}{3}$	....	0
	:	:	:	:	....	:
	p-1	$\frac{1}{3}$	0	0	....	$\frac{1}{3}$

 ©FORMATEX 2007

	p	$\frac{1}{3}$	$\frac{1}{3}$	0	....	$\frac{1}{3}$
Quaternary mixtures	1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	....	0
	2	0	$\frac{1}{4}$	$\frac{1}{4}$	....	0
	3	0	0	$\frac{1}{4}$	....	0
	:	:	:	:	....	:
	p-2	$\frac{1}{4}$	0	0	....	$\frac{1}{4}$
	P-1	$\frac{1}{4}$	$\frac{1}{4}$	0	....	$\frac{1}{4}$
	p	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	....	$\frac{1}{4}$
Mixtures for p components for $5 \leq p \leq p$-1 values	1	$\frac{1}{p}$	$\frac{1}{p}$	$\frac{1}{p}$	....	0
	2	0	$\frac{1}{p}$	$\frac{1}{p}$	....	0
	3	0	0	$\frac{1}{p}$	....	0
	:	:	:	:	....	:
	p - n	$\frac{1}{p}$	0	0	....	0
	P - n+1	$\frac{1}{p}$	$\frac{1}{p}$	0	....	0
	:	:	:	:	....	:
	p -1	$\frac{1}{p}$	$\frac{1}{p}$	$\frac{1}{p}$	....	$\frac{1}{p}$
	p	$\frac{1}{p}$	$\frac{1}{p}$	$\frac{1}{p}$	....	$\frac{1}{p}$
Centroid	1	$\frac{1}{p}$	$\frac{1}{p}$	$\frac{1}{p}$	....	$\frac{1}{p}$

Simplex centroid mixture design versus fractional mixture design

Each experimental mixture design strategy has advantages and disadvantages, every case should be analyzed separately in order to select the best strategy. In general, problems that involve the analysis of components in systems with or without reaction should be analyzed using simplex centroid designs or fractional mixture design. But, when the study include six or more components; a fractional mixture design must be constructed and develop in order to realize a screening of the variables in order to identify a smaller number of components as relevant for the problem at hand. In almost every case, further, more refined, studies should be undertaken to guarantee the mixture optimization. For example, a study case that include 9 variables require that 511 assays are realized to be analyzed by simplex centroid design ($N = 2^p - 1$), and only 73 are necessary to be analyzed by fractional mixture design [$N = p(p-1) + 1$]. As a result, the fractional mixture design reduces the experimental assays up to 85%, and in addition, contains the freedom degrees to allow constructing a second order mathematical model to variable analysis.

Mixture design analysis via mathematical models

Independent to the mixture design used, the experimental results must be used to construct a mathematical mixture model to facilitate understanding, interpretation, and establish the functional relationship between the variables of study. Mixture model differ from the usual polynomials employed in response surface techniques. Such differences are related to the effect of the constrain $\sum x_i = 1$ that

affects the form of the polynomial models fitted to data from mixture experiments. Including the constrain, standard forms for mixture models, are then:

$$Linear: \quad E(y)=\sum_{i=1}^{p}\beta_i x_i$$

$$Quadratic: \quad E(y)=\sum_{i=1}^{p}\beta_i x_i+\sum\ \sum_{i\langle 1}^{p}\beta_{ij}x_i x_j$$

$$Cubic: \quad E(y)=\sum_{i=1}^{p}\beta_i x_i+\sum\ \sum_{i\langle 1}^{p}\beta_{ij}x_i x_j+\sum\ \sum_{i\langle 1}^{p}\delta_{ij}x_i x_j(x_i-x_j)+\sum\ \sum_{i\langle 1\langle k}^{p}\sum\ \beta_{ijk}x_i x_j x_k$$

$$Special\,cubic: \quad E(y)=\sum_{i=1}^{p}\beta_i x_i+\sum\ \sum_{i\langle 1}^{p}\beta_{ij}x_i x_j+\sum\ \sum_{i\langle 1\langle k}^{p}\sum\ \beta_{ijk}x_i x_j x_k$$

The terms in these models have relatively simple interpretations. The linear terms describe the expected response for pure cultures (x_i = 1 and x_j = 0, when $j \neq i$), as well as the linear blending portions ($\sum_i^p \beta_i x_i$ is the mixture linear portion). When there is curvature arising from nonlinear mixing between component pairs (quadratic terms), the terms β_{ij} represent either synergistic or antagonistic effects. Higher-order terms are frequently necessary in the mixture model because of the constraint $\sum x_i = 1$. Furthermore, complex phenomena that require descriptions encompassing large experimental regions often require elaborated models.

