

Toy Example with Adaptive Substitution Analysis

Hanxi Sun

Toy Data

```
ALPHABET: A C G T
LENGTH: ~ 100,000 bp
BACKGROUND: 0.325828 0.191345 0.182691 0.300136
RATE_MAT:
-0.805948    0.144819    0.512276    0.148853
 0.246602   -1.248680    0.160256    0.841822
 0.913638    0.167846   -1.331401    0.249917
 0.161594    0.536683    0.152123   -0.850401
TREE: ((human:0.103078,(mouse:0.0751206,rat:0.0788608):0.279636):0.101507,cow:0.101507);
```

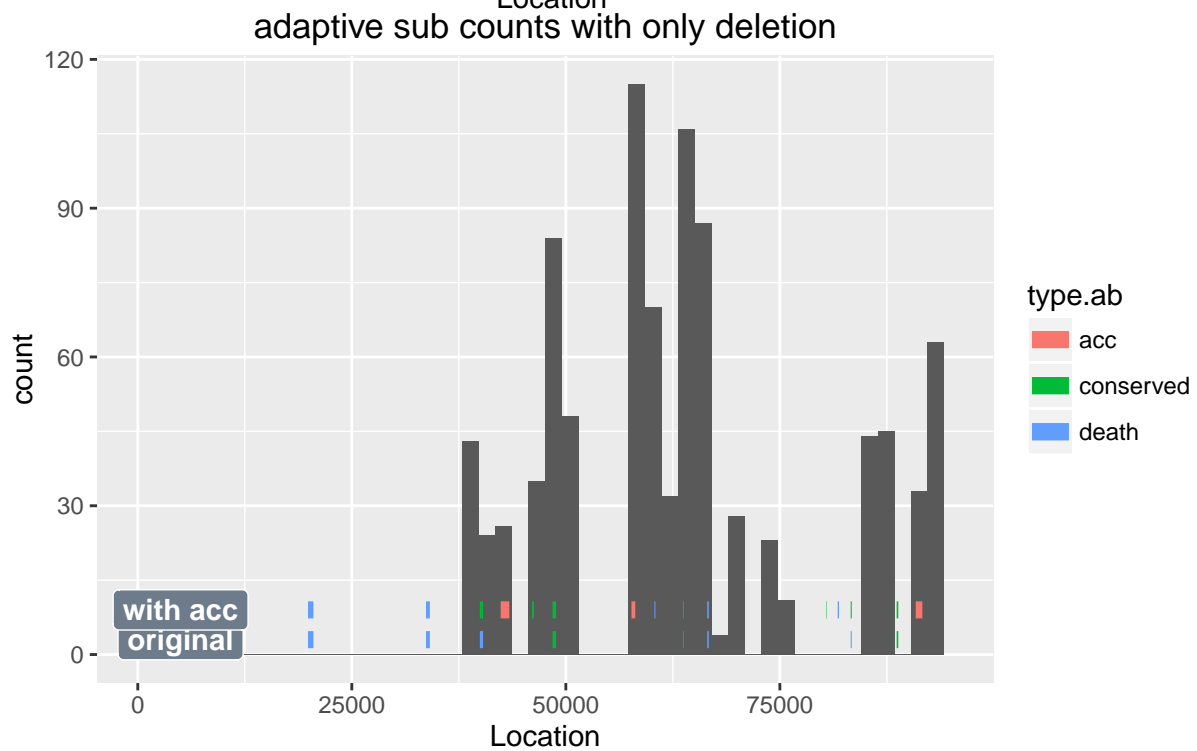
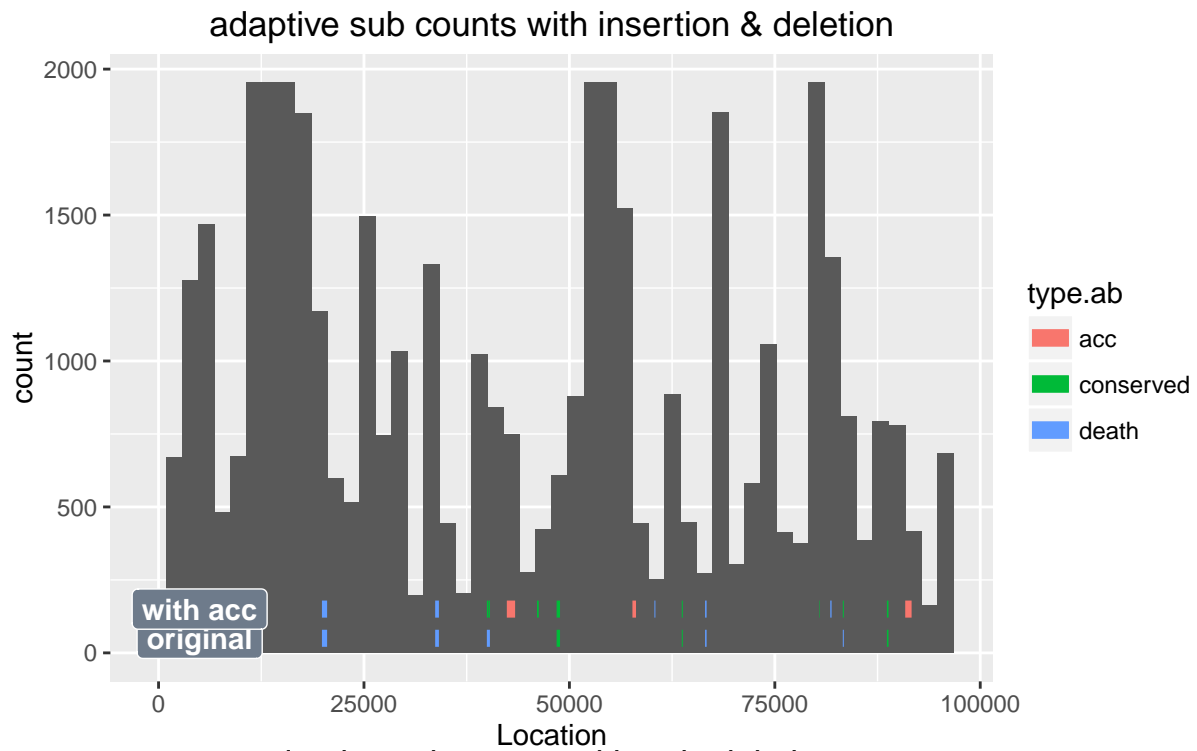
Original DLESS (phyloHMM) output

