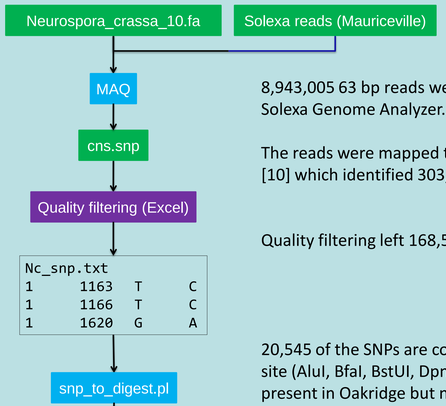
A SNP map for *Neurospora crassa* Mauriceville

Kyle R. Pomraning[1], Kristina M. Smith[2,3] and Michael Freitag[2,3]. [1]Molecular and Cellular Biology Program, [2]Dept. of Biochemistry and Biophysics, [3]Center for Genome Research and Biocomputing, Oregon State University, Corvallis, OR 97331, U.S.A.

contacts: pomranosaurus@gmail.com; freitagm@cgrb.oregonstate.edu

Identification of SNPs between the *N. crassa* Mauriceville and Oakridge strains by genomic sequencing



8,943,005 63 bp reads were generated for Mauriceville using a Solexa Genome Analyzer.

The reads were mapped to the Oakridge genome using MAQ [10] which identified 303,834 potential SNPs.

Quality filtering left 168,599 SNPs with ≥ 5x coverage.

20,545 of the SNPs are covered by a common restriction enzyme site (AluI, BfaI, BstUI, DpnII, HaeIII, HhaI, MseI, MspI, RsaI, TaqI) present in Oakridge but not Mauriceville.

SNP validation

1) SNPs in Mauriceville EST data

95/100 SNPs screened are present in a BLAT aligned Mauriceville EST dataset.

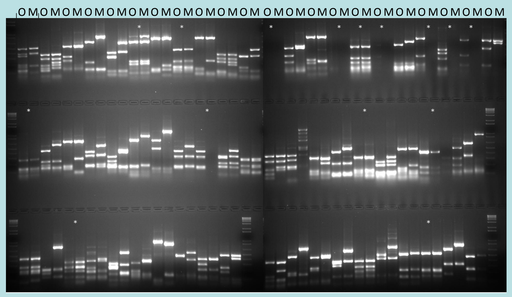
A verified C to G SNP

```
Mauriceville: 193 agagatgtgtattggtgtcaaagggcaatgactgatgattc 115
                  |||||| ||||||||||||||||||||||||||||||||||
Oakridge:  119025 agagatctgtattggtgtcaaagggcaatgactgatgattc 119065
```

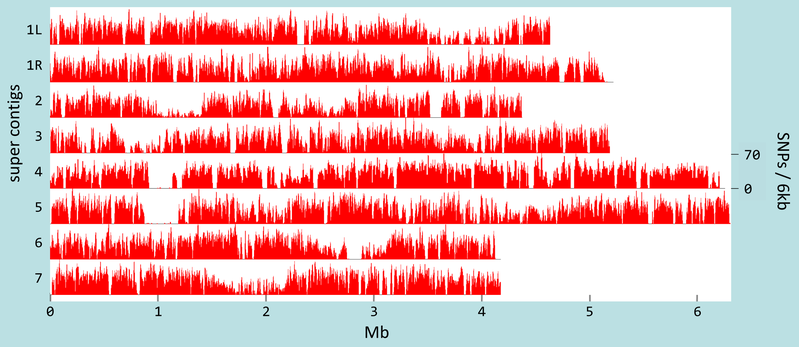
2) RFLP patterns after digesting PCR products covering TaqI digestible SNPs

48/48 SNP containing regions amplified by PCR from Mauriceville (M) and Oakridge (O) have the expected RFLP pattern.

*Ambiguous RFLP where the greatest difference in band size is < 10bp.