Mixed models

The least square method is widely used to find or estimate the numerical values of the parameters to fit a function to a set of data and to characterize the statistical properties of estimates. Several variants of the method can be used: the simplest version is called ordinary least squares (OLS), a more sophisticated version is called weighted least squares (WLS), which often outperforms OLS because it can modulate the importance of each observation in the final solution. Recently additional variations, alternating least squares (ALS) and partial least squares (PLS), had also been proposed. The OLS method is typically used to fit regression models in designed experiments, where a single dependent variable or response [y] depends on *K* independent or regressor variables [$x_1, x_2, \ldots x_k$], and therefore their relation can be represented by the general following model [$y=\beta_1 x_1+\beta_2 x_2+\beta_3 x_3+.......\beta_k x_k+\varepsilon$]. This model describes a hyperplane in the *K*-dimensional space of the regressor variables [x_j]. The parameters β_j (j=0,1,2,…k) are called the *regression coefficients* and represent the expected change in response [*y*] per unit change in x_j when all the remaining independent variables x_i (i≠j) are held constant.
The method of least squared chooses the β's in the previous equation so that the sum of the squares of the errors [ε_i] is minimized. The least squares function is: $L=\sum_{i=1}^{n}\varepsilon_i^2$ or $L=\sum_{i=1}^{n}\left(y_i-\sum_{j=1}^{k}\beta_i x_{ij}\right)^2$. The function *L* is to be minimized with respect to β_0, β_1,…. β_k. the least squares estimators, say $\hat{\beta}_0$, $\hat{\beta}_1$,…. $\hat{\beta}_k$, must satisfy:

$$\left.\frac{\partial L}{\partial \beta_I}\right|_{\hat{\beta}_1,\ldots\hat{\beta}_k}=-2\sum_{i=1}^{n}(y_i-\sum_{j=1}^{k}\hat{\beta}_j x_{ij})=0 \text{ and } \left.\frac{\partial L}{\partial \beta_j}\right|_{\hat{\beta}_1,\ldots\hat{\beta}_k}=-2\sum_{i=1}^{n}(y_i-\sum_{j=1}^{k}\hat{\beta}_j x_{ij})x_{ij}=0 \text{ for } j=1,2,....k$$

From which one obtains:

$$\hat{\beta}_1\sum_{i=1}^{n}x_{i1}+\hat{\beta}_2\sum_{i=1}^{n}x_{i2}+\cdots\hat{\beta}_k\sum_{i=1}^{n}x_{ik}=\sum_{i=1}^{n}y_i$$

 ©FORMATEX 2007

$$\hat{\beta}_1 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^{n} x_{i1}x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{i1}x_{ik} = \sum_{i=1}^{n} x_{i1}y_i$$

$$\hat{\beta}_1 \sum_{i=1}^{n} x_{ik} + \hat{\beta}_i \sum_{i=1}^{n} x_{ik}x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{ik}x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{ik}^2 = \sum_{i=1}^{n} x_{ik}y_i$$

These equations are known as the least squares normal equations. Note that the number of equations is correlated to the number of unknown regression coefficients. Solution of the normal equations will be the least squares estimators of the regression coefficients $\hat{\beta}_1, \hat{\beta}_2, \ldots \hat{\beta}_k$. A simple alternative to obtain and solve the normal equations, obtaining the least squares estimators, relies on a matrix formulation. Any experimental design constructed and developed in the laboratory can be expressed in this manner.

Mixture formulations

Once selected the experimental mixture design to be constructed, the microorganisms to be evaluated must be cultured on specific solid culture medium and propagated on specific liquid culture medium to increase de biomass. Next, from each culture, one must take the number of viable cells that will form the total of the mixture (100%). For example, if the mixture contains 1,000,000 viable cells, the sum of all microorganisms presented in each mixture always must be equal to 1,000,000 viable cells [$X_1 + X_2 + X_3 + \ldots X_p = 1{,}000.000$].

Mixed models construction

Once the experimental assays are realized a mathematical mixture model that establishes the functional relationship between the variables of study must be constructed.