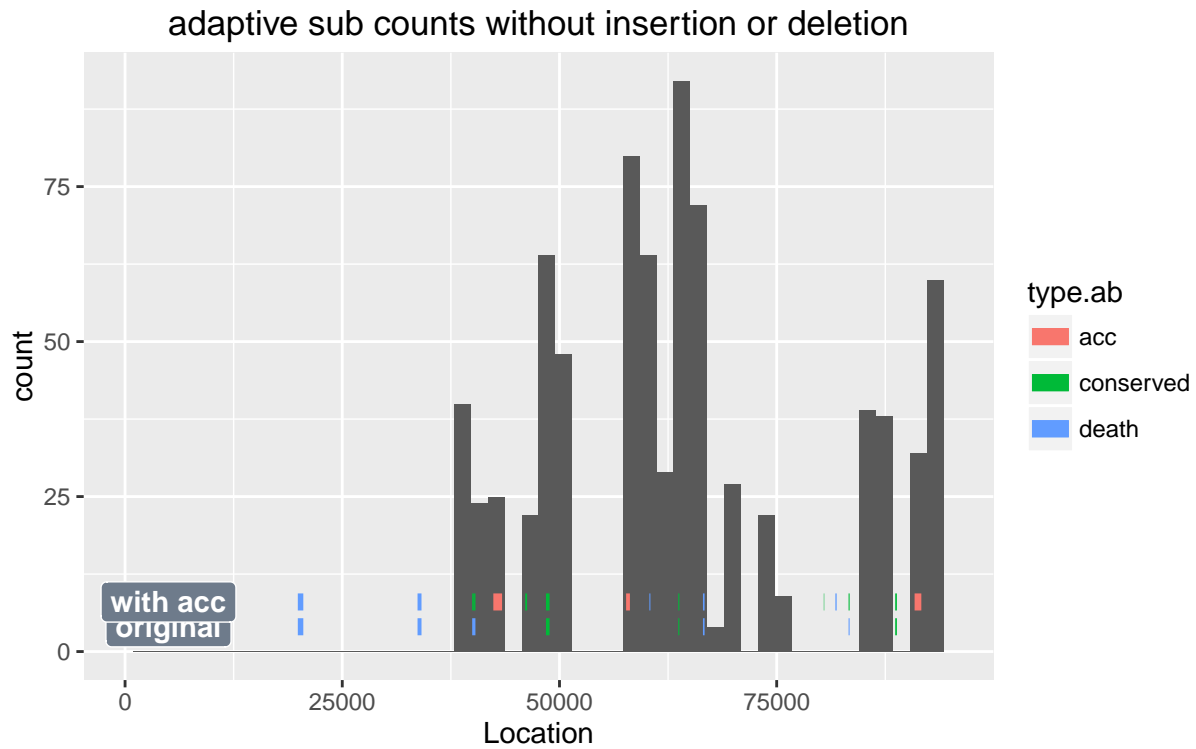
```
##           type start  end
## 1    cow-death 19883 20514
## 2    cow-death 33663 34117
## 3 mouse-death 39953 40326
## 4   conserved 48444 48857
## 5   conserved 63675 63798
## 6 human-death 66512 66699
## 7    cow-death 83272 83397
## 8   conserved 88638 88835
```

