Genome research of gut bacteria, how to analyze and how to apply

Tohru Suzuki, Kouta Sakaguchi, and Kazumasa Yasui

The United Graduate School of Agricultural Science, Gifu University, Yanagido, Gifu, 501-1193, Japan

ABSTRACT: After development of new generation DNA sequencers, the analyses of bacterial genome have been easy and familiar to microbiologist. However, to utilize vast amounts of data, it need bioinformatics and reverse-genetics, including transformation of plasmid DNA into target bacterium and gene knockout technique using homologous recombination. The authors proposed novel transformation technique to control restriction enzyme reaction using plasmid artificial modification (PAM). For the knockout gene, an error prone PCR has been employed to construct temperature sensitive plasmid. Using this plasmid the knockout technique become extremely easier, than conventional method.

INTRODUCTION

Recent innovation of DNA sequencing technology makes radical change of biological research, not only human medical but also animal, plant, microbiological science. The whole genome sequencing of bacterium is now very easily and quickly obtained using new-generation genome sequencers. In the case of human gut microbial society (so called 'microflora'), more than 200 species of symbiotic and commensal bacteria have been sequenced.

However, the species number of gut microflora has estimated over 1000. To understand, relationship between each bacterium of microflora and host animal gut intestinal tract and also between bacterium and bacterium.

The microflora of animal GIT gives many effect on host health, prevents pathogenic infection, immunological reaction, such as preventing or enhancing allergic responses, obesity, arterial sclerosis and hypertension, mental condition such as depressive symptomatology. Tessier found *Bifidobacterium* from human infant feces and its assist human health, now the Bifidobacterial species are major in human large intestine and many species are found from other animals. It is thought Bifidobacterium act many good effects on human health.

From the unrecorded history, people take many fermentation foods, such as pickles and yogurt, fermented fish, 'Sushi' etc. Some lactic acid bacteria, eg. *Lactococcus,* and fungi, eg. *Aspergillus oryzae*, are used in the key-process of fermentation foods. They have assist human nutrition and health controlling. Now, these microorganism are actively taken as probiotics.

I will show recent progress of genome biology in microbiology and its impact on animal science and production.

GENOME SEQUENCER

When the human genome project was reported by two groups, HUGO and Celera Genomics in 2004. The sequence analyses were performed using capillary electrophoresis based automated sequencer, which enable 96 electrophoresis at once, which greatly improved the sequencing speed, then s slab gel based automated DNA sequencer.

October 2005, 454 released the Genome Sequencer 20, the first next-generation sequencing system on the market. In this sequencer, Template DNA fragment are attached on Beads on which surface are coated with a Primer DNA. Then the PCR reactions are performed on the beads in oil emulsions. The amplification is proceed clonally in each emulsion. After the PCR, beads are put on micro honeycomb plate (PicoTiterPlate), sequencing chemistry is based on pyrosequencing PCR. During elongation of DNA polymerase released pyrophosphate is chaptured by sulfurylase and luciferase. The luminescence of million beads is detected by CCD camera. The newest Genome Sequencer FLX Titanium, released in October 2008, featuring 1 million reads at 400 base pairs in length.

Two other next generation sequencing systems are available now. SOLiD, provided by ABI is using immobilized PCR and repeated ligation reaction. Illumina also using immobilized PCR. Amplified templates are sequenced using a robust four-color DNA sequencing-by-synthesis technology that employs reversible terminators with removable fluorescent dyes.

These high through put DNA sequencer enabled human genome (3 x10^9 bp) re-sequencing during single running of the system. The sizes of bacterial genome are 2 x 10^6 –8 x 10^6, several hundred of bacterial genome is able to analyze using the 2nd generation genome sequencer.

Table 1. Comparison of the next generation genome sequencer

Equipment	Read Length	Run Time	Throughput
454 FLX Titanium	400 bp	10 hours	600 Mb
ABI SOLiD	2x 150 bp	~14 days	85-95 Gb
Illumina HiSeq2000	2 x100 bp	~8 days	150200 Gb

Thus, bacterial genome sequencing is now not difficult and not need big money. To obtain bacterial draft genome sequence it need only $1500 and 1 days. (However, to finish genome sequencing is much more difficult because, usually assembled data still have several hundred of unreaded regions. It needs manual amplification and sequencing to close the missing zones). It is much easy and lower cost to identify bacterial strain, than that using growing ability on selective medium set and chemical and biochemical characterization. In this point (Sep, 2010) 1371strains have been finished and 4941 strains are on going.

REVERSE GENETICS

As described above, vast amount of the genome sequence have been available today. However, much of the data has been used inefficiently in molecular biological studies since reverse genetic tools, such as convenient shuttle vectors, an efficient transformation method, gene knockout and random mutagenesis techniques, etc., have not been available. Accordingly, we have been working towards developing simple methods that would establish transformation techniques for bacteria for which the genome sequence is available.

