



## Read construction



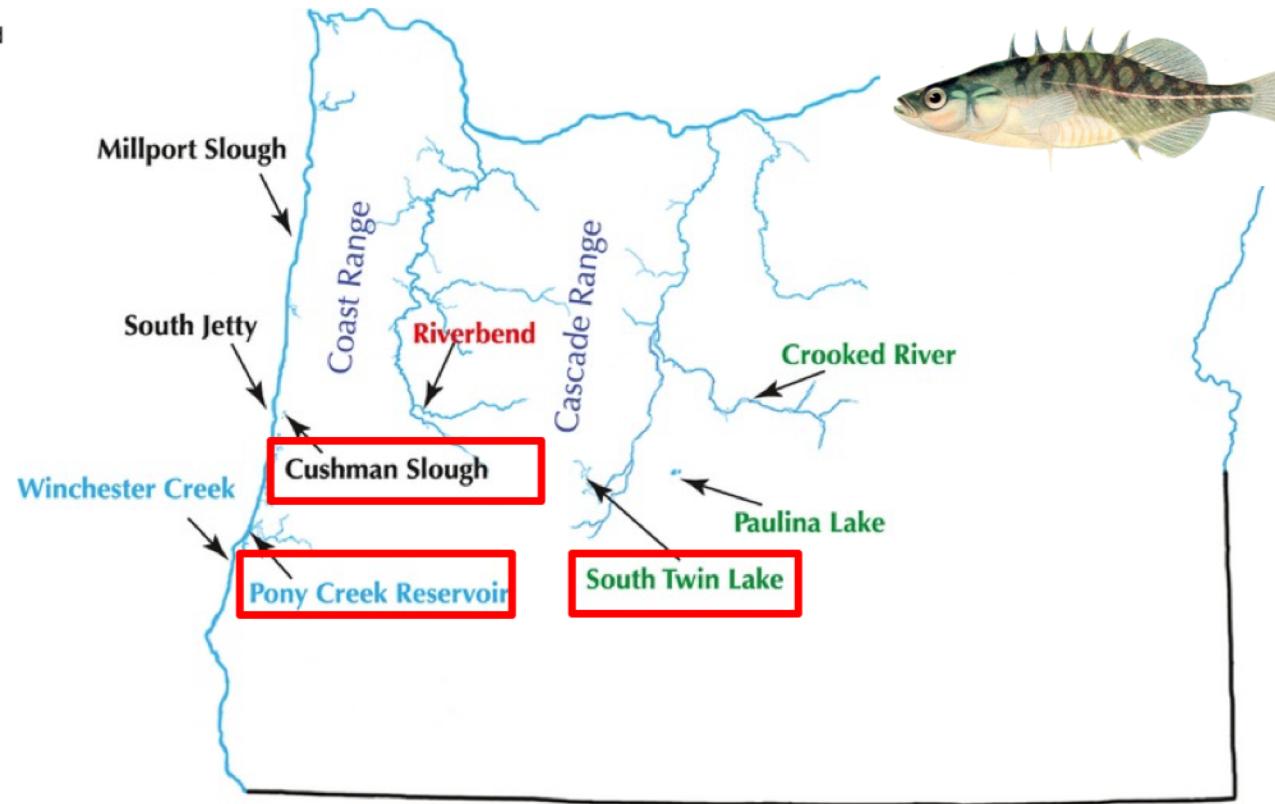
# Read construction



A)



