Report from the Genetics Working Group

ISAPP Meeting, May 3-5, 2002

Group members:

Dan O' Sullivan
Martin Kullen
Gwen Allison
Richard Hull
Elaine Vaughn
David Mills
Chris Kitts
David Pridmore
John McCormick
Todd Klaenhammer, Chair

Since Eli Metchnikoff published the "*The Prolongation of Life: Optimistic Studies",* the field of probiotics is now nearly 100 years old. Today it remains a theoretically based science that has not been deciphered scientifically or mechanistically. However, this is rapidly changing on a number of fronts. With the rapid progress over the past decade on the genetics of lactic acid bacteria and pending release of complete genome sequence information for the major probiotic species, the field is now armed with sophisticated microbiology and genetic tools. Incorporation of these genetic tools within a multidisciplinary scientific approach is expected to reveal the contributions of probiotics and commensals to general health and well being, and explicitly identify the mechanisms and corresponding host responses that provide the basis for their positive roles.

Major Probiotic strains/species. There are numerous probiotic strains, species, and genera that have been defined and characterized over the past one hundred years. Most notably, *Lactobacillus* and *Bifidobacterium* are most often considered in the probiotic category. While the list of potential probiotics continues to grow (Tannock, 1999; Sanders, 1999), we currently have a very small view of the possible collection of organisms that may constitute beneficial commensals or potential probiotic cultures. More than 500 microbial species are believed to occupy the human gastrointestinal tract and this composition remains largely unknown and highly variable within different locations and among different individuals. The microbial content of the small and large intestine is not adequately reflected by fecal analysis (Zoetendal et la., 2002), which has been the predominant sample analyzed to date. The application of of PCR-Denaturing Gradient Gel Electrophoresis (PCR-DGGE), Terminal Restriction Fragment Length Polymorphisms (TRFLP), and high-throughput sequencing of 16S rRNA libraries to the study of the microbial ecology of the GI tract has begun to identify the major culturable and non-culturable populations, and provides the means to study their changes over time and under different conditions (Kitts, 2001; Zoetendal, 1998). Fluorescent *in situ* hybridization in combination with flow cytometry is also facilitating the high-throughput enumeration of groups within the microbiota (Vaughan et al., 2002). Methods have now become available for whole genome amplification of uncultured cells (lower limit of approximately 1000 cells) where the functionality of 60% of the genes in the genome can be predicted based only on sequence analysis (Hawkins et al., 2002). As a result of these genetic approaches, our view of the microbial composition of the human GI tract will be expanded considerably in the years ahead, particularly in cataloging the collection of unculturable organisms occupying mucosal tissues.

These approaches will undoubtedly contribute vast taxonomic information about the microbial composition of the GI tract and other mucosal tissues, and provide a more complete landscape upon which one can measure the impact of probiotics to alter, protect or reestablish that collective flora. As the list of newly discovered commensal organisms constituting the normal

microflora continues to grow, it is also anticipated that potentially new probiotic organisms may be revealed.

