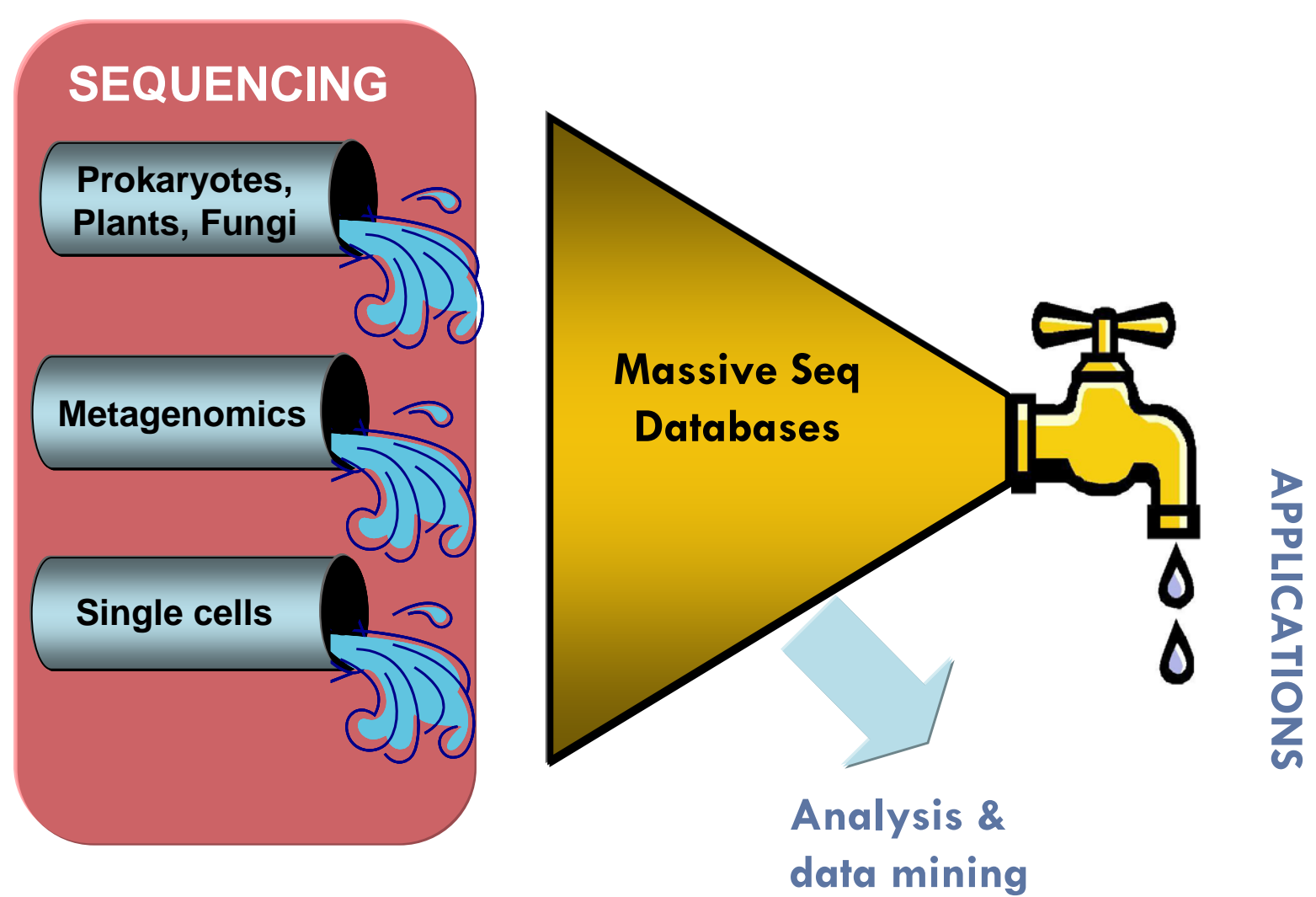


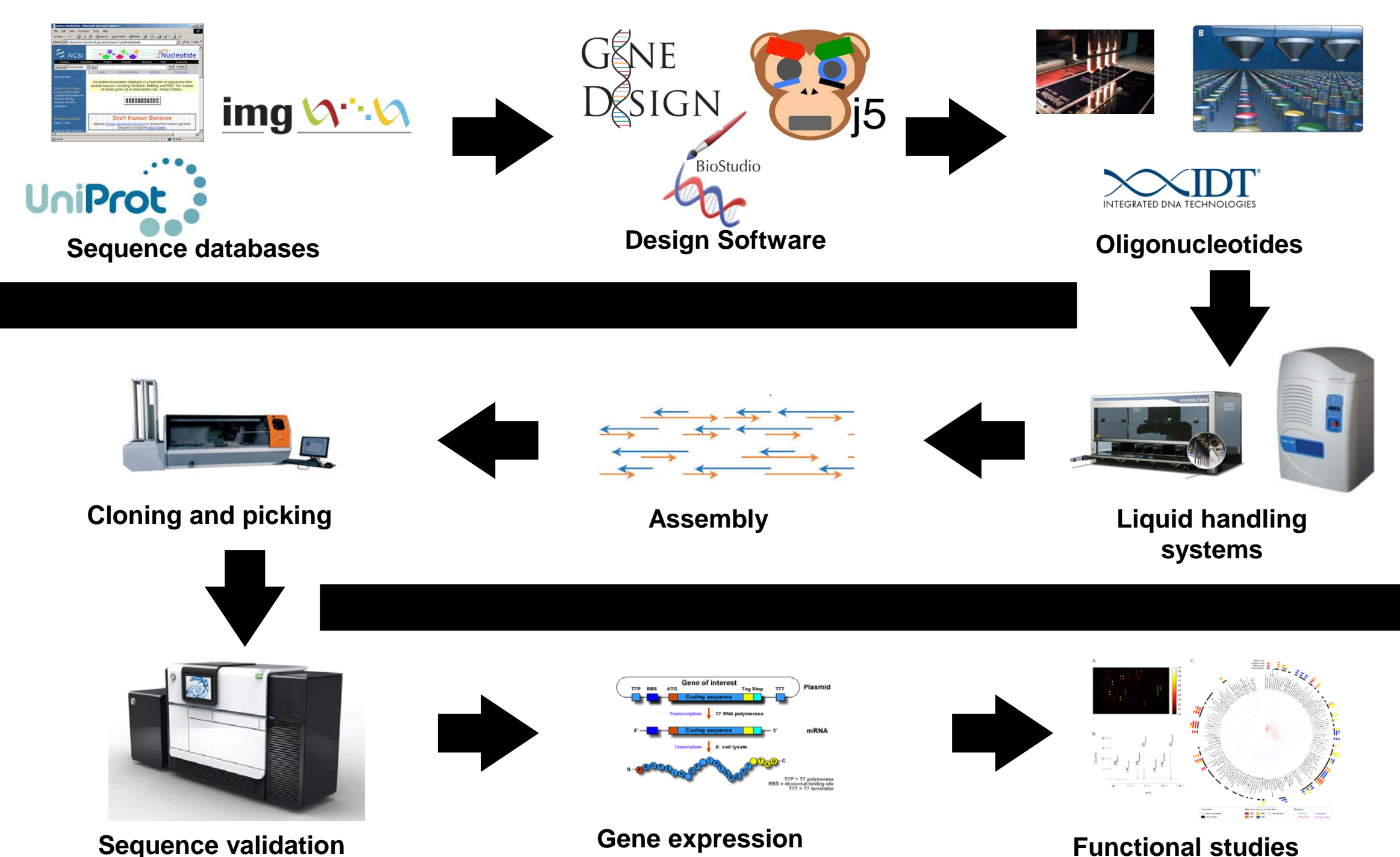
Adding Value to Sequence Information



Sequence information specially from metagenomes and single cell genomes is essentially **digital information** that cannot be easily accessed for biological hypothesis testing. This represents a limiting factor in the development of sequence based applications.

DNA synthesis allows digital sequence information to be 'translated' into biology by generating DNA in a template-independent manner.