Modified DLESS output

```
##           type start  end
## 1    cow-death 19883 20514
## 2    cow-death 33663 34117
## 3    conserved 39953 40140
## 4    conserved 40170 40326
## 5    human-acc 42385 43392
## 6    conserved 46057 46272
## 7    conserved 48444 48857
## 8      cow-acc 57668 58108
## 9    rat-death 60346 60465
## 10   conserved 63675 63798
## 11   human-death 66512 66699
## 12   conserved 80421 80474
## 13 mouse-rat-death 81751 81905
## 14   conserved 83272 83397
## 15   conserved 88638 88835
## 16   mouse-acc 90869 91646
```

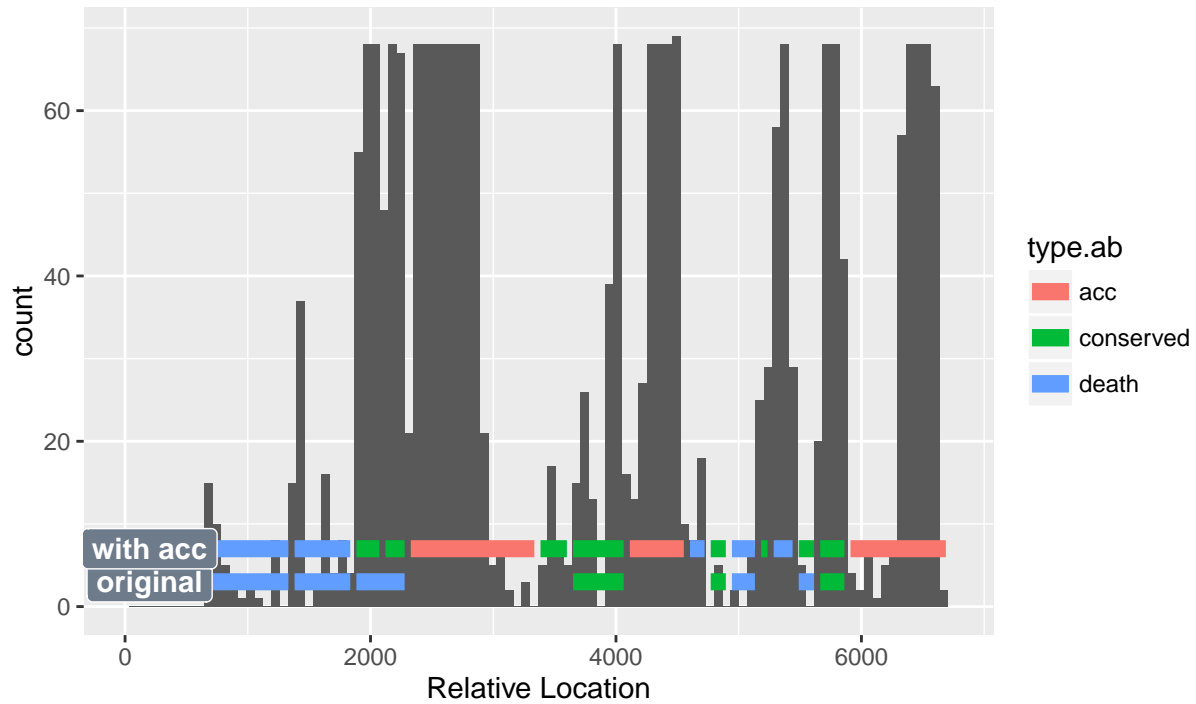
Plots on the Chain



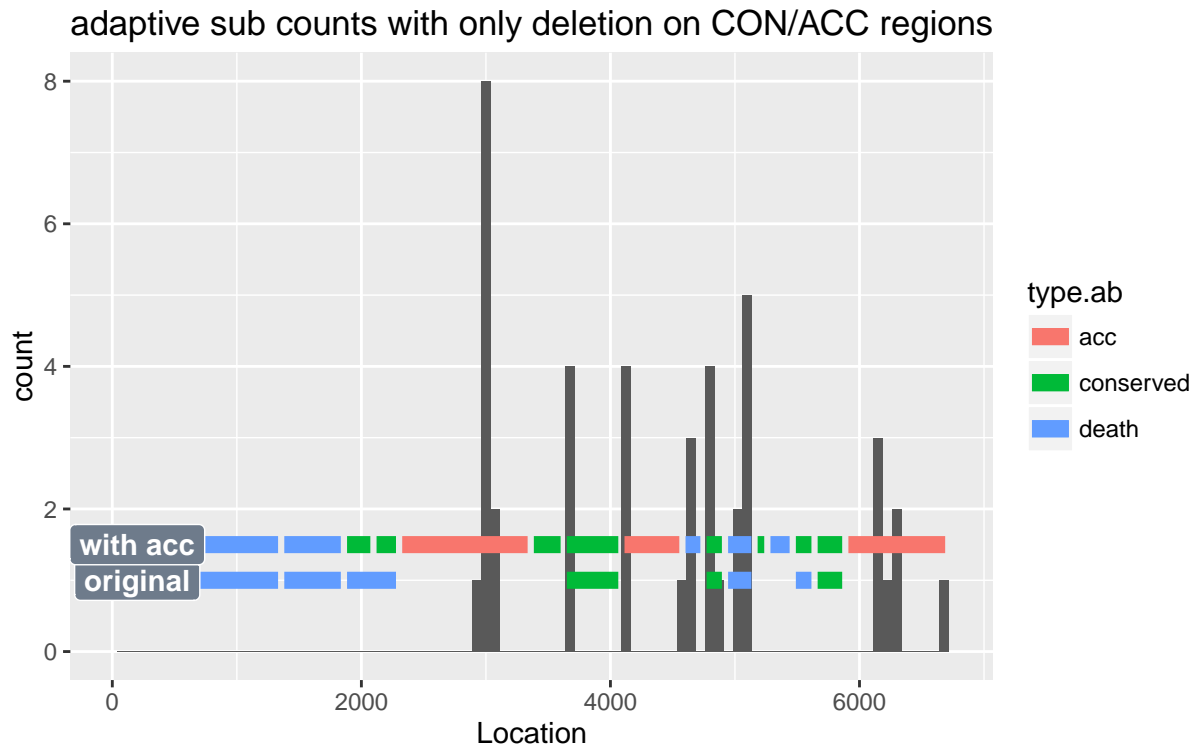


Plots on the Conserved/Accelerated Regions

adaptive sub counts with insertion & deletion on CON/ACC regions



Warning: Removed 1 rows containing missing values (geom_bar).



Warning: Removed 1 rows containing missing values (geom_bar).

adaptive sub counts without insertion or deletion on CON/ACC regions