Genome sequencing of known Probiotics/Commensals

Realizing their practical significance in fermentation, bioprocessing, agriculture, food, and more recently, medicine, the lactic acid bacteria have been the recent focus of intense genomic research. The first complete genome of the LAB group was published on *Lactococcus lactis* subsp. *lactis* IL1403 by Bolotin et al. (2001). Analysis of the 2.4 Mb genome revealed a number of unexpected findings: biosynthetic pathways for all 20 amino acids, albeit not all are functional; a complete set of late competence genes, 5 complete prophages, partial components for aerobic metabolism, and a wealth of ABC transporters reflecting the organisms fastidious lifestyle. Noting that some of these systems are not functional or complete, the genomic analysis of *Lactococcus* suggests an evolutionary trend toward minimization of the chromosome and elimination of unnecessary systems during adaptation to nutritionally complex environments, such as milk. International efforts have focused recently on probiotic lactic acid bacteria, with the genomes of three species now completed (*Lb. plantarum, Lb. johnsonii,* and *Lb. acidophilus)* and 7 other genomes underway including *Lb. gasseri, Lb.casei (2 strains), Lb. rhamnosus, Bifidobacterium longum (2 strains)* and *Bifidobacterium breve* (Klaenhammer et al., 2002). However, these results have yet to reach the public domain, with the exception of the draft genome sequence data for *Lb. gasseri* and *B. longum,* provided in 2002 by the U.S. Department of Energy-Joint Genome Institute (JGI). As part of their microbial genomes program (see http://www.jgi.doe.gov/JGI_microbial/html/index.html) 11 genomes are being sequenced in collaboration with the Lactic Acid Bacteria Genome Consortium (LABGC), composed of 10 scientists representing 7 universities in the U.S. Of the genomes being sequenced, three represent probiotic species (*Lb. gasseri, Lb. casei,* and *B. longum*), and three others (*L. lactis, St. thermophilus,* and *Lb. delbrueckii*) represent organisms that may be potentially used as intestinal delivery vehicles for biologics. As these sequences are generated, they will be available on the JGI website for public use. Timely public availability of genome information for various LAB species will catapult the fields' collective efforts to carry on with comparative and functional genomic analyses of the probiotic species within the LAB group.

More importantly, the opportunity now presents itself to compare the genetic content and organization of probiotic organisms against the growing number of genomes from commensal and pathogenic organisms (eg. *Bacteriodes, Streptococcus mutans, Streptococcus bovis, Streptococcus pneumonia, Clostridium, Listeria*, etc) (Hooper and Gordon, 2001). This analysis is expected to reveal key similarities and differences that reflect on both the habitat occupied and the lifestyle within that habitat. Ideally, our viewpoint will be augmented considerably by comparisons of closely related species, occupying similar versus dissimilar habitats, multiple genomes of the different strains within the same species, and multiple genomes of strains found in varying environments. Genomic regions that are expected to be identified in these analyses include:

- Conserved vs. distinct gene sets
- Putative virulence factors
- Horizontal evolution – gene transfer
- Minimal genome of probiotic cultures
- Altered GC content – islands/regions of adaptability (surrounding prophages, IS elements; expolysaccharides, bacteriocins, transposons) that may be critical to the

survival and functionality of commensal/probiotic organisms in their corresponding habitats

Genome sequences in the public domain are essential to the scientific progress of this field. In 2003 a greater collection of genome sequences will be publicly available , but even then the sequences are not likely to reflect the biodiversity that occupies these complex ecosystems. As a result, expanded sequencing capacity will continue to support genomic efforts to determine the microbiomes of organisms inhabiting the mouth, vagina, and regions within the GI tract (deVos, 2001; Hooper and Gordon, 2001). Viewing the metagenome, defined as the collective genomic content of a diverse "cell-wall less" population within an environment, is likely to reveal key functions and activities essential for survival, competition, and activity in that environment. Translating this information into the interactions, roles and functions of probiotic cultures promises to be an exciting frontier of science in the decade ahead. As these data accumulate, one major challenge will be continuous updates of genomes and genome sequences as each new organism is sequenced. The quality of the bioinformatic view, essential to deciphering probiotic mechanisms and functional roles, will rely heavily on continuously up-dated databases and comparative analyses.

Genetic Tools:

Over the past decade, efforts in plasmid biology and biotechnology of lactic acid bacteria have supported the development of genetic tools (e.g. transformation systems, cloning and expression vectors, integration vectors and systems for gene inactivation) in a select number of probiotic cultures, that are relatively well developed commercially or scientifically (Kullen and Klaenhammer, 1999). There remain, however, many model probiotic strains that are recalcitrant to genetic manipulation. Genetic accessibility is an important selection trait to consider for any new probiotic strains, recognizing the powerful impact that genomic information and genetic approaches will now play in establishing gene function and the mechanistic basis of functionality. For example, recombinant strategies such as the *in vivo* expression technology (IVET) and signature tagged mutagenesis (STAG) (Chiang et al., 1999) are designed to identify and investigate gene regulation and function *in vivo*. These techniques have been used extensively to study host-pathogen-host interactions, and have more recently been used to study probiotics (Bron et al., 2002) and other beneficial organisms in various habitats (Rainey et al., 1999).