DNA Synthesis Workflow at JGI



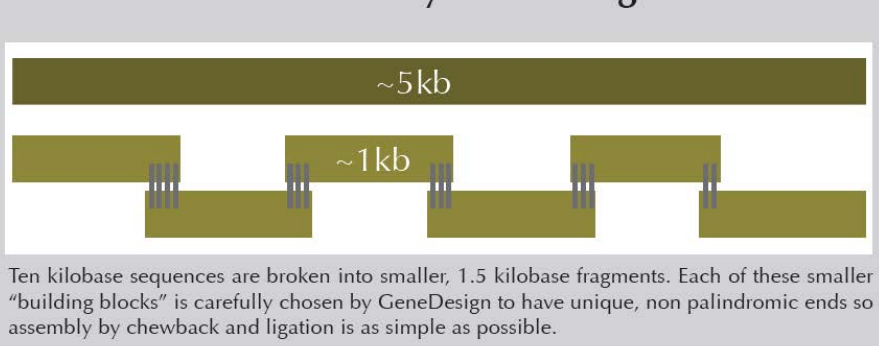
There are many steps involved in the DNA synthesis and assembly. We have implemented a suite of informatics tools to assist the design of oligos, the assembly of larger constructs, and potential strategies for combinatorial libraries. For the lab processes we have utilized robotic workstations, next-generation sequencing, and novel enzymology solutions to significantly improve the scale and reduce the cost.