Firstly one proposes the mathematical model to capture the functional relationship between dependent variable (y) and independent variables (x_i), for example a polynomial model with interaction and quadratic terms:

$$y = \beta_1 x_1 + \beta_2 x_2 + .. + \beta_p x_p + \beta_{12} x_1 x_2 + .. + \beta_{1j} x_1 x_j + .. + \beta_{2j} x_2 x_j + \ldots + \beta_{i1} + \beta_{i2} + \ldots + \beta_{ij} x_i x_j + \varepsilon$$

The experimental scheme can be expressed as a matrix (*n x p*) with the levels for each independent variable:

$$\begin{array}{c} \\ X = \end{array} \begin{array}{c} \begin{array}{ccccc} \beta_1 & \beta_2 & \beta_3 & & \beta_k \end{array} \\ \begin{vmatrix} 1.0 & 0.0 & 0.0 & \cdots & 0.0 \\ 0.0 & 1.0 & 0.0 & \cdots & 0.0 \\ 0.0 & 0.0 & 1.0 & \cdots & 0.0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1/p & 1/p & 1/p & \cdots & 1/p \end{vmatrix} \end{array}$$

Where every column (x_i) corresponds to a regressor (β_i).

The matrix must be expanded to contain the same number of columns as regressors included in the proposed model. Additional columns are added to the matrix to account by the nonlinear terms, i.e. with values that are result of the product between variables according to the nonlinear regressor (for example, the values in the column of $x_i x_i$ interaction must be the square product of x_i).

$$X = \begin{array}{c} \begin{array}{ccccc|ccccc} \beta_1 & \beta_2 & \beta_3 & \cdots & \beta_k & \beta_{12} & \beta_{1k} & \beta_{2k} & \cdots & \beta_{ij} \end{array} \\ \left| \begin{array}{ccccc} 1.0 & 0.0 & 0.0 & \cdots & 0.0 \\ 0.0 & 1.0 & 0.0 & \cdots & 0.0 \\ 0.0 & 0.0 & 0.0 & \cdots & 1.0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1/p & 1/p & 1/p & \cdots & 1/p \end{array} \right| \left| \begin{array}{ccccc} 0.0 & 0.0 & 0.0 & \cdots & 0.0 \\ 0.0 & 0.0 & 0.0 & \cdots & 0.0 \\ 0.0 & 0.0 & 0.0 & \cdots & 0.0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1/2p & 1/2p & 1/2p & \cdots & 1/ij \end{array} \right| \end{array}$$

This matrix is used directly in the least square method.
The experimental results should be defined as vector **y** (*n x 1*). Finally, according to the mathematical model proposed, it is necessary to define a vector of the regression coefficients **β** (*p* x 1), and a vector of random errors **ε** (*n* x 1).

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix}, \qquad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \beta_k \\ \beta_{12} \\ \vdots \\ \beta_{1k} \\ \vdots \\ \beta_{2k} \\ \vdots \\ \beta_{kk} \end{bmatrix}, \qquad \text{and} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_p \end{bmatrix}$$

Allowing us to find the vector of least squares estimators, $\hat{\beta}$, that minimizes

$$L = \sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon'\varepsilon = (y - X\beta)'(y - X\beta)$$

Also, L may be expressed as

$$\begin{aligned} L &= y'y - \beta'X'y - y'X\beta + \beta'X'X\beta \\ &= y'y - 2\beta'X'y + \beta'X'X\beta \end{aligned}$$

since $\beta'X'y$ is a (1 x 1) matrix, or a scalar, and its transpose $(\beta'X'y)' = y'X\beta$ is the same scalar. The least squares estimators must satisfy the expression:

$$\left.\frac{\partial L}{\partial \beta}\right|_{\hat{\beta}} = -2X'y + 2X'X\hat{\beta} = 0$$

which simplifies to:

$$X'X\hat{\beta} = X'y$$

Thus, the least squares estimator of β is

$$\hat{\beta} = (X'X)^{-1} X'y$$

In this manner the mathematical model is defined, and used to describe the expected response for pure cultures ($x_i = 1$ and $x_j = 0$, when $j \neq i$), linear blending portions ($\sum_i^p \beta_i x_i$ is the mixture linear portion), and nonlinear mixing between component pairs (quadratic terms) for either synergistic or antagonistic effects. Next, and in order to evaluated interaction and discriminate unnecessary microorganisms, considered as potential candidates to form the microbial consortium, the model must be solved using any numerical method, for example Levenberg-Marquardt method, considering restrictions [($0 \leq X_i \leq 1$), and ($X_1 + X_2 + X_3 + \ldots X_p = 1$ or 100%)] under conditions to maximize the value of dependent variable (y). Likewise, the model must be used to construct contour graphics as is showed in the figure 3, in order to show the expected response for pure cultures (figure 3-a), antagonistic effects (figure31-b) or synergistic effects (Figure 3-c).

