Commercial Gene Synthesis

John Mulligan

BLUEHERON® BIOTECHNOLOGY

NSABB, July 13, 2006

Topics

- Commercial gene synthesis today
- Issues and technology for the future
- Current screening practices

Access to DNA is Central to Modern Biology

- Biomedical Research
- Biology
- Agriculture
- New areas such as Synthetic Biology

Acquiring and Modifying DNA is Costly

- **Researchers spend $300 to $500 million a year on reagents to clone and modify genes**
- **Every $1 spent on reagents represents an additional $3 to $5 of fully loaded costs**
  - Labor, overhead, facilities, etc.
- **Fully-loaded costs of $1 billion or more annually**

Commercial Gene Synthesis

- **A potential substitute for $1 billion in costs**
- **The current market is $20 to $30 million a year**
  - Revenues rowing at 30% to 50% a year
    - Volume growth much higher
  - Highly fragmented: 50 or more companies in this area world wide
- **Still a tiny fraction of the overall molecular biology market**
  - We expect it to grow rapidly but to take 5-10 years to reach a significant fraction of the molecular biology market
  - Demand will drive rapid improvement of the technology

Gene Synthesis Technology

- **In use since the late ’70’s but only beginning to be widely used**
- **Challenges**
  - Error rate: ~1/300
  - Mismatched hybridization can lead to scrambled order
  - Reliability impacts speed and cost
- **Three general approaches**
  - Standard PCR synthesis
  - Array-based PCR synthesis
  - Solid phase assembly

PCR-based Gene Synthesis

- **Synthesize an overlapping set of oligonucleotides that cover the desired sequence**
  - Full coverage of both strands or partial
- **Pool the oligos and PCR amplify**
  - Some protocols start with a ligation step
  - Most use a secondary amplification with outside primers
- **Clone into a plasmid vector**
- **Sequence and choose a correct clone**
- **Assemble larger fragments by fusion PCR**

PCR-based Gene Synthesis

- **Most commercial synthesis and essentially all synthesis in individual labs is based on PCR**
- **Many published protocols**
  - Most work on a subset of genes
- **A substantial fraction of natural sequences require or benefit from other approaches**
  - Not ”PCR-able”
  - High GC, repetitive, very long, etc.

Array-Based PCR Synthesis

- **Start with a pool of oligonucleotides synthesized on an array**
  - Church, Gao, et al. (2004)
- **Could be a very cheap source of oligos: 4,000 to 800,000 oligos for $500 to $1,000**
- **Currently limited by the quality of array-based oligo synthesis**
- **Codon Devices is commercializing this technology**

Solid-Phase Gene Synthesis

- **Conceptually similar to oligonucleotide synthesis**
  - Monomers = duplex DNA fragments
  - Add molar excess to drive the reaction
  - Wash away failures and side reactions at each step
- **Works on almost any sequence**
- **Method used at Blue Heron**
  - Developed under an NIGMS-funded SBIR grant
  - ~5 megabases synthesized

Error Removal

- Raw oligonucleotides have an error rate after cloning of 1/20 to 1/500 base pairs, depending on the source
- Economical synthesis depends on reduction in error rates
- Most commercial groups use a proprietary error removal technology to improve reliability and reduce costs

Complex Manufacturing Process

- **Every order is different**
  - Every gene is made from a dozen to several thousand parts
- **Every part is new and used for only one order**
- **The smallest parts are chemicals**
  - Mixed populations of good and bad parts
  - Error rate of on in a few hundred
- **Larger parts are biological**
  - Unpredictable behavior
- **The final product must be perfect**

Existing Manufacturing Tools are Inadequate

- **Commodity market**
  - Prices drop 30% to 50% / year
  - Must drop production costs at least this fast
- **Mass customization used in some industries**
  - Have not found one where every part is new
- **Handling high failure rates is critical to controlling manufacturing costs**
- **Existing tools focused on assembly-line production, “job shops”, custom engineering**
  - None can address this process fully

