

# Potato killer! Digging dirt on its cousin...

## Introduction

The eukaryotic group of oomycetes includes several plant pathogens such as *P. infestans*, which caused the potato blight in the 19th century. Therefore, oomycetes possess great destructive potential to crops, and are of concern to industries. Chitin synthase (CHS) genes of oomycetes may also provide clues as to plant cell wall synthesis; additional assemblies will aid the comparative genomics investigation. *P. andina* Illumina reads (SRA\_SRR1817238) were used to prepare an assembly for this purpose.

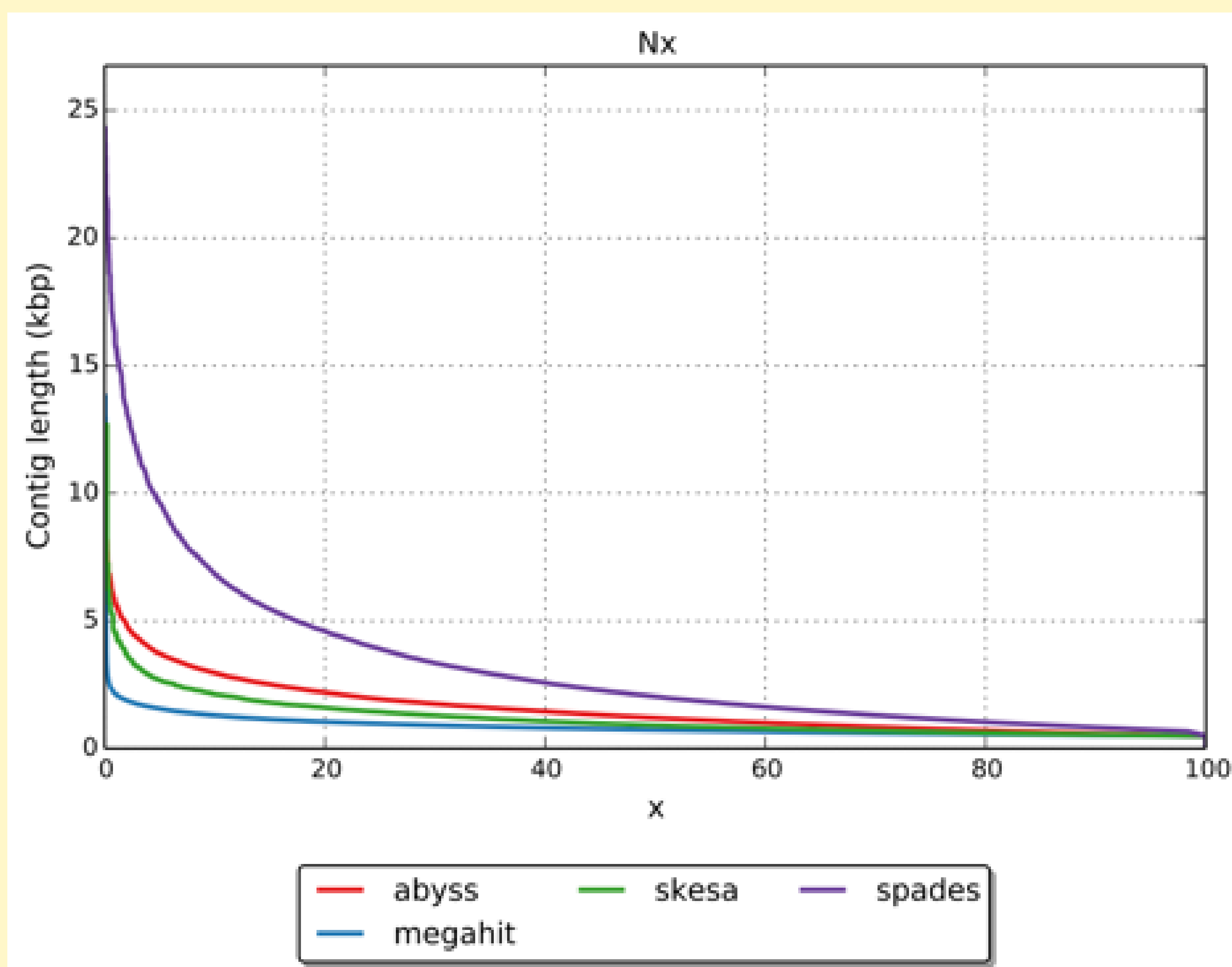
## Project Goals

- Assemble *P. andina* genome
- Identify & analyze chitin synthase (CHS) genes in assembly
- Comparison of largest contigs with *P. infestans*

## Tools

- Genome assemblers (SKESA, MEGAHIT, SPAdes, ABySS)
- Genome assembly validator - QUAST
- BLAST+ for gene comparison