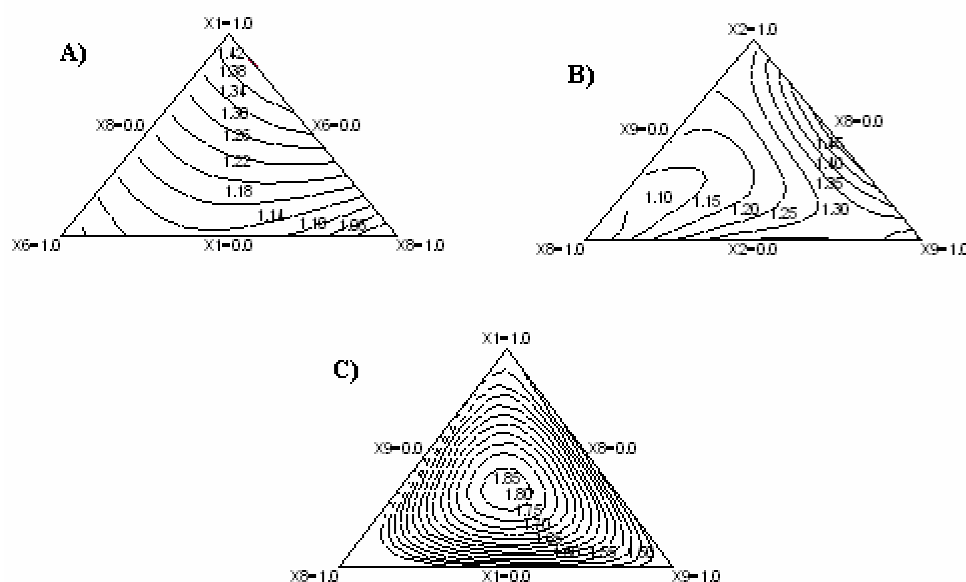

Figure 3. Response estimated (y) using the mathematical model.

Optimization strategy to select the starter inoculum

The experimental strategy described above can be used to optimize the starter inoculum based on two considerations associated with the number of variables to analyze. For mixtures of four or fewer components, the resulting scheme is similar to a simplex centroid design, and therefore higher-order models should be constructed in order to describe the response surface and perform the optimization (figure 3-c). For mixtures with five or more components, the response values can be described with second order model; this model shall be used for discrimination and selection of variables. Next, with a reduced number of variables a new scheme shall be constructed and analyzed. This strategy to use sequential experimental design, mathematical model construction, and analysis of variables up to optimization is similar to classical method of variables optimization defined as response surface methodology (Montgomery 2005).

Microorganisms identification

Once the microbial consortium has been evaluated and the microorganisms forming the starter inoculums selected, the identification of the microorganisms that will form the mixture must be performed. Several identification strategies at the phenotypic (Biochemical test, Fatty Acid Metil Ester analysis (FAME),

Matrix-Assisted Laser Desorption Ionization Time-Of-Flight Mass Spectrometry (MALDI-TOF-MS or MALDI), and genotypic (Ribotyping, Rep-PCR analysis, 16S or D2 rRNA analysis) levels have been developed leading to complete characterization of microorganisms. However, phenotypic methods based on cellular morphology and biochemistry tests rely on a more subjective identification of microorganisms. Identification based on FAME or MALDI is very reliable, it is based on mature technologies that are marketed by multiple companies, prepackaged kits with well-established quality control procedures, and mature instrumentation with extensive databases to identify the most commonly encountered microorganisms. In addition, the underlying techniques for phenotypic methods are familiar to microbiologists working in routine microbiological testing laboratories. In contrast, genotypic methods that are based on nucleic acid analysis are less subjective, less dependent on the culture method, and more reliable because nucleic acids are highly conserved by microbial species. However, the instrumentation required is expensive; the unit testing cost is greater than that for phenotypic methods, and a significant investment in training, validation, and database development is required to successfully implement these methods. Thus, genotypic identification is mainly restricted, for now, to research efforts. In general, one hopes that the results of phenotypic tests will be correlated with the genotypic characteristics of a colony, bringing about accurate identification of microorganisms (Latouche, et al. 1997).
Once the microbial consortium has been selected, and identified, fermentation assays must be developed using the microbial consortium as a starter inoculum in order to reach a predictable, controllable, efficient, and stable fermentation process.