Automated Laboratory vs. Manufacturing

- **Most or all gene synthesis today is carried out in sophisticated laboratories with some automation**
  - PhDs involved
  - Difficult to scale rapidly
- **Within a few years, nearly all commercial gene synthesis will be carried out in manufacturing facilities**
  - Largely automated
  - Robots for production
  - People for process development
  - Highly sophisticated process control and scheduling
- **Interesting, meaty problems for operations research…**

The Future: Centralized Commercial Synthesis

- **Industrialization and ability to scale the critical competitive arena for commercial providers**
- **New technologies**
  - Array-based synthesis
  - New oligonucleotide synthesis technology
  - New assembly technology for large fragments
- **Centralization of commercial synthesis**
  - Two to four companies, each with a capacity to produce 20 to 50 megabases a year
  - A small number of specialized “boutique” operations
  - Most using a mix of technologies

The Future: A Dispersed Technology

- **Technology access is easy**
  - Robust, world-wide market for used equipment
  - Simple hardware for all aspects of the technology- could be built from scratch by a few engineers
  - Oligonucleotide chemistry is feasible for companies or laboratories in many (nearly all?) countries
  - Molecular biology and bacteriology kits available from many different companies in many countries
  - Protocols on the internet
- **Governments or NGOs**
  - Any country or moderately well-funded group could put together synthesis capacity FROM SCRATCH with a moderate investment ($1 million and 3-6 PhDs)

Gene Synthesis Technology is Widespread

Bioneer Corporation
49-3, Munpyeong-dong,
Daedeok-gu, Daejeon 306-220,
Korea

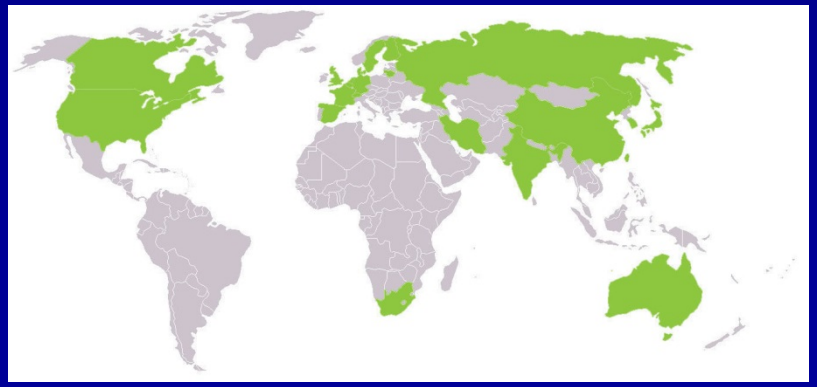
"The capacity of this facility is to produce 7.2 tons of phosphoramidite per year…

Currently we have (the) capacity of producing 20,000 oligos per day…

Bioneer offers a special gene synthesis service."

Controlling Synthesis Technology is Difficult

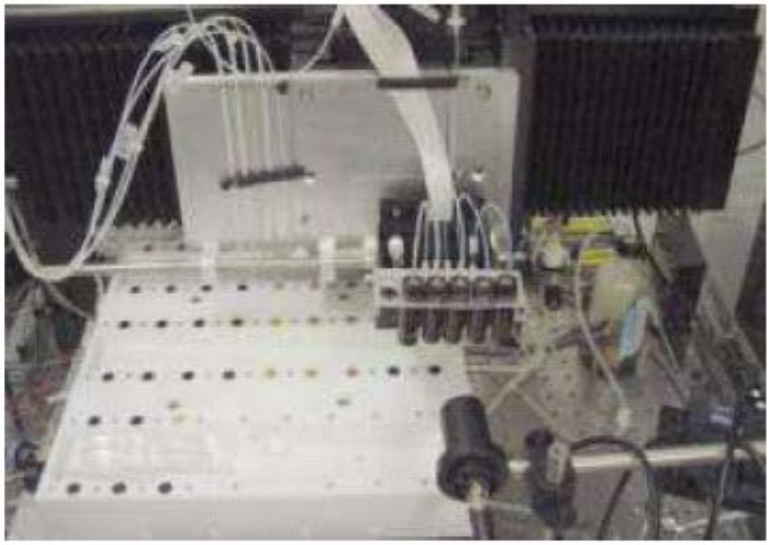
- **Synthesis materials are easy to acquire**
  - Any sophisticated chemistry group could build oligonucleotide synthesis capacity from scratch
  - For large-scale synthesis groups the “drop at the bottom of a reagent bottle” can add up to kilograms of phosphoramidite per year- tracking the materials is not feasible
- **PCR-based synthesis works on many sequences**
- **Transforming and growing bacteria is low-tech**



