Many people, especially those from the scientific community, have come to know of Beijing Genomics Institute (BGI) through its participation in the Human Genome Project (HGP). Since publishing the working draft of the indica rice genome in Science on 5 April 2002 and releasing the genomic data unconditionally, BGI or “Huada” (its Chinese equivalent) has become well-known in China and in many parts of the world where rice is a staple food.

To most people outside China, it is indeed phenomenal that BGI has come this far in such a short time (see Table 1). We hope that our experience, as described in this short essay, would be a useful reference to people of the developing world, who are seeking advances in science and technology for the betterment of their own society and economy as well as living standards.

Participation in the Human Genome Project

The Human Genome Project was proposed in 1990 by a group of scientists from the US with the aim of sequencing the human genome in 15 years and with a budget of US$3 billion. This was a very ambitious plan with immense scientific and technological challenges. Its momentum picked up in the later part of the 1990s with two technological breakthroughs.

Firstly, automatic capillary DNA sequencers were introduced in 1998 for the specific purpose of large-scale DNA sequencing required by projects of the HGP scale and with throughput that was at least one magnitude higher than any previous technologies. Secondly, advances in the IT industry saw increasing computational power and the popularization of the Internet enabled information to be stored and exchanged efficiently across the world.

The third element was that large pharmaceutical and biotechnological companies and venture capitals, foreseeing commercial gains generated by the HGP, started to invest in human genome sequencing and its related fields. This trend peaked in 1998, when the instrument company Perkin-Elmer and Craig T. Venter, one of the pioneers (though mostly in unconventional ways) in the field, established Celera Genomics. This company was set up for the specific purpose of sequencing the human genome and seizing its commercial potentials through patent seeking and discovery of drug candidates.
A Success Story of Jumpstarting Genomic Projects

With competition from the private sector, the HGP stepped up on its own project. Working drafts of the human genomes were published simultaneously in the journals *Nature* and *Science* by both the public and private sectors. Then the scenario changed, as Celera switched to become a drug discovery company. However, the HGP continued on its path towards producing a finished map of the human genome by April 2003, which is the 50th anniversary of the discovery of the double helix structure of the DNA.

Apart from the HGP and the International Human Genome Sequencing Consortium (IHGSC), the Human Genome Center (HGC) of the Institute of Genetics, Chinese Academy of Sciences came onto the scene. Part of this center later evolved into the Beijing Genomics Institute (BGI) for the specific purpose of sequencing a portion of the human genome. One percent of the human genome or the Beijing Region in the short arm of chromosome three was assigned to BGI at the Fifth Strategic Meeting of IHGSC at the Sanger Institute (then the Sanger Center) in Hinxton, UK on 1 September 1999.

No additional collaborators were accepted to the consortium thereafter. China, an emerging power in both economy and science, is the only developing country in the pact. Other participating countries include Germany, France, Japan, the UK and US, in ascending order according to their contributions to the project.

The significance of our contribution to the HGP is that it has, for the first time in modern history, propelled China onto the international stage in a major scientific event. The HGP is not only critical to the advancement of science, but it has also broken new grounds in terms of economic development, in which China will play a major role.

Since the Opium War nearly two hundred years ago, China has not been able to escape from the image of a nation in war and chaos and its people suffering constantly from poverty and diseases. China has missed many opportunities to keep up with the world or even its neighbors in the 1930s, 1950s and 1960s. Through reforms in the late 1970s and large part of the late 20th
century, a relatively relaxed political atmosphere and the encouragement of private enterprises, China has gradually recovered and picked up steam in the area of economic development.

One of the most notable trends in the past twenty years is that more than half a million students and scholars have been trained in the Western developed nations, mainly in the US, UK, Japan, France, Germany, Australia and Denmark. A large number of them have since returned to China. In addition to advanced knowledge and skills in science and technology, they have also brought back new ideas. The participation of HGP is the first sign that China, as one of the largest nations and the most populous one in the world, will contribute a major part to human civilization and world development, as it has always been in the past several thousand years.

Launching the Rice Genome Project

After accomplishing the sequencing task of the HGP, we were faced with a dilemma as to where we should head next. It appeared that continuation on the genomics path was an obvious choice. At that time, we had about two dozen DNA sequencers and not so adequately equipped computers. Therefore, the challenges were about setting goals and upgrading equipment. We are obliged to do research that has direct impact on the economic development of the nation. Any research we do as far as genomics is concerned has to be fundamental and have direct impact on applications.

Figure 2. Chinese-made supercomputers

To choose rice as our next target has its practical, logical and scientific reasoning. Its size of around 400 Mb is about one magnitude larger than the Beijing Region. It is a good model organism to the study of other plants, and any information we learn from the project has the potential of being applied to rice breeding. We also decided that we would sequence a hybrid cultivar, different from the direction of the International Rice Genome Sequencing Project (IRGSP) consortium that started their project three years earlier.