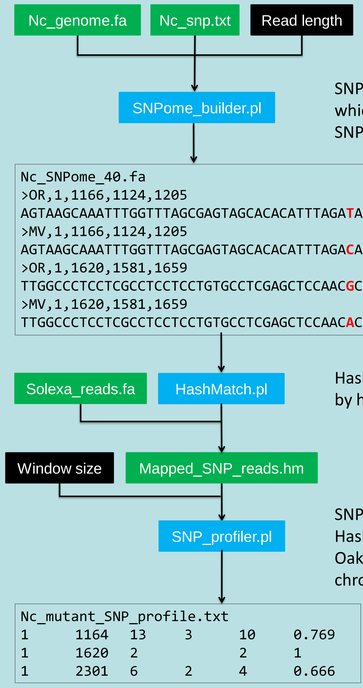
SNP coverage



Abstract

With the advent of high-throughput DNA sequencing it is now straight-forward and cheap to generate high-density small nucleotide polymorphism (SNP) maps that can improve on mapping strategies based on RFLP [1], CAPS [2,3] or RAD-mapping [4,5] analyses. Direct identification of single point mutations has been described in fission yeast [6] but in most organisms bulk segregant analyses followed by SNP mapping are used [7]. Here we present a high density SNP map between *Neurospora crassa* Mauriceville-1-c (FGSC2225) [8] and OR74A (FGSC987) [9], the strains most typically used by Neurospora researchers to carry out RFLP mapping crosses.

Mapping mutations by bulk segregant analysis using high-throughput sequencing

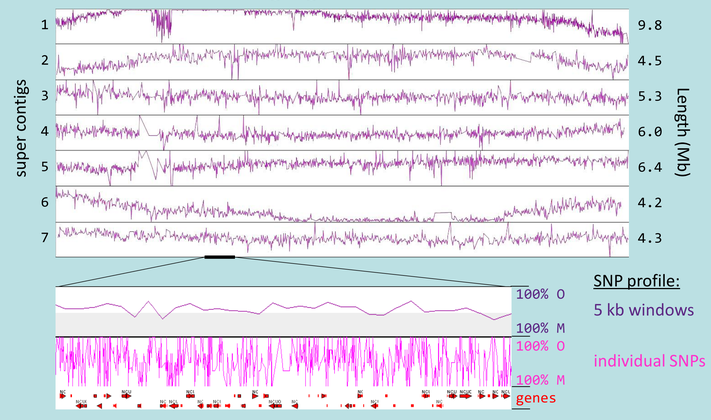


SNPome _builder.pl is used to build the SNPome which is a file containing each version of all the SNPs that will be used for mapping.

Hashmatch.pl [11] maps the short reads generated by high-throughput sequencing to the SNPome.

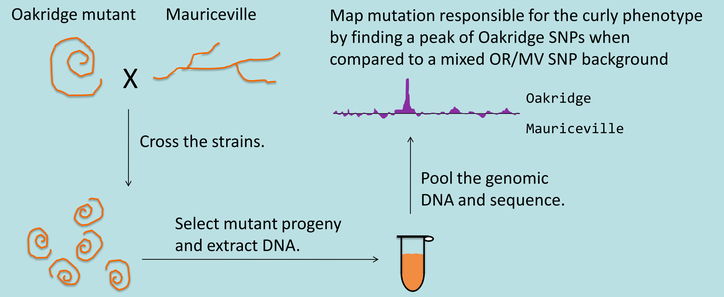
SNP_profiler.pl uses the alignment data from Hashmatch [11] to map the SNP ratio between Oakridge and Mauriceville along the chromosomes.

