

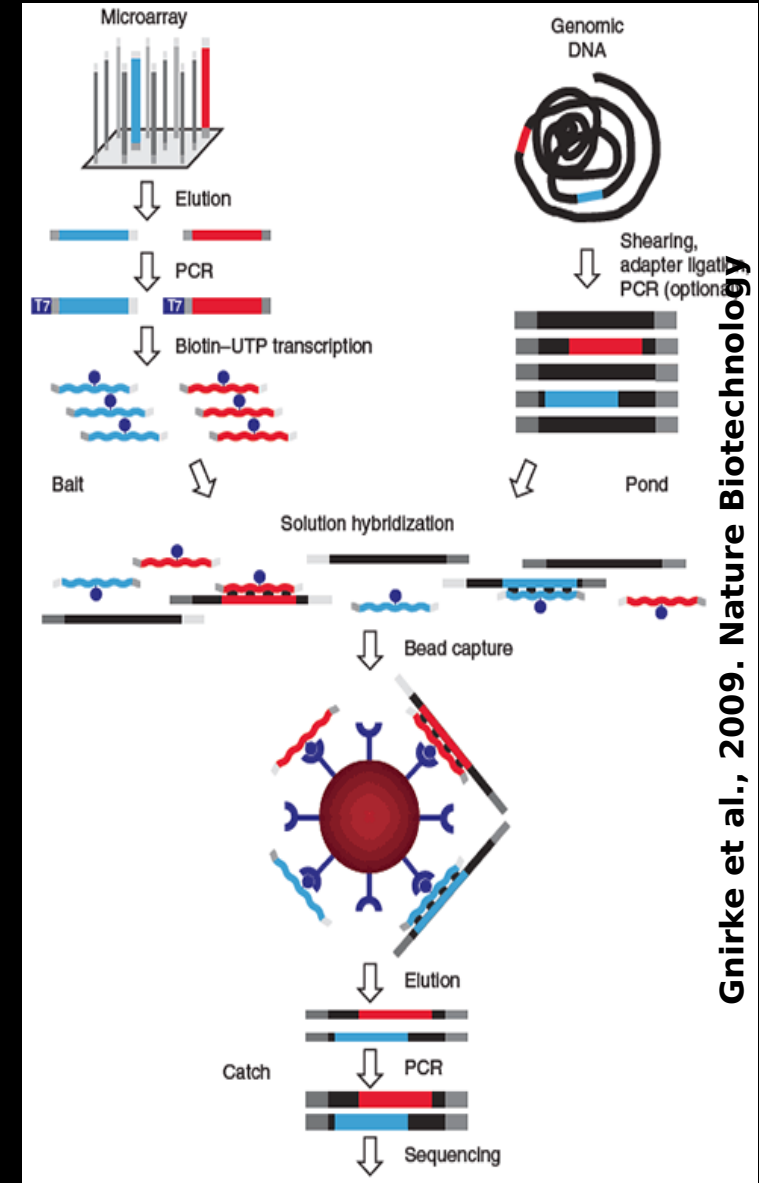
Targeted Enrichment through Hybridization Sequencing