Therefore, what we needed was just more sequencers, more powerful computers and securing funds. We did not know at that time that private companies had already started to eye the rice genome. Their announcements, coming a few months after we announced our hybrid rice genome project in May 2000, changed the scenario drastically.

Hence, we took a rather different approach to the project. Instead of working on a single species, sequencing commenced from two parental strains of the hybrid rice cultivar that occupies the largest hybrid rice plantation in the country. The idea was to use this cross breed to reveal the mechanism of heterosis potential, which was the basis of high yield and high quality of the hybrid.

We set three goals:

1. Study rice as a model organism to monocots and plants in general. Later, analysis of the genome revealed that the organization and features of the rice genome is quite different from any other genomes known. The rice genome has twice as many genes than Arabidopsis, a dicot plant; it also has many features that do not exist in other genomes.

2. The indica rice genome, which is a more popular subspecies of rice and cultivated mainly in China and many Southeast Asian countries, would be first sequenced from our genomics pipeline. The information would then be compared to the maternal strain and the japonica rice genome, which was sequenced by private companies and the IRGSP consortium.

3. Comparison of the two species which make the hybrid rice cultivar yield rich information about the secret of heterosis and many other elite traits which make hybrid rice the first choice of rice farmers.

Rice is food to more than half of the world population, hence any information we gain through its study should benefit rice farmers. The study of rice genomics, especially, could significantly change the next era of rice breeding and food security. However, the emergence of private companies onto the arena has changed the playground.

Companies, with their legitimate claim of return to their investment, should be allowed to apply for patents and withhold information. However, it is hard to determine if this will be to the best interests of most rice farmers in the Third World countries. We were
concerned when the private companies announced that
they had finished the sequencing of the *japonica* rice
genome. Instead of withholding genomics information
to the benefit of one single institution or one single
nation, we chose to release the information of our rice
genome data to the public domain. We believed that
the faster and wider the information disseminates, the
more advanced rice research will be and more support
will be garnered from nations and people who would
benefit from the information.

BGI has recently announced that the rice genome
gene expression microarray which covered most rice
genes will be made available to research laboratories
all over the world. We hope that this will accelerate the
progress of rice research for better quality and yield.

**International collaboration**

Almost all of the large-scale genome projects are
carried out through multi-national and multi-institutional
collaborations, largely due to the nature of these
projects. From *E. coli*, yeast, *C. elegans*, fruit fly,
*Arabidopsis*, to human, rice and more recently fish,
mosquito, and malaria genome projects, all these are
done with scientists from many countries.

Sequencing, assembly and annotation may be
processed in one big genomics center, but the results
and their consequences will be shared by everyone and
accessible to anyone who needs it. The *indica* rice
genome project was mainly done in China, but
 collaborators came from across the nation and some of
the co-authors were working from abroad. China is also
one of main partners of the IRGSP working on the
*japonica* rice genome.

International collaborations are vital to small
nations who have less economic prowess to access
advanced knowledge like genomics. It will also shorten
their distance from the economic superpowers or at least
help them keep the distance from widening.
Participation is a good way to jump start genomics
projects. One of the greatest achievements from our
participation in the HGP is its effects on public education
of human genomics and its relevance to the everyday
life of ordinary people.

The recently announced Human HapMap Project
is also an international collaboration. We are also
working with scientists from Denmark on the Pig
Genome Project (PiGP) and will continue to seek other
international collaborations. In these collaborations,
resources, expertise, technology, funding, infrastructure,
and manpower from all the partners have helped to bring
the projects quickly to fruition.

Geographical distributions of species and
agricultural adaptations of crops are localized due to
climate, soil, culture and migration of population.
Industrialized countries will select their study subjects
according to scientific merits and its relevance to the
local economy. These elements will be the main
determinants in selecting a genomics program.

On the agricultural front, China and Japan are going
to finish the rice genomes by the end of this year, China,
Denmark and the US are researching into pig genome,
the US, UK and China into chicken genome, the US
into maize and cattle genomes, and Australia is into
sheep genome. As for the BGI, we will lead a study into
those species which have ultimate importance and not
on the high priority lists of other major genomics centers
(see Yu J. et al., An International Campaign for
Agricultural and Livestock Genomics (CALG), on page
957 of this special issue).

**Conclusion**

Common perceptions are that genomics is an
expensive business. This might be true to some extent,
especially at the early stage of the Human Genome
Project, where the cost of sequencing a base pair was
estimated to be about US$1 dollar.