SNP profile in bulk segregants

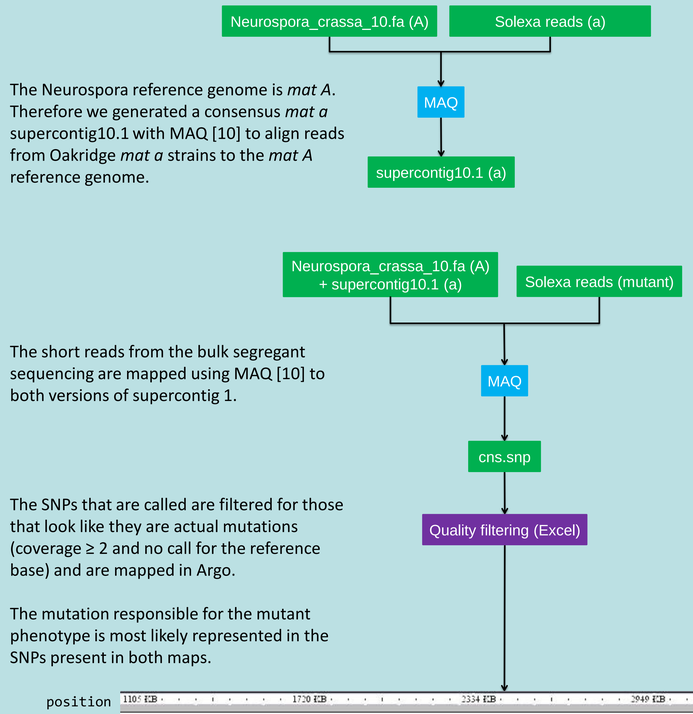


Eighteen *mat a* progeny were obtained from a mapping cross between an Oakridge mutant (O) and Mauriceville (M). Their DNA was mixed in equal amounts, sequenced on a Solexa Genome Analyzer and mapped using the pipeline described.

Mapping a mutation by bulk segregant analysis



Mapping a mutation linked to mating type

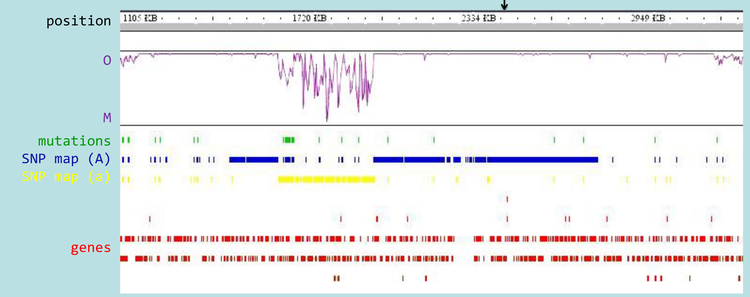


The Neurospora reference genome is *mat A*. Therefore we generated a consensus *mat a* supercontig10.1 with MAQ [10] to align reads from Oakridge *mat a* strains to the *mat A* reference genome.

The short reads from the bulk segregant sequencing are mapped using MAQ [10] to both versions of supercontig 1.

The SNPs that are called are filtered for those that look like they are actual mutations (coverage ≥ 2 and no call for the reference base) and are mapped in Argo.

The mutation responsible for the mutant phenotype is most likely represented in the SNPs present in both maps.



Acknowledgements

We thank Mark Dasenko and Chris Sullivan for assistance with Illumina sequencing. We also thank the Neurospora Functional Genomics Project (NIH P01 GM068087) for primers and the Fungal Genetics Stock Center for strains. This work is funded by start-up funds from the Computational and Genome Biology Initiative at OSU and grants from the American Cancer Society (RSG-08-030-01-CCG) and DOE (DE-FG02-08ER64665).

Citations

1. Metzenberg, RL et al. 1984. N. Newst. 31:35-39
2. Jin Y et Al. 2007. Fungal Genet Biol. 44(6):455-465
3. Lambreghts R et al. 2009. Genetics. 181:767-781
4. Lewis ZA et al. 2007. Genetics. 177:1163-1171
5. Baird NA et al. 2008. PLoS ONE. 3(10):e3376
6. Irvine DA et al. 2009. Genome Res. 19:1077-1083
7. Cuperus JT et al. 2010. PNAS. 107(1):466-471
8. Beauchamp et al. 1977. PNAS. 74:1172-1176
9. Case et al. 1965. Neurospora Newsl. 8:25-26
10. MAQ. http://maq.sourceforge.net/ accessed: 2/1/2010
11. Filichkin SA et al. 2010. Genome Res. 20:45-58