Targeted Sequencing: HybSeq Enrichment

Solution Phase Hybridization

*Examples: Agilent SureSelect,
NimbleGen SeqCap EZ,
Microarray MyBait, etc*

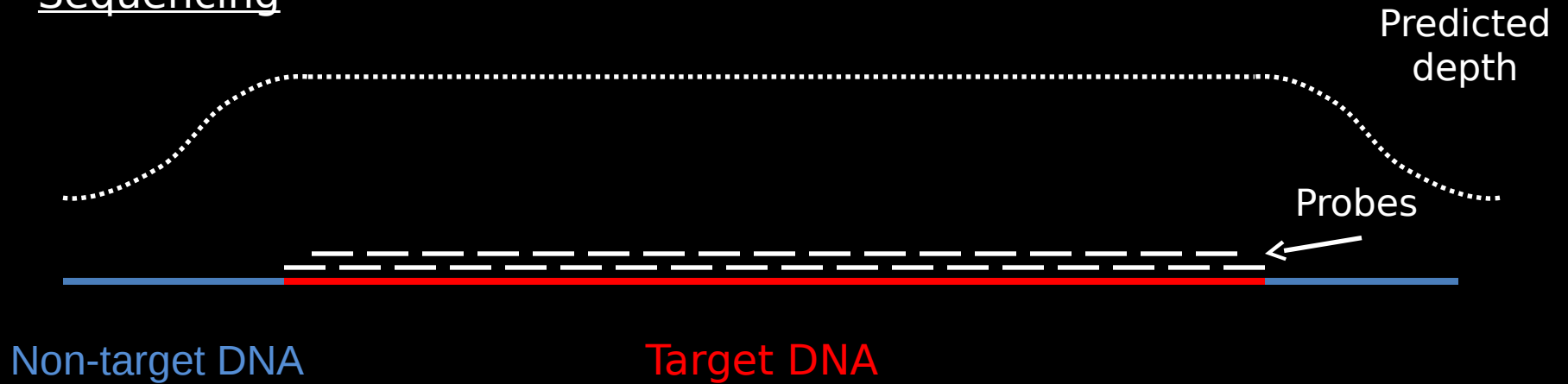
- 'Baits' synthesized on arrays
- 120 bp RNA probes.
- Hybridization in solution
- Immobilization via biotin-streptavidin capture
- 1 – 5 ug DNA of input library.
- Scalable to large samples, targets



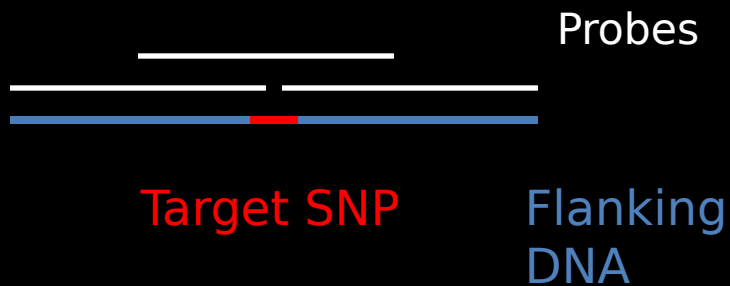
Gnirke et al., 2009. Nature Biotechnology

