

# SEQUENCING THE COTTON GENOMES: THE NEXT STEP

**A report on the outcomes of an open forum on  
cotton genome sequencing held Saturday 13  
January 2007 (6:30 to 9:30 PM) at the XV Plant &  
Animal Genome Conference (San Diego  
California, USA)**

---

## **Background**

During the September 18-20, 2006 ICGI Research Conference in Brazil there was considerable discussion regarding optimum strategies for sequencing the cotton genomes.

While no clear consensus emerged regarding an optimum strategy, a small group volunteered to develop, on an *ad hoc* basis, a white paper that would collate the background and technical details regarding cotton genome sequencing in a single document.

The *ad hoc* "Cotton Genome Sequencing White Paper Group", led by Jeffrey Chen (University of Texas) comprised Xiaoya Chen, Elizabeth Dennis, Andrew H. Paterson, Brian E. Scheffler, David A. Stelly, Christopher D. Town, and Tianzhen Zhang. Approximately 50 "Consulting Writing Members", representing 11 countries, received early drafts of the White Paper for comment and consideration (Appendix I).

The "Cotton Genome Sequencing White Paper Group" in collaboration with Bayer CropScience organized an evening session during the XV Plant and Animal Genome Conference to present the initial draft of the white paper and to discuss the implications of the accumulated information.

Professor Joe Ecker (Salk Institute), one of the pioneers of the *Arabidopsis* genome sequencing effort and a reviewer of the USA Department of Energy (DOE) genome sequencing proposals, was invited to moderate the discussion.

---

---

## Outcomes

The meeting was well attended (~50 interested participants), filling the entire meeting room. Participants included stakeholders from Academia, US government Research institutes, Industry, and non-cotton Scientists, including international representatives.

The agenda for the evening (Appendix II) focused on two presentations: (1) an industry perspective of cotton genome sequencing (Appendix III), and (2) a report of the White Paper: Cotton Genome Sequencing (<http://algodon.tamu.edu/sequencing>).

Professor Joe Ecker oversaw a lively discussion that covered the following topics:

- What tools are available for genome sequencing?
- What tools are needed for cotton genome sequencing?
- What strategies have been taken for sequencing repeated regions in other crops?
- How will the repeated regions in the cotton diploids and polyploids impact on sequencing?
- Which cotton genomes should be sequenced, and which one should be first?
- What is a practical and economical course for cotton genome sequencing?
- Do we need new knowledge or tools to enable cotton genome sequencing?
- What are the funding models to consider for cotton sequencing initiatives?
- How do we engage with national or international networks?

At the conclusion of the meeting, Professor Ecker called for a non-binding resolution to summarize the consensus of the discussion participants, noting that it was only representative of the discussion participants not of the entire cotton community.

---

---

The following were determined to be the consensus of the meeting participants: 45 yes votes and 5 abstentions.

1. There is a clear desire among all participants to sequence the cotton genomes (as opposed to alternate strategies such as reduced representation sequencing).
2. Four short term research/funding efforts were to be encouraged.
  - a. Sequencing the *G. raimondii* genome as a stepping stone to the other cotton genomes
  - b. Developing the genetic and genomic resources needed to sequence the *G. hirsutum* sequence
  - c. Sequencing homoeologous BACS from *G. hirsutum* ( $A_t$  and  $D_t$  subgenomes), a few representative diploid genomes and *Gossypioides kirkii*.
  - d. Preliminary sequencing of other diploid genomes
3. In the longer term, *i.e.*, as the requisite genetic and genomic resources become available, the following were identified as key goals.
  - a. Complete sequencing of the *G. hirsutum* genome
  - b. Complete re-sequencing of related *Gossypium* genomes, *e.g.*, A, B, C, D, E, F, and K genome representatives and *Gossypioides kirkii*.

A number of working strategies and issues were discussed in light of the considerations above.

1. Three organization and funding possibilities were considered.
  - a. *G. raimondii* continue under the auspices of the DOE-JGI, where a pilot project is underway
  - b. Other federal agencies should be approached for the other genome sequences.
  - c. There may be some opportunity for International initiatives.
2. A plan for the dissemination of the genome sequence information (the bioinformatics platform) needs to be in place very quickly, as preliminary sequence information is emerging quickly. This will require a public domain "facility" easily accessible to all researchers.
3. The next steps in the process include
  - a. Summary of the outcomes of this meeting made publicly available.
  - b. White Paper revised as needed.
  - c. Offer by Chinese colleagues to host Sequencing Workshop evaluated.
  - d. Applications for funds for next Sequencing Workshop developed and submitted to NSF/USDA and potential Commercial partners.

