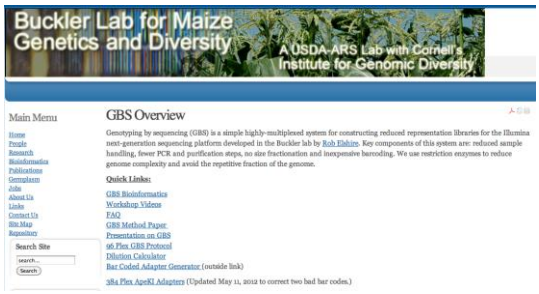
Alternatives to whole genome sequencing:

Genotyping by sequencing (GBS), Restriction site associated DNA markers (RAD-seq), and SNPs for mapping

Dustin Mayfield
University of Missouri

Resource: http://www.maizegenetics.net/



Why reduce representation

- Large, repetitive genome species
- Little or no known genomic information
- Large scope of sampling – population level

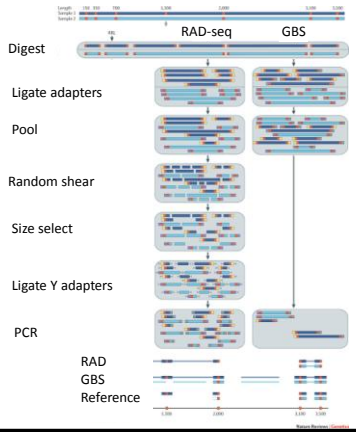
How to reduce representation

1. Target Enrichment
- Sequence Capture – oligonucleotide baits designed to bind to regions of interest and these regions are selected or enriched
- Long range PCR – specific genes and genomic subsets
- Molecular inversion probes – probes are single stranded DNA molecules that contain sequences complimentary to target in genome, and they hybridize to and capture the genomic target

How to reduce representation

2. Restriction Enzymes
- Different enzymes for different organisms

Both RAD-seq and GBS methods begin with restriction enzyme digestion



Modified from Davey et al. 2011

Dealing with Missing Data

Impute

Reduce multiplexing level

Sequence the library again

Choose less frequently cutting enzyme

How many SNPs will one get?

Species	Enzyme	Sample #	Marker #
Corn	ApeKI	54,000	681K
Rice	ApeKI	850	>60K
Grape	ApeKI	1000	200K
Shrub Willow	ApeKI	459	23K
Shrub Willow	EcoT22I	463	2K
Scrub Jay	PasI	107	>9K
Bank Vole	PstI	283	53K
Fox	ApeKI	48	147K
Fox	EcoT22I	48	16K
Cow	ApeKI	48	13K
Cow	PstI	48	64K

http://www.maizegenetics.net/