Genome sequences and bioinformatics will present volumes of information for rational selection of genes for identification, confirmation, and characterization of their functional roles. Those currently presumed important for colonization, survival, and functionality include the following:

- Acid tolerance
- Bile tolerance
- Stress tolerance
- Surface proteins
- Lipoteichoic acid
- Extracellular proteins
- Exopolysaccharides
- Adherence factors
- Aggregation
- Biofilm formation
- Immunomodulation
- Putative Autoimmunity promoting factors
- Bacteriocin production
- Carbohydrate (prebiotic) utilization & metabolism
- Gene transfer potential
- Antibiotic resistance
- Putative virulence factor homologs
- Siderophores, scavengers of Fe++

- Quorum sensors and response regulators
- Prophages, prophage remnants, lysogenic conversion characters
- Mobile genetic elements

Functional genomic analyses of these properties will create opportunities to establish cause and effect relationships, but it is also expected that, global, pleiotrophic, and cascading effects will result from some gene knock outs. Redundant proteins encoded in the genome are also expected to have cumulative effects that are not resolved by a one-gene, one phenotype analysis (e.g. there may be hundreds of surface proteins that impact immunomodulation, attachment, agglutination, retention).

Probiotics as delivery vehicles for biologics

The use of lactic acid bacteria as delivery vehicles for biological compounds has been considered and actively investigated for a number of years (Wells et al., 1996; Mercener et al., 2000; Thole et al., 2000). Successful examples of metabolic engineering (Hols et al., 1999) and expression of vaccines and cytokines (Gilbert et al., 2000; Steidler, 2002) have already been reported.
Compounds targeted for possible delivery in food or *in vivo* include vaccines, enzymes, proteins, cytokines, vitamins, exopolysaccharides, and metabolites. The generally recognized as safe status (GRAS) of the lactic acid bacteria and their suitability for oral consumption at levels as high as 10^9cfu/gram makes them attractive candidates for this application in both human and animal models. Probiotic cultures may offer additional advantages for enhanced delivery of biologics to specific locations in the GI tract, mouth, vagina, or other selected tissues. Genomic information and genetic tools continue to be critically important to furthering the development of these applications, and provide opportunities such as tailored gene expression (regulated promoters; intracellular, anchored, secreted) for targeted and regulated delivery of specific biological compounds.

Host Tissue Expression

The first study on the genetic level indicated that a probiotic *Lactobacillus* could induce intestinal gene expression in the HT29 cell line that reduced binding by enteric pathogens to the intestinal cells (Mack et al., 1999). The development of high throughput DNA arrays for transcript profiling will have an enormous impact on detection of host-microbe interaction mechanisms. This approach was elegantly demonstrated in a landmark study of intestinal transcriptional responses of germfree mice with the commensal *Bacteroides thetaiotaomicron* using DNA arrays, and further used to show bacterial species-specific host responses (Hooper and Gordon, 2001; Hooper et al, 2001). DNA array technology will allow determination and parallel analyses of a large number of biomarkers or genes that are indicative of microbe-host, as well as microbe-microbe interactions in the complex intestinal ecosystem. Besides the messenger RNA, the analysis of the host and microbe proteomes, secretomes and metabolomes are expected to reveal further functionalities. (Graves and Haystead, 2002; Phelps et al., 2002).

Genetic Committee Recommendations

1. Proceed with compilation of complete genome sequencing of probiotic cultures, commensals, and unculturable flora of human mucosal habitats. Characterisation of the microbiome and metagenome of the GI tract, mouth, and vagina will be a valuable platform to assess interactions, roles, functionality, and impact of probiotic cultures.

