FINAL

MEMORANDUM OF UNDERSTANDING
BETWEEN
CELERA GENOMICS
AND
REGENTS OF THE UNIVERSITY OF CALIFORNIA, BERKELEY
ON BEHALF OF THE
BERKELEY *DROSOPHILA* GENOME PROJECT GROUP
CONCERNING
RESEARCH ON DETERMINING THE SEQUENCE AND ANNOTATION OF THE REFERENCE *DROSOPHILA* GENOME

ARTICLE I

Objectives of the memorandum

As previously announced*, in order to produce a complete, annotated, and publicly accessible sequence of the *Drosophila* genome at reduced costs and at an accelerated pace, Celera Genomics and the Berkeley *Drosophila* Genome Project Group (BDGP) (hereafter referred to as the Parties) have agreed to cooperate as defined in this Memorandum of Understanding. The BDGP is a consortium of research groups working at the University of California at Berkeley, Lawrence Berkeley National Laboratory, Baylor College of Medicine and Carnegie Institution of Washington and funded by the National Institutes of Health, the Department of Energy and the Howard Hughes Medical Institute.

The purpose of this Memorandum is to establish a framework for the collaboration and the exchange of non-confidential technology and information (sequence data exchanged will be limited to data already made available to the public by either party). The parties believe that by bringing their extensive experience and resources to bear on the finishing and initial annotation of the euchromatic portion of the *Drosophila* genome, it is reasonable to expect to substantially complete these tasks within calendar year 1999. The early availability of the *Drosophila* genome is a meritorious goal. *Drosophila* has played a key role in biological research for over 80 years and among the model organisms, *Drosophila* is particularly well-suited for the role of adding value to the human sequence. In terms of DNA sequence similarity as well as morphological, physiological, and behavioral complexity *Drosophila* is by far the closest to humans of the genetically tractable, invertebrate model organisms. Moreover, the large *Drosophila* research community has provided a wealth of information and understanding unequaled in its depth and intellectual breadth. For these reasons the Parties believe that the annotated sequence of the *Drosophila* genome will provide an invaluable key to understanding the sequence of the human genome and human biology. In addition, it will provide information about insect genomes that will have great utility in agricultural research.

---

* Shotgun Sequencing of the Human Genome ,J. Craig Venter, Mark D. Adams, Granger G. Sutton, Anthony R. Kerlavage, Hamilton O. Smith, and Michael Hunkapiller, *Science* 1998 June 5; 280: 1540-1542.

The parties believe their collaboration is more than a test of the scientific efficacy of the random shotgun sequencing approach to sequencing of the human genome. It is also a test of whether a public/private collaboration can expedite and lower the cost of the generation and use of genomic information that will ultimately benefit agriculture and the prevention, diagnosis, and treatment of important human illnesses.

This Memorandum describes changes in the BDGP and Celera programs that will be made to facilitate the scientific collaboration outlined in this document. BDGP will modify its strategy to sequence the *Drosophila* genome by producing a larger number of Bacterial Artificial Chromosome (BAC) end sequences that will provide a necessary component of the whole genome shotgun strategy proposed by Celera Genomics. During the first few months of this collaboration, BDGP will also produce unfinished sequence from an increased number of BAC clones. Celera Genomics will modify its previous plans to sequence the *Drosophila* genome to include BAC end sequences produced by BDGP. Both parties will work together to fill gaps and resolve ambiguities in the sequence that results from the whole genome shotgun assembly to produce finished sequence and will collaborate in the initial annotation of this sequence. The goal will be that the finished sequence will be of the highest scientific standard of quality as demonstrated in other privately or publicly funded genome sequencing projects.

The BDGP recognizes the enormous contribution that will be made by Celera Genomics, a for-profit company, in providing the results of their whole genome shotgun sequencing of the *Drosophila* genome to this collaboration and by making these data publicly available by deposition in GenBank. Celera Genomics recognizes the important contributions of the BDGP, and other academic groups, in producing a physical map of the genome and by having already determined approximately 20% of its sequence. Both parties agree that participation in this collaboration by other members of the international community doing research with *Drosophila* would be beneficial to all. Therefore, the parties resolve to encourage the participation by other *Drosophila* genome research programs in this effort.

ARTICLE II

Scope and Depth of Cooperation

Possible areas of cooperation may include but are not limited to those technical topics of DNA sequencing, mapping, resource development, genomic analysis and annotation described below. While this Memorandum does not make any funding commitments or constitute a legal obligation, by either party, to perform specific tasks, both parties agree to put forward their best efforts to accomplish the listed tasks.