Informatics Tools for Designing Oligos and Assemblies

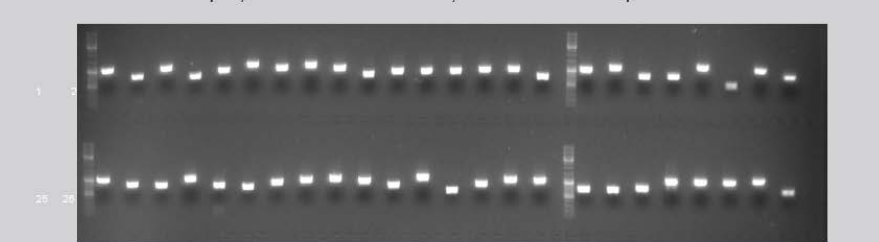
User sequences are submitted to the JGI synthesis pipeline and manipulated by a suite of software tools.



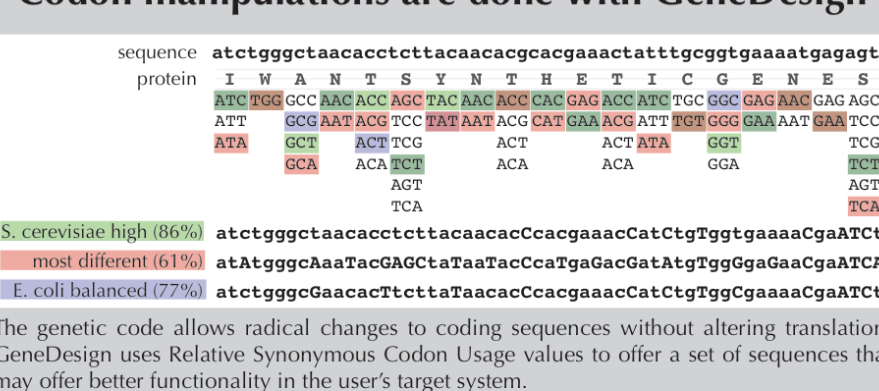
Smaller sequences are partitioned into assembleable constructs by GeneDesign



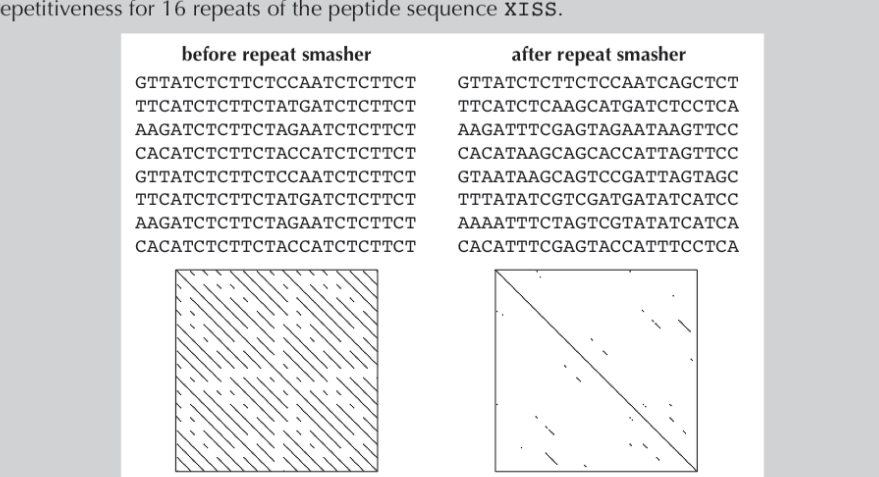
1.3 kilobase fragments are broken into overlapping 300 base oligonucleotides. Each oligo overlap is chosen by GeneDesign to have unique, non-palindromic ends. Small oligos are included to drive the polymerase chain assembly towards the final product.