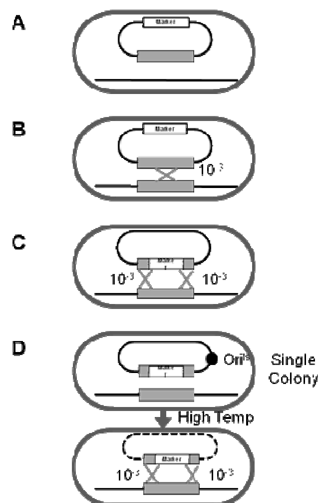

Figure 1. The reverse genetics tool for bacterial research. Transformation (A), single crossover recombination (B), double crossover recombination for gene knockout (C), double crossover recombination for gene knockout using temperature sensitive plasmid (D).

PLASMID ARTIFICIAL MODIFICATION (PAM) TO IMPROVE TRANSFORMATION EFFICIENCY

Transformation is essential to develop molecular biological techniques of each bacterium. However, restriction-modification (R-M) systems prevent efficient shuttle vector plasmid introduction into target bacterium. Recently, the whole genome DNA sequences of many bacteria have been reported. Using homology and motif analyses, possible R-M genes are able to find from genome sequence. By introducing DNA methyltransferase genes into *E.coli* cell, the plasmids will modified by these enzymes.

After propagation of the shuttle vector between *E.coli*–target bacterium in the PAM host, the plasmid will protected from the digestion by restriction enzyme of target bacterium during transformation, and gaves higher efficiency. We propose to designate this method Plasmid artificial modification (PAM).

Here, we described the method of transformation with *Bifidobacterium adolescentis* ATCC15703, using PAM method and electroporation, in which the efficiency was improved 10^5 times by introducing 2 genes encoding the modification enzymes (Fig. 2).

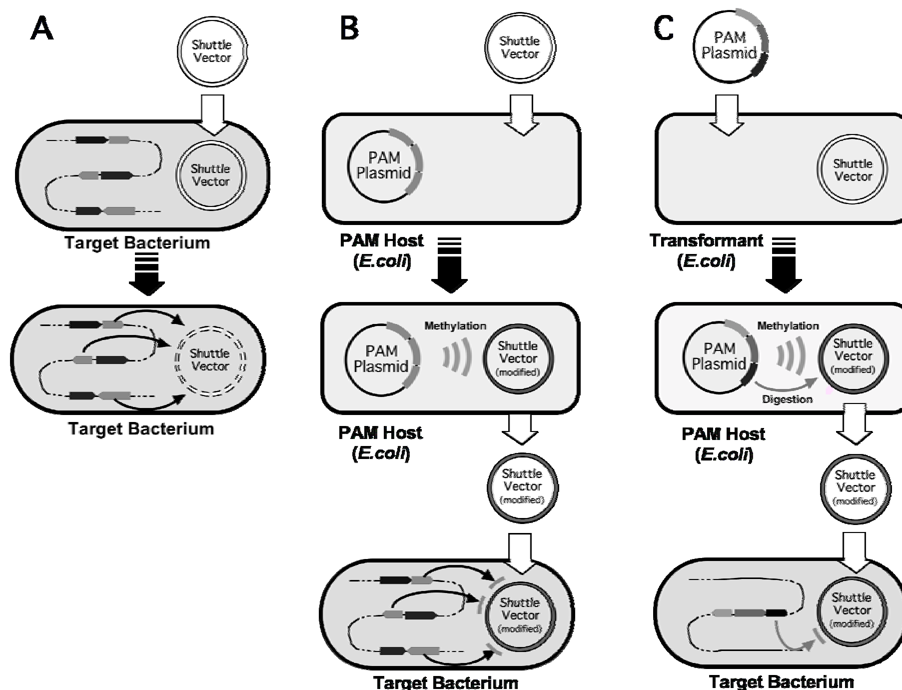

Figure 2. The PAM concept. *Panel A*: The conventional method for the transformation of bacteria.

The introduced shuttle vector is degraded by a restriction enzyme of the target bacterium. A small amount of vector survives and replicates in the target bacterium. *Panel B*: A PAM plasmid expressed by *E. coli* (the PAM host) carries all of the modification methylase genes expressed by the target bacterium. A shuttle vector plasmid is introduced into the PAM host and is methylated by the appropriate modification enzymes. The shuttle vector then is isolated and introduced into the target host by electroporation. The vector plasmid is protected from host restriction enzymes and yields a higher transformation efficiency. *Panel C:* The R-M system is a complicated structure composed of a gene cluster that may include subunits or unknown accessory genes. The PAM plasmid, containing the known modification gene(s) as well as the uncharacterized components, is introduced into an *E. coli* transformant harboring a shuttle vector. Restriction enzyme digestion occurs, but some copies of the plasmid survive in the PAM host. The plasmid is then isolated and introduced into the target bacterium. (Reproduced from ref.4 with permission from Oxford Journals).