Targeted Sequencing: HybSeq Enrichment

Multi-Locus Sequencing

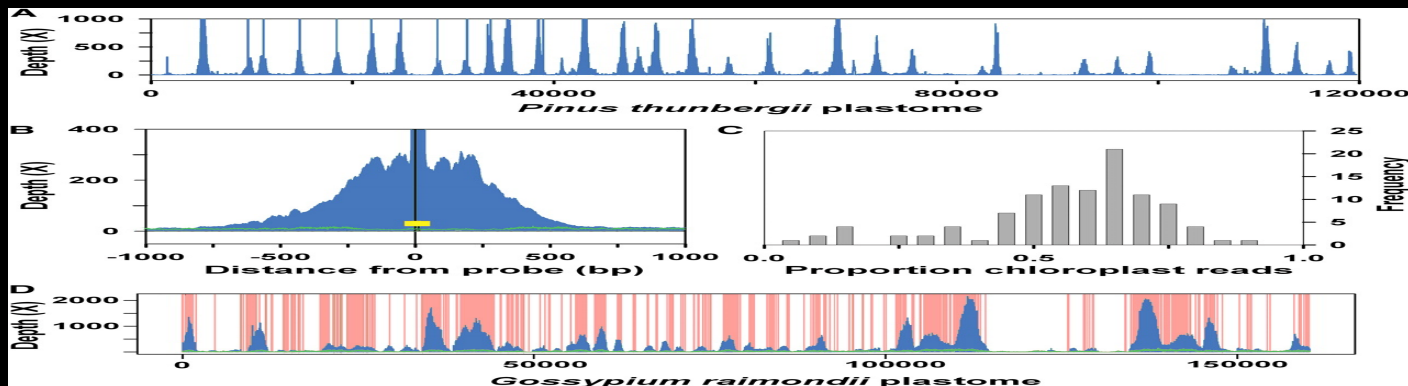
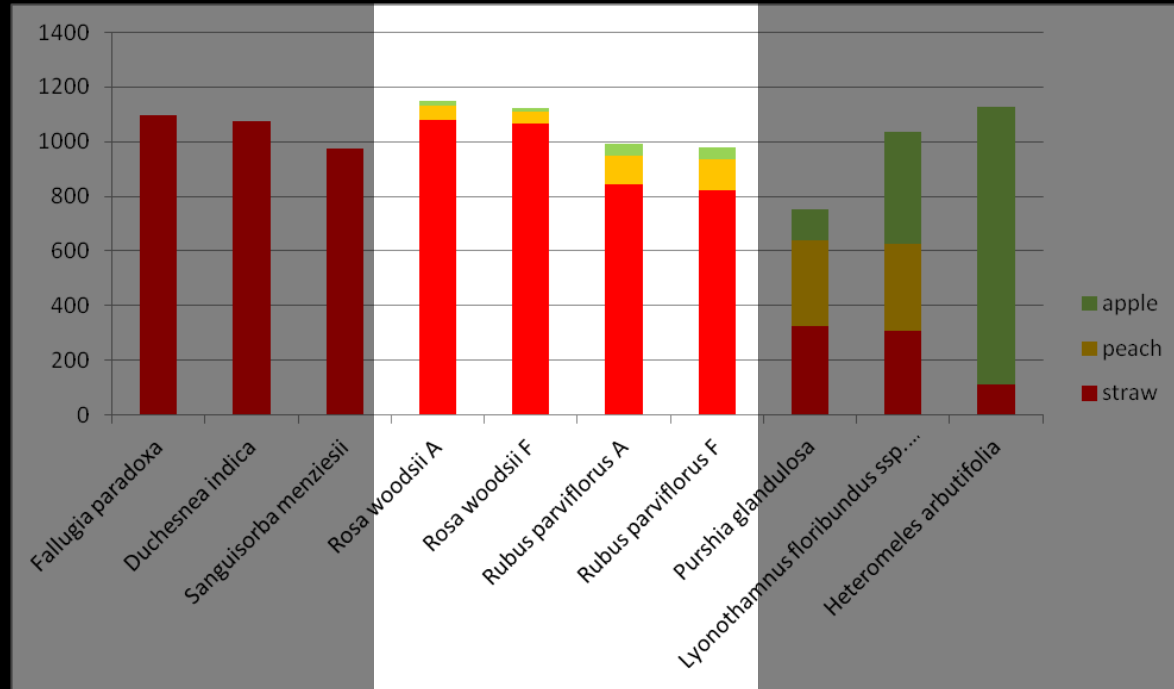


SNP Genotyping



Targeted Sequencing: HybSeq Enrichment

Probe specificity allows enrichment in distantly related organisms:



Cronn et al
2012

**Size-selected library
Adapters Ligated
100 - 500 ng**

Targeted Sequencing: HybSeq Enrichment

Hybridization to probes

Cleaning & Recovery

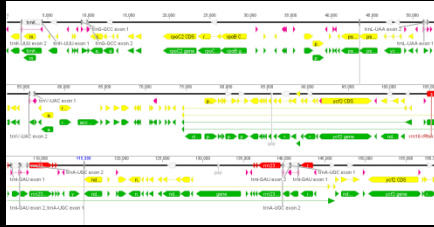
Elution of enriched library

PCR Enrichment

- Use precautions to prevent RNA degradation.

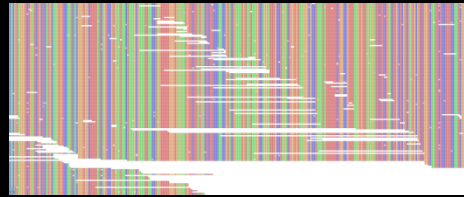
- User provides Streptavidin magnetic beads