Codon manipulations are done with GeneDesign

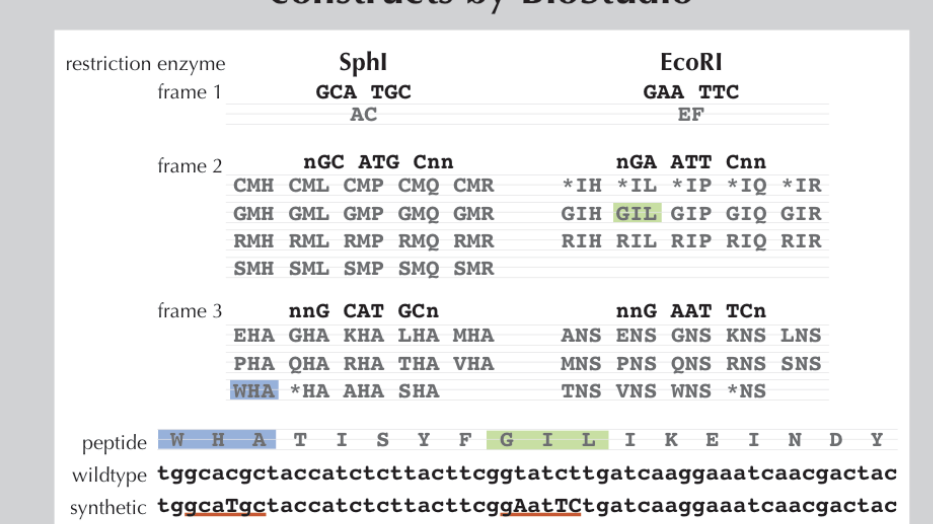


Highly repetitive sequence is a significant obstacle to synthesis. GeneDesign 'smashes' repeats by preserving repetitive protein sequence while altering the underlying nucleotide sequence. The diagram of before and after sequences below illustrate the removal of nucleotide repetitions for 16 repeats of the peptide sequence XEES.



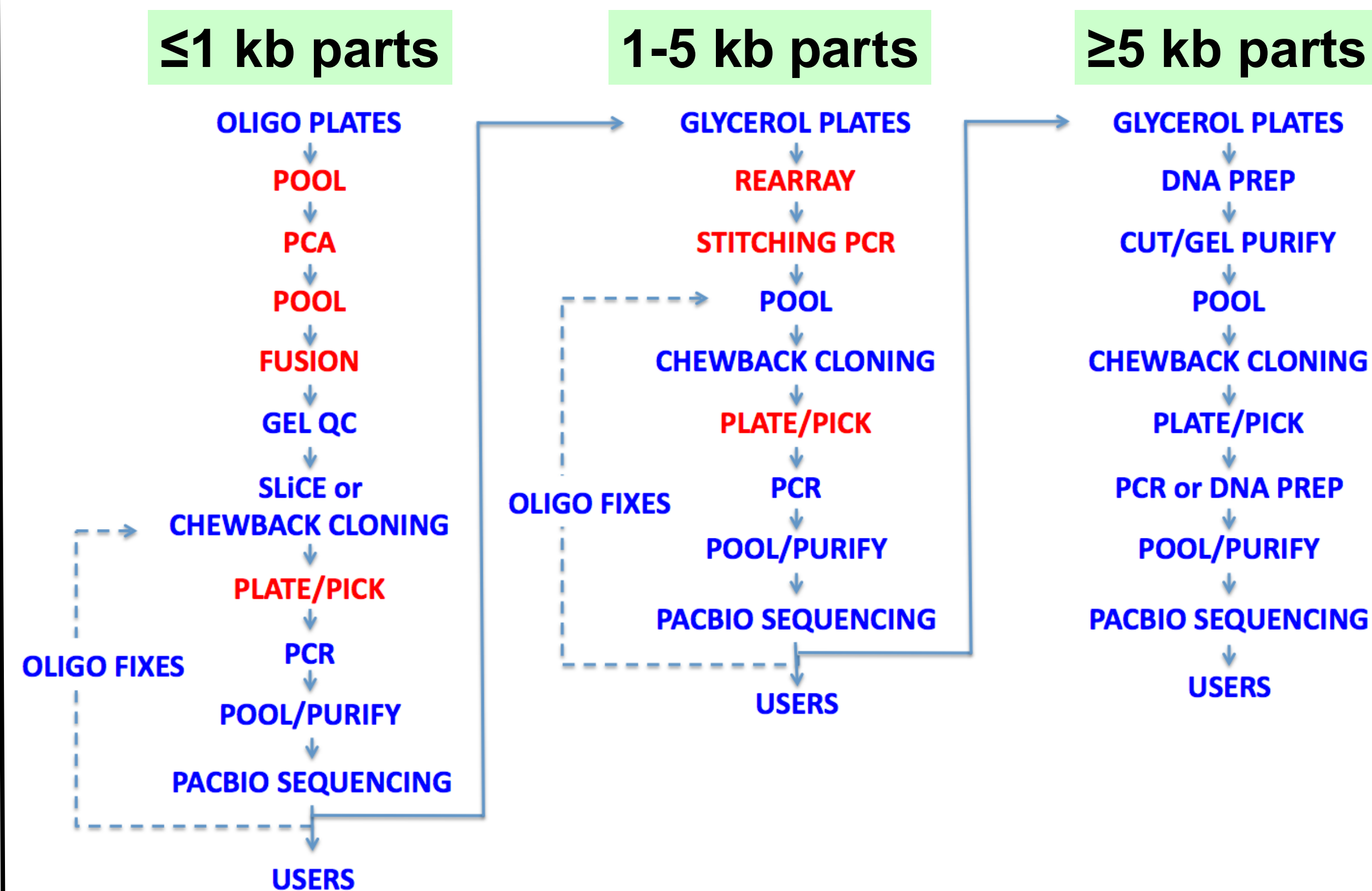
These informatics tools are publicly available. Listed here are some examples of what these tools can assist in the design of DNA parts and assembly. Users have the options to use any of these tools to design their own constructs. J5 is an automated DNA assembly software developed at JBEI (j5.jbei.org). It provides a web-base user interface to perform multi-part, combinatorial assembly.