BDGP researchers and their collaborators in Europe will independently accomplish the following tasks that are of interest to Celera:

- Generate paired BAC (or P1) end sequences from *Drosophila* DNA that would provide an average density of paired ends of 1 per 5kb (a total of 12,500 BACS and P1s for which both end sequences were successful). These sequences will be available to Celera Genomics by April 1, 1999.

- Redirect the majority of their genomic sequencing effort to the generation of low pass sequence (between 0.5 and 3X coverage) from BACs throughout the genome until contigs from the whole genome shotgun process are available for finishing and then focus their sequencing efforts on finishing.

- Construct an ordered physical map using these BACs which will assist Celera Genomics in the ordering and orientation of contigs assembled from shotgun sequencing and BAC end sequences.

- Fingerprint BACs using single-enzyme digestion to provide a resource for evaluating the accuracy of the sequence assembly.

- Fill gaps and resolve ambiguities to produce a final finished sequence using contigs placed in GenBank by Celera Genomics.

Celera Genomics, at no cost to the federal government, will independently accomplish the following tasks that are of interest to BDGP:

- Conduct whole-genome shotgun random sequencing of the *Drosophila* genome that will result in 10X coverage to be completed, depending on the start date, by about July 1, 1999.

- As previously committed by Celera Genomics and because it is a specific condition of the federal grant award to the BDGP that, "Sequence data supplied by others will be publicly available before being used for activities in this grant," Celera Genomics will begin depositing contigs, as unfinished sequence, in GenBank upon the completion of the random sequencing phase. The rate of deposits to GenBank will, at a minimum, be sufficient to allow BDGP to perform sequence finishing and all contigs greater than 2 kb will be deposited by January 1, 2000 at the latest.

BDGP and Celera will accomplish the following tasks as they relate to the Drosophila genome:

- As the sequences are improved through finishing efforts, their entries will be updated in GenBank so that BDGP can remain in compliance with the specific condition of their grant award that they, "adjust their data release policy to be the same as the National Human Genome Research Institute has adopted for their human sequencing grants." The Parties agree to exchange all primary electropherograms for sequence data relevant to the collaboration.

- Validate, evaluate, and correct as well as compare, study, and annotate sequence data produced under this Memorandum.

- Jointly and concurrently publish at least two papers. One paper will describe the consensus DNA sequence of *Drosophila* and the other will provide a comparative analysis of the *Drosophila* and other genomes. Related papers dealing with other aspects of the work, such as construction of the physical map, development of software for sequence assembly, etc., may also be published. Authorship on all publications will be based on relative contributions to the work.

The scope of cooperation may be expanded by amendment to this Memorandum pursuant to ARTICLE VI.

ARTICLE III

Funding

The activities carried out under this Memorandum will be subject to and dependent on funds and manpower available to each party. Each party will bear the costs of its own participation in cooperative activities under this Memorandum. This Memorandum shall not be used to obligate or commit funds or as the basis for the transfer of funds.

ARTICLE IV

General Provisions

Cooperation under this Memorandum shall be in accordance with the laws, regulations and or bylaws under which each party operates.
Pre-existing and newly developed intellectual property shall remain in the ownership of the discovering party. As this collaboration is limited by the restriction that BDGP exchange only nonproprietary technical information and only sequence data already available to the public any new intellectual property will necessarily remain that of the respective parties.

ARTICLE V

Responsibilities

Both parties will designate an individual to function as the point of contact to resolve particular problems or policy matters pertaining to this Memorandum. The parties have designated the following individual to serve this purpose:

BDGP
Gerald M. Rubin
John D. MacArthur Professor of Genetics
Investigator, Howard Hughes Medical Institute
University of California, Berkeley
Berkeley, CA 94720-3200

Celera Genomics
J. Craig Venter, President
Celera Genomics
45 West Gude Drive
Rockville, MD 20850

ARTICLE VI

Entry into Force, Amendments, and Termination

This Memorandum shall enter into force upon the later date of the signature by both parties and remain in force for two (2) years from the effective date, unless extended by agreement. This Memorandum may be amended at any time by mutual written agreement of the parties.

This Memorandum may be terminated by either party upon thirty (30) days written notice of termination to the other party.

Approved and Accepted

For Celera Genomics	For the Regents of The University of California, Berkeley	For the BDGP
By:_____	By:_____	By:_____
J. Craig Venter	Marion Lentz	Gerald Rubin
Title: President	Title: Manager, University/ Industry Research	Title: Professor of Genetics
_____	_____	_____
date	date	date