Mazzarella et al. 2016



Catchen et al 2013

# Denovo pipeline

P  
i  
p  
e  
l  
i  
n  
e

**USTACKS**

Build loci de novo

**CSTACKS**

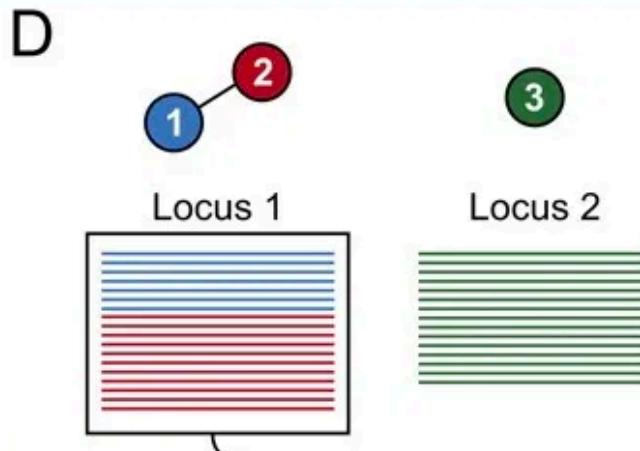
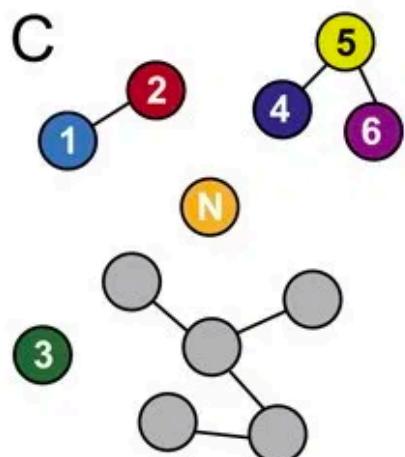
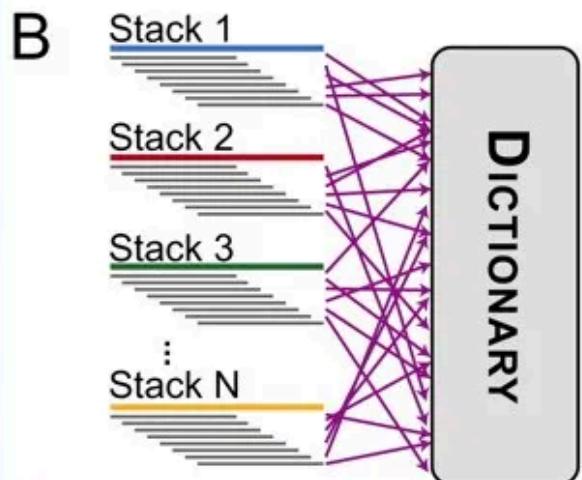
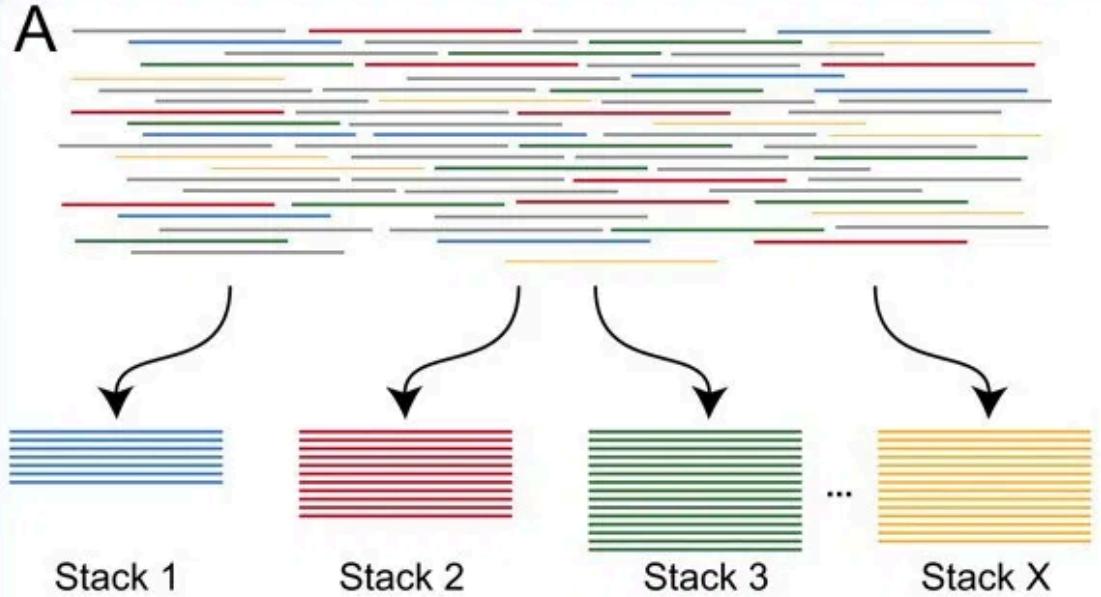
Assemble catalog

**SSTACKS**

Match to catalog

**Populations**

Filter loci and compute  
Population statistics



**E**

1° reads

```
TGCAGGGACACACAGGAGCTGAGCCATT CCT GCGGCTCC C GACCAACGTTTG
```

2° reads

```
TGCGGGGACACACAGGAGCTGAGCCATT CCT GCGGCTCC C GACCAACGTTTG
TGCAGGGACACACAGGAGCTGAGCCATT CCT GCGGCTCC C GACCAACGTTGCTG
TGCAGGGACACACAGGAGCAGAGCCATT CCT GCGGCTCC C GACCAACGTTTG
TGCAGGGACACACAGGAGCTGAGCCATT CCT GCGGCTCC C GACCAACGTTTG
TGCAGGGACACACAGGAGCTGAGCCATT CCT GCGGCTCC C GACCAACGTTTG
```