Larger sequences are partitioned into assembleable constructs by BioStudio



BioStudio automates this process for chromosomes, placing unique restriction enzyme recognition sites as landmarks that divide the chromosome into constructs of custom size, each of which is divided into smaller, modular pieces to a minimum of about 10kb. The overlaps left by two consecutive landmark sites are different and non-palindromic, in order to support physical construction.

Laboratory Processes



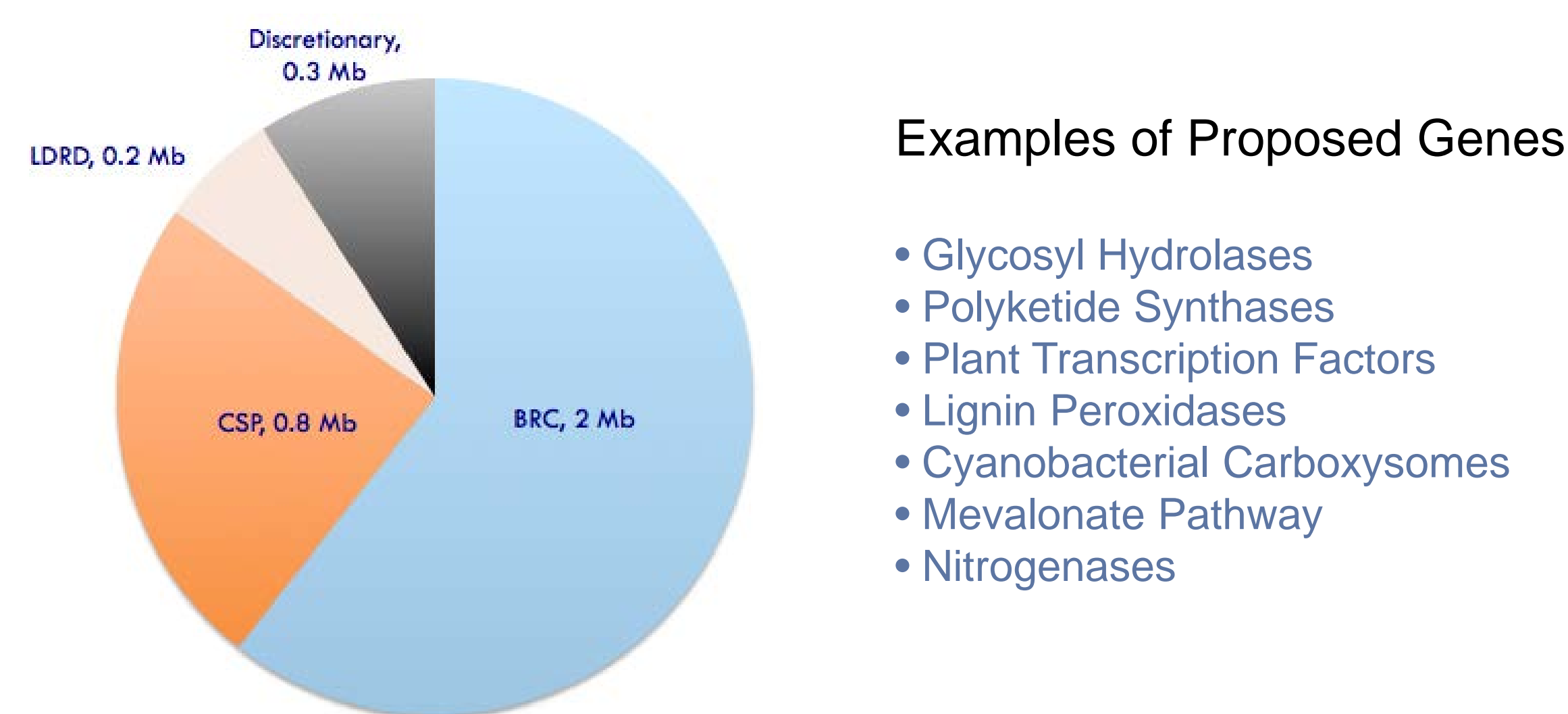
The laboratory processes begin with the construction of basic 1Kb DNA parts. Typically 48-96 basic parts can be synthesized in a week by a single FTE. Several robotic workstations are applied to handle the labor intensive steps shown in red. The basic parts are then assembled to larger size constructs in a similar process.