Probe development for genic loci



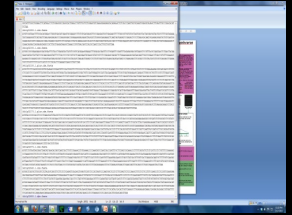
Asclepias syriaca
genome assembly

-Filter for single copy



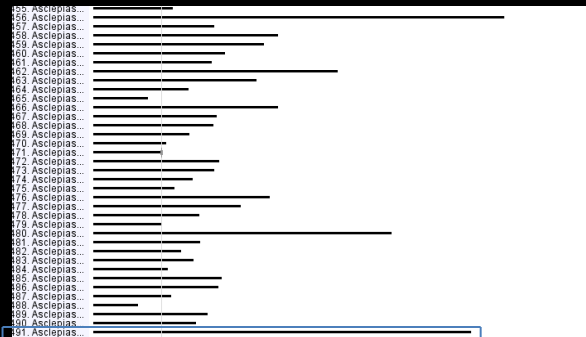
A. syriaca transcriptome

-Target genes, exons

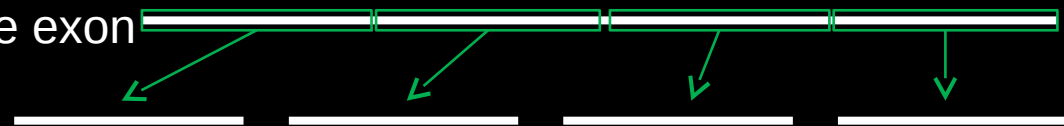


Sets of
conserved,
low-copy loci.

Set of
Asclepias
exons



Single exon

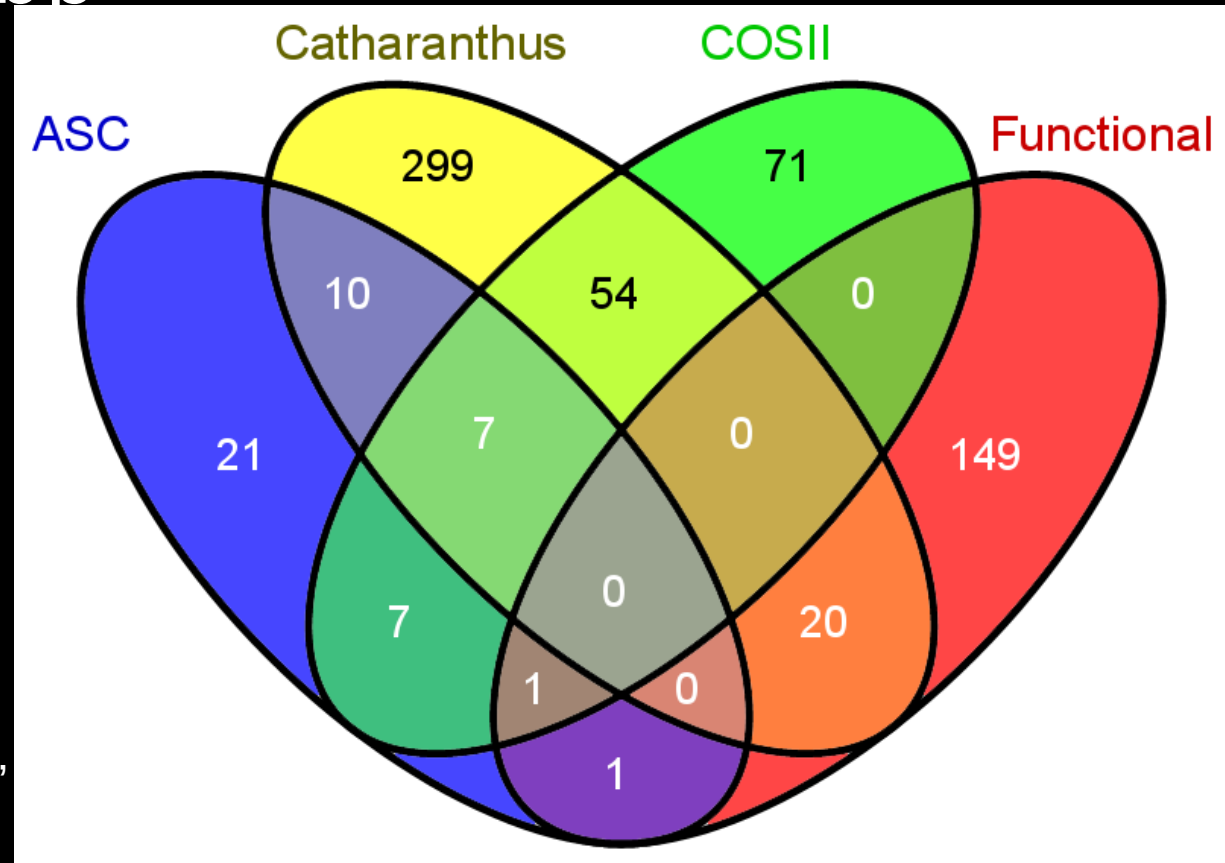


Probes for targeted sequencing.