## White Paper Consulting Writing Members

### Australia

Danny Llewellyn  
<Danny.Llewellyn@csiro.au>

### Belgium

Tony Arioli  
<tony.arioli@bayercropscience.com>  
Curt Brubaker  
<curt.brubaker@bayercropscience.com>

### Brazil

Edina Moresco <edina@facual.org.br>

### China

Wangzhen Guo <moelab@njau.edu.cn>  
Zhiying Ma <mzhy@hebau.edu.cn>  
Gui-Xian Xia <guixianx@yahoo.com>  
Shuxun Yu <yu@cricaas.com.cn>  
Xianglong Zhang  
<xlzhang@mail.hzau.edu.cn>  
Jun Zhu <jzhu@zju.edu.cn>  
Shuijin Zhu <shjzhu@zju.edu.cn>  
Yuxian Zhu <zhuyx@water.pku.edu.cn>

### Egypt

Essam Zaki <eazaki@link.net>

### France

Jean-Marc Lacape <marc.lacape@cirad.fr>

### India

Ishwarappa Katageri  
<ikatageri@yahoo.com>  
B. M. Khadi <cicrngp@rediffmail.com>  
Vidyasagar Parchuri <pvsagar@sify.com>  
Vijay N. Waghmare <vijayvnm@yahoo.com>

### Israel

Yehoshua Saranga  
<saranga@agri.huji.ac.il>

### Pakistan

Mehboob Rahman  
<mehboob\_pbd@yahoo.com>

### USA

John. E. Bowers <jebowers@uga.edu>  
Roy Cantrell <RCantrell@cottoninc.com>  
Peng Chee <pwchee@uga.edu>  
Leon S. Dure III <ldure@uga.edu>  
Deqiu Fang  
<Deqiu.Fang@deltaandpine.com>  
Alan R. Gingle <agingle@uga.edu>  
Candace Haigler  
<candace\_haigler@ncsu.edu>  
Ramesh Kantety  
<ramesh.kantety@aamu.edu>  
Russell J. Kohel <kohelrj@neo.tamu.edu>  
Siva Kumpatla <spkumpatla@dow.com>  
Gerald O. Myers  
<GMyers@agcenter.lsu.edu>  
Richard Percy <RPercy@uswcl.ars.ag.gov>  
Daniel G. Peterson  
<dpeterson@pss.msstate.edu>  
Keerti Rathore <rathore@tamu.edu>  
Umesh Reddy <ureddy@wvstateu.edu>  
Jun-kang Rong <jkron@yaho.com>  
Sukumar Saha <ssaha@msa-  
msstate.ars.usda.gov>  
J. McD. Stewart <jstewart@uark.edu>  
Barbara Triplett  
<btripler@src.ars.usda.gov>  
Mauricio Ulloa <mulloa@pw.ars.usda.gov>  
Allen Van Deynze  
<avandeynze@ucdavis.edu>  
Jonathan Wendel <jfw@iastate.edu>  
Thea Wilkins <thea.wilkins@ttu.edu>  
Robert J. Wright <robert.wright@ttu.edu>  
Jinhua Xiao <jinhua.xiao@monsanto.com>  
John Yu <zyu@qutun.tamu.edu>  
Hongbin Zhang <hbz7049@tamu.edu>

### Uzbekistan

Abdusattor Abdugarimov  
<genomics@uzsci.net>  
Ibrokhim Abdurakhmonov  
<ibr58@hotmail.com>

**Agenda for White Paper Presentation at the XV<sup>th</sup>  
PAG Conference 13 January 2007**

**SEQUENCING THE COTTON GENOMES**

**Saturday 13 January 2007; 6:30 to 9:30 PM  
Pacific Salon 3 XV PAG Conference**

The *Cotton Genome Sequencing White Paper Group* invite you to an open forum designed to foster the growing coalition of cotton researchers interested in sequencing the cotton genomes.