New Methods Extend Synthesis Capabilities

Assembly manual for the POSaM:

THE ISB Piezoelelctric Oligonucleotide Synthesizer and Microarrayer



Accurate multiplex gene synthesis from programmable DNA microchips

Jingdong Tian[1], Hui Gong[1], Nijing Sheng[2], Xiaochuan Zhou[3], Erdogan Gulari[4], Xiaolian Gao[2] & George Church[1]

Microfluidic PicoArray synthesis of oligodeoxynucleotides and simultaneous assembling of multiple DNA sequences

Xiaochuan Zhou[3,4], Shiying Cai[3], Ailing Hong[3,4], Qimin You[3], Peilin Yu[1], Nijing Sheng[1], Onnop Srivannavit[2], Seema Muranjan[3], Jean Marie Rouillard[2], Yongmei Xia[2], Xiaolin Zhang[3,4], Qin Xiang[3], Renuka Ganesh[1,4], Qi Zhu[1], Anna Matejko[1], Erdogan Gulari[2] and Xiaolian Gao[1,*]

Amplification and assembly of chip-eluted DNA (AACED): a method for high-throughput gene synthesis

Kathryn E. Richmond[1], Mo-Huang Li[1], Matthew J. Rodesch[2], Madhusudan Patel[1,3], Aaron M. Lowe[4], Changhan Kim[5], Larry L. Chu[1], Narasimhar Venkataramaian[5], Shane F. Flickinger[3], James Kaysen[1], Peter J. Belshaw[3,6], Michael R. Sussman[2,6] and Franco Cerrina[1,5,*]

Build genes with a modified ink-jet printer?

Reducing the Potential for Nefarious Uses

- Centralization will simplify monitoring and regulation of gene synthesis
- Dispersion of the technology makes complete control implausible
- Screening orders will be increasingly important

Gene Synthesis Screening

- **Some companies screen orders**
  - Different software
  - Different databases
  - Different criteria
- **Some companies do not screen**
  - Cost
  - Liability
  - Effort

Order Screening at Blue Heron

- **Screen all orders against a database of select agent genes**
  - Black Watch, Craic Computing
- **Review orders that resemble select agent genes**
  - A Ph.D. reviews several positive hits per day
  - Most hits are not select agent genes
- **Detailed analysis of select agent genes**
  - Most select agent genes are OK to provide
  - Some require significant review
  - Check the literature
  - Discuss with customer
- **Decide if we will build the sequence**

Screening Tools

- **Current tools very simple**
  - Homology search (Blast)
  - High false positive
  - Low or zero false negative
- **No database of “Select Sequences”**
- **Rules require interpretation**
- **Therefore, screening is expensive**

Industry Consortium

- **International Consortium for Polynucleotide Synthesis (ICPS)**
- **Goals**
  - Develop improved screening software and other tools to simplify and improve screening
  - Encourage the widespread use of these tools
  - Provide an industry point of contact with government
- **Status**
  - Established in June
  - Founding members recruiting other companies
  - Establishing operational group