**F**

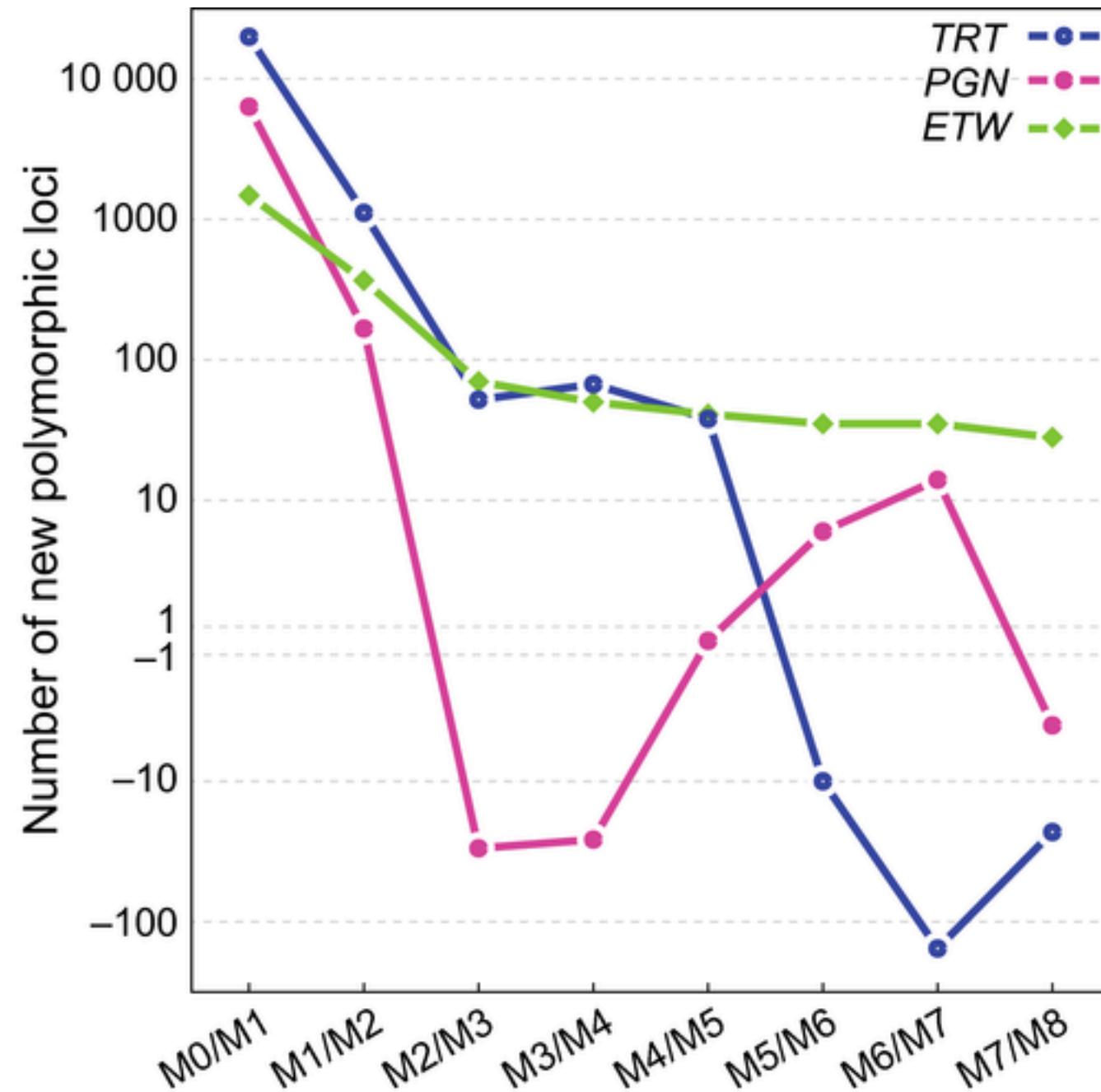
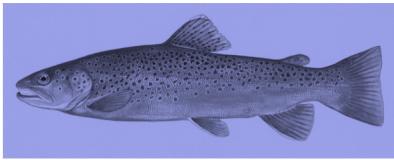
$\Sigma$  ↑

```
TGCAGGGACACACAGGAGCTGAGCCATT CCT GCGGCTCC C/GACCAACGTTTG
```

Parameter Description	<b>denovo_map.p</b> <b>1 Parameter</b>	Pipeline component	Component Parameter	Default Value
Minimum stack depth / minimum depth of coverage	<code>-m</code>	<b>ustacks</b>	<code>-m</code>	3
Distance allowed between stacks	<code>-M</code>	<b>ustacks</b>	<code>-M</code>	2
Distance allowed between catalog loci	<code>-n</code>	<b>cstacks</b>	<code>-n</code>	1

*"If  $-M$  is too low, you start separating variants of a single gene as different genes. If you set them too high, you start merging paralogs."*

We will find collectively  
the parameters that find  
the most new loci at  $r = 0.8$



reference pipeline

P  
i  
p  
e  
l  
i  
n  
e

**USTACKS**

**CSTACKS**

**SSTACKS**

**Populations**

Build loci de novo

Assemble catalog

Match to catalog

Filter loci and compute  
Population statistics

# Reference pipeline

P  
i  
p  
e  
l  
i  
n  
e

**Populations**

**Alignment to a reference genome**

Filter loci and compute  
Population statistics