Synthesis Design and Status Reporting Platform

Name	Project	Status
RAS	Rosmarinic Acid	Hedgepodge (2, 5, 10)
B	Limonene	Complete
A	Limonene	Changes requested
TAT	Rosmarinic Acid	Waiting for reagents
Thermosaurus aurantiacus	Fungal Enzymes	Failed

We plan to integrate the DIVA (Design, Implementation, Validation Automation) platform developed at JBEI with the DNA synthesis pipeline. This will provide a web-based graphical user interface for external users of the JGI's DNA Synthesis Program

Synthesis Allocation FY13

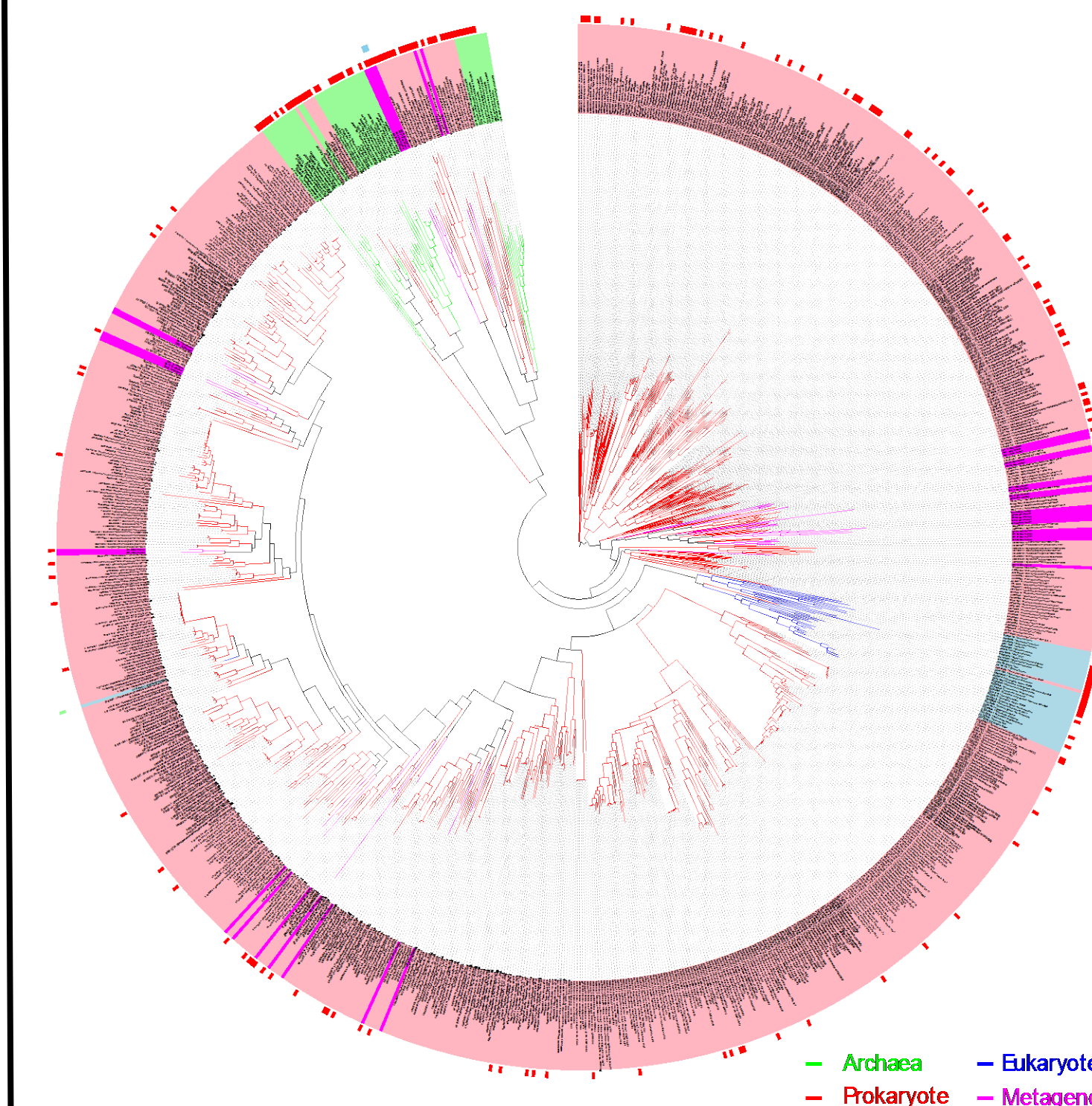


Examples of Proposed Genes

- Glycosyl Hydrolases
- Polyketide Synthases
- Plant Transcription Factors
- Lignin Peroxidases
- Cyanobacterial Carboxysomes
- Mevalonate Pathway
- Nitrogenases