Now, the cost has dropped to about a few US cents
a base pair at most of the large genomics centers. To
sequence a genome with a size of five million base pairs,
it costs less than half a million US dollars and takes
about a few days instead of months, though the
annotation of a finished genome will take a longer time.
For microbes, the job should be considerably easier,
since there are close to one hundred such genomes with
complete sequences. At times, solution to one particular
question could come out much easier from the other
side of the coin, i.e., pest resistance of crops to
pathogenic microbes is hundred thousand smaller in
size and complexity to deal with, and it could be as
effective.

Streamlining of procedures and more automation
will decrease the cost of genomics projects even further.
By operating in developing countries like China and
taking advantage of favorable local government policies
towards innovation of science and technology, the
overall cost of genome sequencing would be further
reduced (it is no coincidence that all three major
genomics centers in China are all established in
designated industrial development zones).
Genomics will help us to close the knowledge gap between genome and biology. It has become a discipline of its own. When setting up a genomics program, one has to consider more on its scientific merits, rather than the cost of sequencing or whether it will get enough attention or be funded.

We think that the future will be genome-centric, as genomics is essential to the understanding of health and diseases, the improvement of living standards, evolution, and how to create an ecological sound environment.

<table>
<thead>
<tr>
<th>Time</th>
<th>Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998.09</td>
<td>Establishment of the Human Genome Center in the Institute of Genetics, Chinese Academy of Sciences.</td>
</tr>
<tr>
<td>1999.07</td>
<td>Establishment of the Beijing Genomics Institute (BGI) as an independent, not-for-profit research institution.</td>
</tr>
<tr>
<td>1999.09</td>
<td>Assignment of one percent of the region of human chromosome 3 that is later called the “Beijing Region”</td>
</tr>
<tr>
<td>2000.03</td>
<td>Accomplishment of the sequencing of 37 Mb of the “Beijing Region”.</td>
</tr>
<tr>
<td>2000.05</td>
<td>Announcement of Super Hybrid Rice Genome Project.</td>
</tr>
<tr>
<td>2000.10</td>
<td>Announcement of Sino-Danish Pig Genome Project.</td>
</tr>
<tr>
<td>2001.02</td>
<td>Publication of the human genome working draft in the journal Nature with groups from the US, UK, Japan, France and Germany.</td>
</tr>
<tr>
<td>2001.08</td>
<td>Announcement of the completion of the “Beijing Region” with “finished” quality.</td>
</tr>
<tr>
<td>2001.12</td>
<td>Unconditional release of the indica rice genome data to the public domain.</td>
</tr>
<tr>
<td>2002.04</td>
<td>Publication of the working draft of the indica rice genome in the journal Science.</td>
</tr>
<tr>
<td>2002.05</td>
<td>Publication of the genome of Thermoanaerobacter tengcongensis, the first microbial genome sequence published by Chinese scientists.</td>
</tr>
<tr>
<td>2002.09</td>
<td>Announcement of the rice gene expression microarray and version two of the working draft sequence of the indica rice genome.</td>
</tr>
<tr>
<td>2002.10</td>
<td>Announcement of the initiation of Human HapMap Project as a founding member of the International Research Consortium with groups from the US, UK, Japan and Canada.</td>
</tr>
</tbody>
</table>

Table 1. Chronicles of the Beijing Genomics Institute.

Someone defined nostalgia as a yearning to pay bygone prices on a modern salary. If so, then we in publicly funded rice research are experiencing a sort of reverse nostalgia, as we hanker for the healthy budgets of yesterday to realize today’s exciting new research opportunities.

Our reverse nostalgia is all about building the future, not reliving the past. We cannot turn back the clock to the glory years of the Green Revolution of the 1970s and 1980s — and who would want to? Recent advances in molecular biology and genomics are just now providing the powerful tools with which we can leverage past successes to root out the ancient scourges of hunger, malnutrition and poverty. The irony is that this crucial moment is when many donors’ longstanding commitment to publicly funded agricultural research appears to be faltering.

Sequencing of the rice genome

Central to these recent research advances is the sequencing of the genome, or genetic makeup, of the two main subspecies of rice. The Beijing Genomics Institute and the Swiss company Syngenta published these milestone achievements in April in the journal Science. And there is more to come. Before the end of this year, the International Rice Genome Sequencing Project (IRGSP), a long-running, publicly funded effort coordinated by the Japan Rice Genome Program, will complete a more detailed and accurate sequence, with an error rate of less than 0.01 percent.

The sequencing of the rice genome puts rice in a unique position among food crops. The unparalleled importance of rice as a food crop for human consumption, especially for the poor, means that the resulting improvements in rice production will have tremendous impact. Making rice more affordable and nutritious will improve the health of poor consumers,