Conclusions

A description of a methodology for the analysis and construction of microbial consortia has been given. The strategy allows selecting the appropriate experimental mixture design independent of microorganisms number. Fractional mixture design has been used to construct starter inoculums with, seemingly, successful results (Navarrete-Bolaños et al. 2003). Fractional mixture design and the response surface methodology algorithm has been applied to analyze and construct the microbial consortium for alcoholic fermentation to produce mezcal (Mexican alcoholic beverage related to tequila) using isolated strains associated to natural spontaneous fermentation (traditional method). The results showed that is possible to obtain a fermented extract with 6%(v/v) of alcohol in 72 hours of process time, value higher than the obtained by traditional processes (4%, v/v in 168 hours) (Serrato-Joya, et al. 2006). This approach could become a powerful tool for many problems related to microbial ecology, since it allows for a direct study of the relationship between an ecosystem composition and its productivity. Our efforts reveal the importance of isolation and microorganisms identification on selecting high-performing strains for a given application. Disconnecting isolation and identification from the selection process may lead to suboptimal results by discarding cultures with apparently the same "identity". While phenotypic identification methods are a valuable processing tool, their results may not be conclusive for some microorganisms. The absence of definitive genotypic identification may limit the early discrimination of candidate microorganisms in efforts based on endogenous strains.

References

[1] Achi, O.K. (2005). The potential for upgrading traditional fermented foods through biotechnology. African Journal of Biotechnology. 4 (5): 375-380.
[2] Crueger, W., Crueger, A. (1990). Biotechnology: A Textbook of Industrial Microbiology, 2nd Ed. Science Tech Publishers (Madison).
[3] Ellwood, D. C., Tempest, D. W. (1972). Effects of environment on bacterial wall content and composition. Adv. Microb. Physiol. 7:83–117.
[4] Korneel, R., Nico, B., Steven, D. S., Marc, V., Willy, V. (2004). Biofuel cells select for microbial consortia that Self-Mediate Electron Transfer. Applied and Environmental Microbiol. 70(9): 5373–5382.

 ©FORMATEX 2007

[5] Latouche, G.N., Heide-Marie, D., Lee, O.C., Mitchell, T.G., Sorrell, T.C., Meyer, W. (1997). Comparison of use of phenotypic and genotypic characteristics for identification of species of the anamorph genus *Candida* and related teleomorph yeast species. J. Clin. Microbiol. 35(12): 3171–3180.
[6] Montgomery DC (2005) Response surface methods and designs. In Montgomery DC, editor. Design and analysis of experiments. 6th ed. Wiley & Sons, NY, pp 405–463.
[7] Moo-Young, M., Blanch, H. W. (1981). Design of biochemical reactors: Mass transfer criteria for simple and complex systems. Adv. Biochem. Eng. 19: pp 1-69.
[8] Moo-Young, M., Chisti, Y. (1994). Biochemical engineering in biotechnology. Pure and Appl. Chem. 66(1): 117-136.
[9] Navarrete-Bolaños J.L., Jiménez-Islas H., Botello-Alvarez E., Rico-Martínez R. (2003). Mixed culture optimization for marigold flower ensilage via experimental design and response surface methodology. *J. Agric. Food Chem.* 51(8): 2206-2211.
[10] O. Serrato-Joya, M. Cárdenas Manríquez, E. Botello-Álvarez, H. Jiménez-Islas, R. Rico-Martínez y J. L. Navarrete-Bolaños. (2006) Estudio del Consorcio Microbiano Asociado a la Fermentación Alcohólica Natural del Extracto de Agave Salmiana para la Producción de Mezcal en Guanajuato. Memorias del XXVII Encuentro Nacional de AMIDIQ. Ixtapa-Zihuatanejo. Gro. Mex.
[11] Prescott, L.M,. Harley, J.P., Klein, D. A. (1999). Clinical Microbiology. In *Microbiology*, 4th edn. WCB/McGraw-Hill Inc. Boston. MA. pp 725-750.
[12] Stanier, R., Ingranhan, J., Wheelis, M., Painter, P. (1986). The methods of microbiology. In The Microbial World, 5th ed.; Prentice-Hall: Englewood Cliffs, New Jersey. pp 17-40.
[13] Vandecasteele, F.P.J. (2003). Constructing efficient microbial consortia using a genetic algorithm. Biological applications of genetic and evolutionary computation (BioGEC 2003) - Genetic and evolutionary computation conference (GECCO, 2003).