2. Utilize genome information to build genetic tools and microarray/proteomic capabilities. Genome arrays, phylogenetic arrays, and functional metagenomic arrays will be invaluable to investigate the impact of probiotics on the existing flora, and measure the responses (survival, attenuation, gene regulation) of probiotic cultures in various environments.
3. Employ genetic information and functional genomic approaches to investigate probiotic functionality and establish causative mechanisms through which probiotics impact the microflora and host tissues. Coordinated gene expression, knockouts, and complementation of prioritized targets will facilitate discovery of the mechanisms responsible for probiotic functionality.
4. The complete genome sequence should be deciphered for all commercial probiotic cultures to ensure safety, promote functionality, establish identity and provide a reference base to assess genetic changes that may occur over time, or in changing habitats.
5. Multidisciplinary efforts investigating the behavior of probiotic cultures, *in vitro* and *in vivo*, should include isogenic derivatives for comparative effect analysis.
6. The potential for gene transfer (conjugation, transduction, transformation) from GMO's should be investigated within the "environments" of the GI Tract/Mouth/Vagina.

Genetics/Genomics have a vital role to play in the resolution of the probiotic hypothesis, investigating interactions of probiotics with microbial communities and host tissues, and revealing the mechanisms that underlie the roles and functionalities of organisms that can elicit positive affects on health and well being.

References:

Bron, P. Hoffer, S., de Vos, W.M., and Kleerebezem, M. 2002. GI-Tract specific gene expression in *Lactobacillus plantarum* WCFS1. Abstract H2. 7th Symposium on Lactic Acid Bacteria: Genetics, Metabolism, and Applications. Sept 1-5, 2002.

Chiang SL, Mekalanos JJ, Holden DW. 1999. *In vivo* genetic analysis of bacterial virulence. Annu Rev Microbiol. 53:129-154.

de Vos (2001) Advances in genomics for microbial food fermentations and safety. Curr. Opin. Biotechnol. 12:493-498.

Gilbert, C., Robinson, K., Le Page, R. W., and Wells, J. M. (2000): Heterologous expression of an immunogenic pneumococcal type 3 capsular polysaccharide in Lactococcus lactis. *Infect Immun* **68**, 3251-60.

Graves PR & Haystead TA (2002) Molecular Biologist's Guide to Proteomics. Microbiol Mol Biol Rev. 66: 39-63.

Rondon MR, August PR, Bettermann AD, Brady SF, Grossman TH, Liles MR, Loiacono KA, Lynch BA, MacNeil IA, Minor C, Tiong CL, Gilman M, Osburne MS, Clardy J, Handelsman J, Goodman RM. 2000. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Appl Environ Microbiol. 66:2541-2547.

Hawkins, T. L., Detter, J. C., and Richardson, P. M. (2002): Whole genome amplification--applications and advances. *Curr Opin Biotechnol* **13**, 65-7.

Hols, P., Kleerebezem, M., Schanck, A. N., Ferain, T., Hugenholtz, J., Delcour, J., and de Vos, W. M. (1999): Conversion of Lactococcus lactis from homolactic to homoalanine fermentation through metabolic engineering. *Nat Biotechnol* **17**, 588-92.

Hooper, L. V., and Gordon, J. I. (2001): Commensal host-bacterial relationships in the gut. *Science* **292**, 1115-8.

Hooper LV, Wong MH, Thelin A, Hansson L, Falk PG & Gordon JI (2001) Molecular analysis of commensal host-microbial relationships in the intestine. Science 291: 881-884

Kitts, C. L. (2001): Terminal restriction fragment patterns: a tool for comparing microbial communities and assessing community dynamics. *Curr Issues Intest Microbiol* **2**, 17-25.