The session will be moderated by Dr. Joseph R Ecker (Salk Institute for Biological Studies)

**Meeting Agenda**

Welcome ..... *B. Lambert (Bayer CropScience)*  
Introductory Comments ..... *J. R. Ecker (Salk Institute)*  
Industry Perspective..... *C. Brubaker (Bayer CropScience)*  
White Paper Presentation..... *J. Z. Chen (U. of Texas)*  
Moderated Discussion..... *J. R. Ecker (Salk Institute)*  
Conclusions ..... *J. R. Ecker (Salk Institute)*

**Light Refreshments will be provided.**

**Sponsored by**





## Developing a public-private coalition to sequence the cotton genomes

Presentation • 19 February 2007 • Slide 2

### Why sequence the cotton genomes?

#### Public Researchers

- ♦ A pivotal research tool that will accelerate the pace and scope of scientific discovery
  - Increases our understanding of cotton physiology and evolution, e.g.,
    - ♦ Model of polyploid and comparative genome evolution
    - ♦ Fiber cells are an important experimental system for understanding cell wall development

#### Private Industry

- ♦ Strengthens cotton's position as the leading fiber crop through innovation
  - Maintains the competitive advantage of cotton fiber relative to other fibers
  - Creates new values for cotton on the farm and beyond the farm gate

PAG 2007 Bayer • 19 February 2007 • Slide 3



### Why would commercial companies support public domain cotton genome sequencing?

Sequence data are a . . .

- ♦ Pre-competitive tool for discovery.
  - Public domain genome sequence data will enable commercial companies to create and protect value (novel traits and proprietary germplasm)
  - The pace of genome sequence based innovation is proportional to the number of people using it, hence it must be public
- ♦ Pivotal application tool for translating scientific discoveries into commercial outcomes
  - Accelerates the development of new "products"
  - It allows commercial companies to differentiate themselves

PAG 2007 Bayer • 19 February 2007 • Slide 3

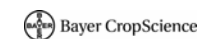


### How will cotton genome sequencing be realized?

#### A Public-Private Coalition

- ♦ The public sector must unite behind a pragmatic strategy, however
  - A single cotton genome sequencing proposal may not be the best strategy
  - A few coordinated & complementary proposals allow a diversity of funding agencies to be approached
- ♦ The private sector will act as a catalyst by providing . . .
  - A forum to bring together Researchers with useful experience and skills
  - Resources for high quality proposals
  - Evidence that the research outcomes are economically relevant
- ♦ Real commitment from the private sector will . . .
  - Maintain the current momentum in the public sector
  - Signal to funding agencies that the Cotton Industry will co-invest in cotton genome sequencing

PAG 2007 Bayer • 19 February 2007 • Slide 4



### What type of proposals will the Industry support?

Criteria for Scientific Partners:

- ♦ Technical capability and experience
- ♦ Intellectual credibility
- ♦ History of constructive collaboration
- ♦ Institutional capability

Criteria for Strategy & Outcomes

- ♦ The most cost-effective approach to obtaining AD sequence data
  - Time to delivery
  - Value for money
  - Applicability to all stakeholders
  - Technical feasibility

*The decreasing cost of re-sequencing (Moore's Law) makes the selection of a genome to sequence less critical, as long as something is sequenced*

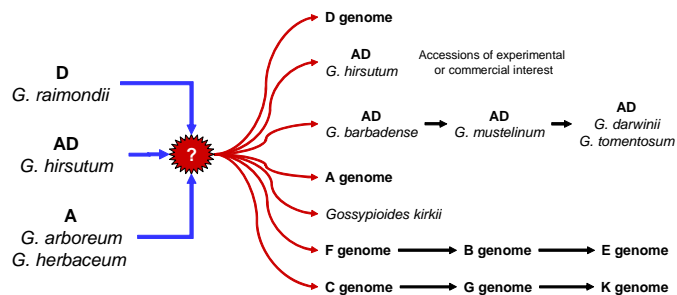
Excluded Criteria

- ♦ Popularity of cultivar or species
- ♦ Cultivated status

### The first sequence is the gateway to rest of the genomes via re-sequencing

Stage One  
Primary Sequence

Stage Two  
Re-sequencing



### What is the Industry going to do next?

- ♦ Continue to facilitate and catalyze the process
- ♦ Provide resources to make useful collaborative projects and proposals successful
- ♦ Work with each other to develop a coordinated private sector support strategy