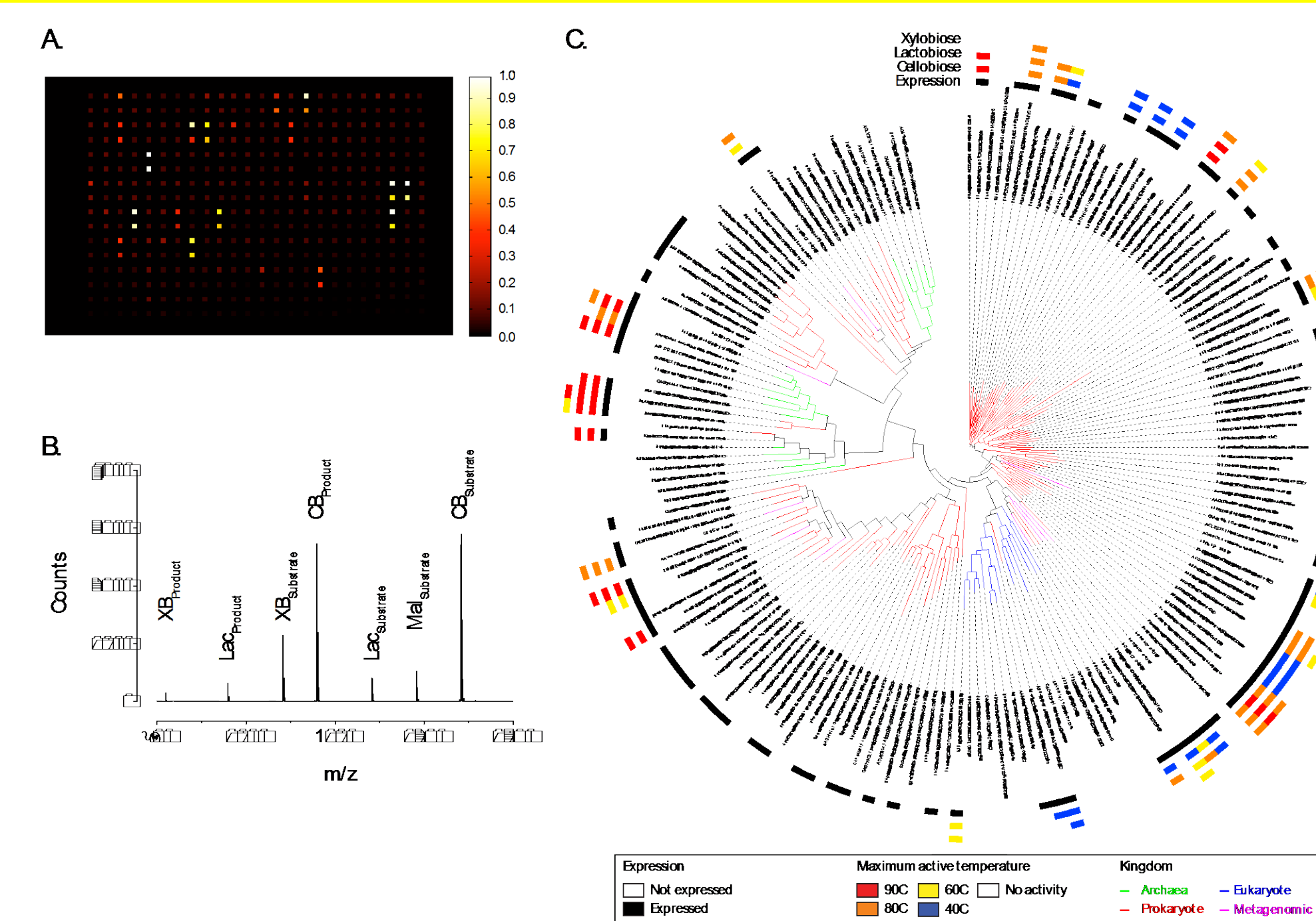
The FY13 synthesis capacity will be used primarily to support Bioenergy Research Centers (BRCs) and our Community Sequencing Programs (CSPs). Additional capacity will be used to develop proof-of-principle projects.