Klaenhammer, T., E. Altermann, F. Arigoni, A. Bolotin, F. Breidt, J. Broadbent, R. Cano, S. Chaillou, J. Deutscher, M. Gasson, M. van de Guchte, J. Guzzo, A. Hartke, T. Hawkins, P. Hols, R. Hutkins, M. Kleerebezem, J. Kok, O. Kuipers, M. Lubbers, E. Maguin, L. McKay, D. Mills, A. Nauta, R. Overbeek, H. Pel, D. Pridmore, M. Saier, D. van Sinderen, A. Sorokin, J. Steele, D. O'Sullivan, W. de Vos, B. Weimer, M. Zagorec, and R. Siezen. 2002. Discovering lactic acid bacteria by genomics. Antonie van Leeuwenhoek 82: 59-71.

Kullen, M.J. and T.R. Klaenhammer. 1999. Genetic modification of intestinal lactobacilli and bifidobacteria. *In* Probiotics: a Critical Review. G. Tannock (*ed)* Horizon Scientific Press, Wymondham, U.K., pp 65-83.

Mack DR, Michail S, Wei S, McDougall L & Hollingsworth MA (1999) Probiotics inhibit enteropathogenic E. coli adherence in vitro by inducing intestinal mucin gene expression. Amer. J. Physiol. 276: G941-G950

Mercenier, A., Muller-Alouf, H., and Grangette, C. (2000): Lactic acid bacteria as live vaccines. *Curr Issues Mol Biol* **2**, 17-25.

Phelps TJ, Palumbo AV & Beliaev AS (2002) Metabolomics and microarrays for improved understanding of phenotypic characteristics controlled by both genomics and environmental constraints. Curr. Opin. Biotechnol. 13: 20-24.

Rainey PB. 1999. Adaptation of *Pseudomonas fluorescens* to the plant rhizosphere. Environ Microbiol 1:243-57.)

Rokbi B, Seguin D, Guy B, Mazarin V, Vidor E, Mion F, Cadoz M & Quentin-Millet M-J (2001) Assessment of Helicobacter pylori gene expression within mouse and human gastric mucosae by real time reverse transcriptase PCR. Infect. Immun. 69: 4759-4766

Sanders, M.E. 1999. Probiotics – scientific status summary. Food Technol. 53: 67-77.

Steidler, L. In situ delivery of cytokines by genetically engineered *Lactococcus lactis*. 2002. Antonie Van Leeuwenhoek 82: 323-331.

Tannock, G.W. 1999. Probiotics: A Critical Review. G.W. Tannock, ed. Horizon Scientific Press. Wymondham. p. 5-14.

Thole, J. E., van Dalen, P. J., Havenith, C. E., Pouwels, P. H., Seegers, J. F., Tielen, F. D., van der Zee, M. D., Zegers, N. D., and Shaw, M. (2000): Live bacterial delivery systems for development of mucosal vaccines. *Curr Opin Mol Ther* **2**, 94-9.

Wells, J. M., Robinson, K., Chamberlain, L. M., Schofield, K. M., and Le Page, R. W. (1996): Lactic acid bacteria as vaccine delivery vehicles. *Antonie Van Leeuwenhoek* **70**, 317-30.

Wells, J. M., Robinson, K., Chamberlain, L. M., Schofield, K. M., and Le Page, R. W. (1996): Lactic acid bacteria as vaccine delivery vehicles. *Antonie Van Leeuwenhoek* **70**, 317-30.

Vaughan, E.E., M.C. de Vries, E.G. Zoetendal, K. Ben-Amor, A.D.L. Akkermans, & W.M. de Vos. 2002. The intestinal LABs. Antonie van Leeuwenhoek 82: 341-352.

Zoetendal EG, Akkermans ADL & de Vos WM (1998) Temperature gradient gel electrophoresis analysis of 16S rRNA from human faecal samples reveals stable and host specific communities of active bacteria. Appl. Environ. Microbiol. 64: 3854-3859