Long locus targets

- 3762 'exons'
- 925 total loci
- 1,720,160 bp

640 "Annotated"



SNP Targets

Identification of SNPs

de Bruijn graph “Bubble”



Bubbles with high, even depth may represent heterozygosity.

4000 SNPs targeted

- 566 paired on 10 kb contigs
- 3434 match 1-10 kb contigs (3717 total contigs)



Targeted Enrichment vs. Restriction Techniques

Advantages:

- Known genomic context
- Targets can include SNPs or candidate genes
- Exclude probable paralogs
- Homologous across taxa
- Less input DNA
- Feasible with degraded DNA

Disadvantages:

- Reference genome or transcriptome required
- Higher cost/sample
- Fewer sites (depends)

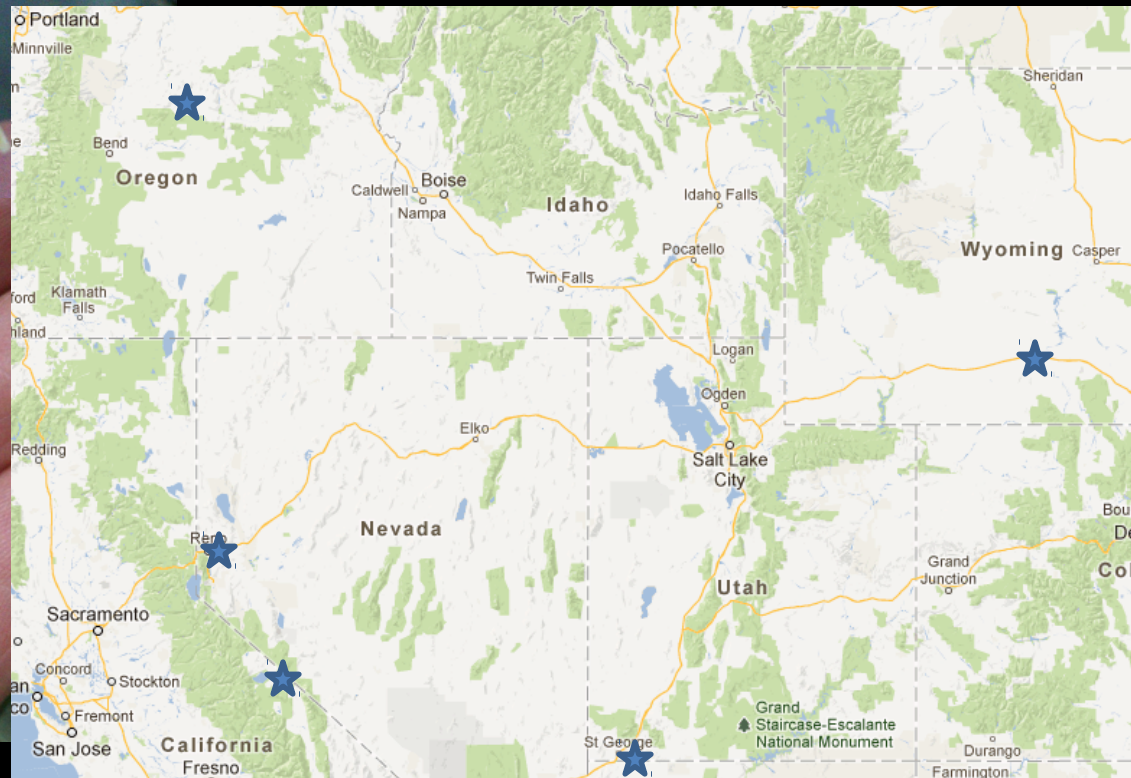
Probes tested in 12 *Asclepias cryptoceras* individuals



MiSeq, 250 bp Paired-End

6.6 million pairs

1.1 million reads / individual



Targeted Capture Success Across 12 Individuals

3717 / 3762 exons hit by assembled reads

921 / 925 loci hit with ≥ 1 hit exon

Percent of reads mapping to targets.



Do interspecific SNP targets work?

3540 / 3717 Contigs hit by reads.

92% Polymorphic

3708 sites are polymorphic and high depth (20x).

-On 295 contigs.

YES!