Systematic Survey of GH1 Enzyme Diversity



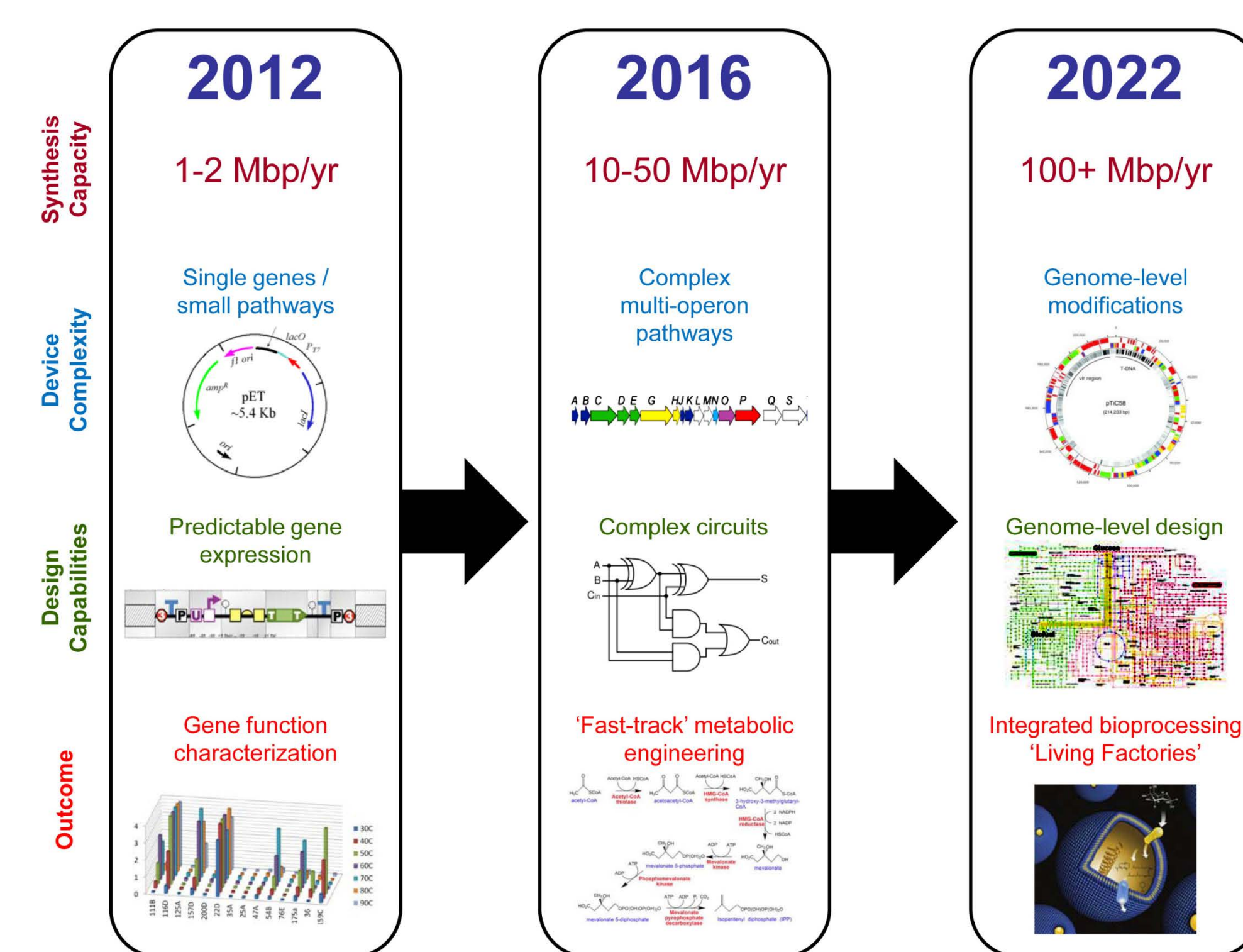
Our initial demonstration project was to functionally characterize the natural diversity of GH1 enzymes. GH1s participate in the last step of biomass breakdown (cellobiose to glucose). We identified GH1 candidate genes from NR, CAZy and JGI metagenomics databases, performed multiple sequence alignments, constructed phylogenetic trees and selected 200 representatives that cover the maximum sequence space (selected GH1 genes are represented by red dots on tree outer circle).

Synthesis and Biochemical Characterization



We have synthesized and over-expressed our 200 GH1 candidate genes. Initial biochemical characterization at different temperatures and pHs has revealed a wide range of functional properties, validating our phylogenomics approach.

Synthetic Biology Vision at JGI



Synthetic biology approaches combined with large scale genomics information, will result in novel sequence driven applications. This will require substantial output and cost improvements as well as increasingly complex regulation and circuitry tools.