Table 2. Comparison of Electroporation Efficiency in *Bifidobacteria* using PAM

Donor host	Recipient	Efficiency* (CFU/µg DNA)
TOP10	B.adolescentis ATCC15703	$1 - 3 \times 10^{0}$
TOP10 /pPAM1233	B.adolescentis ATCC15703	$4 - 6\times10^{4}$
TOP10 /pPAM1283	B.adolescentis ATCC15703	$1 - 2\times10^{4}$
TOP10 /pPAM1233-1283	B.adolescentis ATCC15703	$0.9 - 4\times10^{5}$
B.adolescentis ATCC15703	B.adolescentis ATCC15703	9×10^{4}
TOP10	B.longum 105-A	$1.5\times10^{6} - 5\times10^{6}$
B. longum 105-A	B.adolescentis ATCC15703	$6\times10^{3} - 8\times10^{3}$

*After electroporation, the cells were diluted ×1 or ×100 with MRS, plated on MRS-AC agar, supplemented with 150 µg/ml spectinomycin, then incubated at 37°C under anaerobic condition.

OBTAIN KNOCK-OUT MUTANT

Next to transformation, the site directed mutagenesis is important to construct site directed mutagenesis (so called 'Knock Out' technique. It has been widely applied in the molecular genetics, of Bacteria (*E. coli, Bacillus subtilis*), Plant (*Arabidopsis thaliana etc.*) and animal (mouse). Knocked out organisms gave direct evidences of physiological functions of the deleted genes. In the conventional genetics, firstly obtain some mutants, which showed specific 'phenotype' such as nutritional requirement, morphological disorder, conditional lethality or immortality. Then, perform mapping the position on the genome, cloning and sequencing. In the genome-sequenced organism, all sequence data have been already obtained, and the research strategy will completely be changed from the conventional genetics. Pick up candidates genes, which may be responsible to a specific biological phenomenon. Computer based bioinformatics, including homology analyses are used in this step. Then cloning the candidate genes and express in the other host, such as *E.coli etc*. In the case of simple enzyme, such as hydorogenase, it is enough to understand the function of the gene. If the biological event occurred by multi-gene cooperation, the 'Knock-out' technique is effective to study the mechanism. Using homologous recombination, it is possible to knock out some genes (Fig.3).

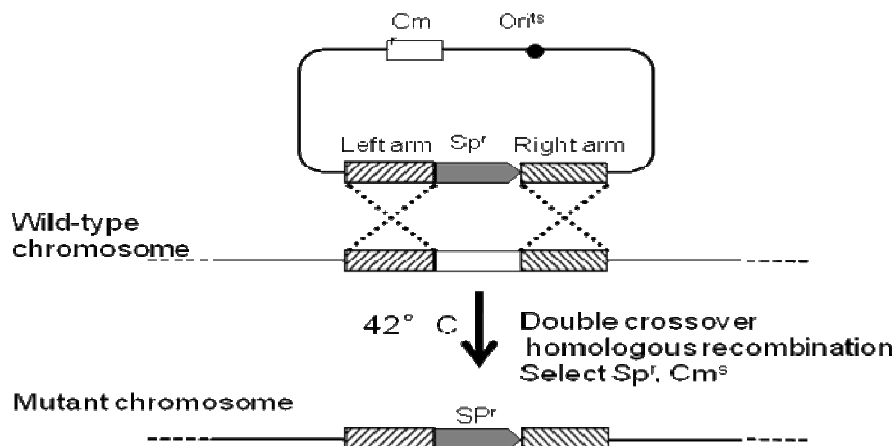

Figure 3. Scheme of homologous recombination with Ts plasmid. Homologous recombination event, screened at 42°C with Spectinomucin (Sp) and Chroramphenicol (Cm), occurred by double crossover at left arm and right arm. Clones that showed Sp^r and Cm^s phenotype were selected as candidates of gene disruptant.

LITERATURE CITED

454 Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature 2005 , 437:376-380.

McKernan KJ et al. (2009) Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two base encoding. Genome Research 19:1527–1541.

Schweizer, HP. (2008) Bacterial genetics: past achievements, present state of the field, and future challenges. BioTechniques, 44, 633–641.

Yasui K, Kano Y, Tanaka K, Watanabe K, Shimizu-Kadota M, Yoshikawa H, Suzuki T. (2009) Improvement of bacterial transformation efficiency using plasmid artificial modification. Nucl. Acids Res. 37, e3. doi: 10.1093/nar/gkn884.