## Picking an assembly



Assembly assessment with QUAST

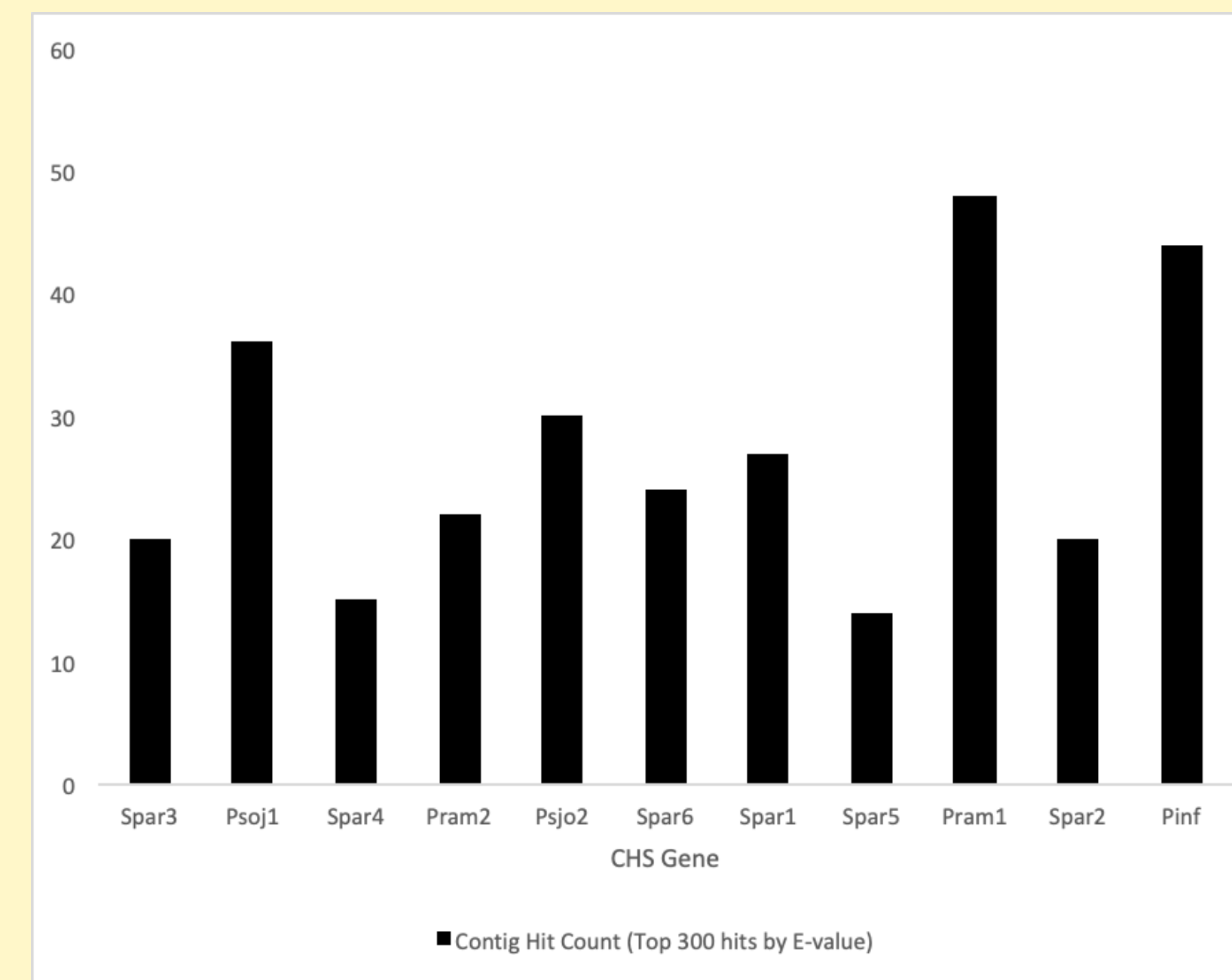
ABySS assembly selected; assembly size closest to desired size.

Assembly	k_30_3
# contigs (>= 0 bp)	4096572
# contigs (>= 1000 bp)	10687
# contigs (>= 5000 bp)	94
# contigs (>= 10000 bp)	0
# contigs (>= 25000 bp)	0
# contigs (>= 50000 bp)	0
Total length (>= 0 bp)	245128706
Total length (>= 1000 bp)	18395721
Total length (>= 5000 bp)	562604
Total length (>= 10000 bp)	0
Total length (>= 25000 bp)	0
Total length (>= 50000 bp)	0
# contigs	28029
Largest contig	9512
Total length	30384137
GC (%)	50.81
N50	1204

Selected Assembly: ABySS with k-mer 30

## Analyses of CHS genes

Using CHS reference files as a BLAST database, the assembly contigs were queried using BLASTx against this database. Hits were then sorted by E-values.



Top 300 contig hits to CHS genes

QSEQID	SSEQID	EVALUE
7586397	chs_Pinf	0.000
7586397	chs_Pram1	0.000
7586397	chs_Psoj1	0.000
7608936	chs_Pinf	0.000
100197	chs_Spar6	0.001
1021674	chs_Spar1	0.001
103572	chs_Pram2	0.001
1042318	chs_Pinf	0.001
1042318	chs_Pram1	0.001
1042318	chs_Psoj1	0.001
1046848	chs_Psoj2	0.001

Top 10 contig hits

```

> 7586397 2177 32954 1557734-,...,4032401+
#Length=2177

#Score = 1205 bits (3118), Expect = 0.0, Method: Compositional matrix adjust.
# Identities = 590/591 (99%), Positives = 590/591 (99%), Gaps = 0/591 (0%)
# Frame = -1

Query 1 MSGAPPPSSGFQPRSIGLPLPSHGPRSSMMSVEYDGIPLPPPSIRSCGSQQYVTSYIPTG 60
      |||
Sbjct 1775 MSGAPPPSSGFQPRSIGLPLPSHGPRSSMMSVEYDGIPLPPPSIRSCGSQQYVTSYIPTG 1596

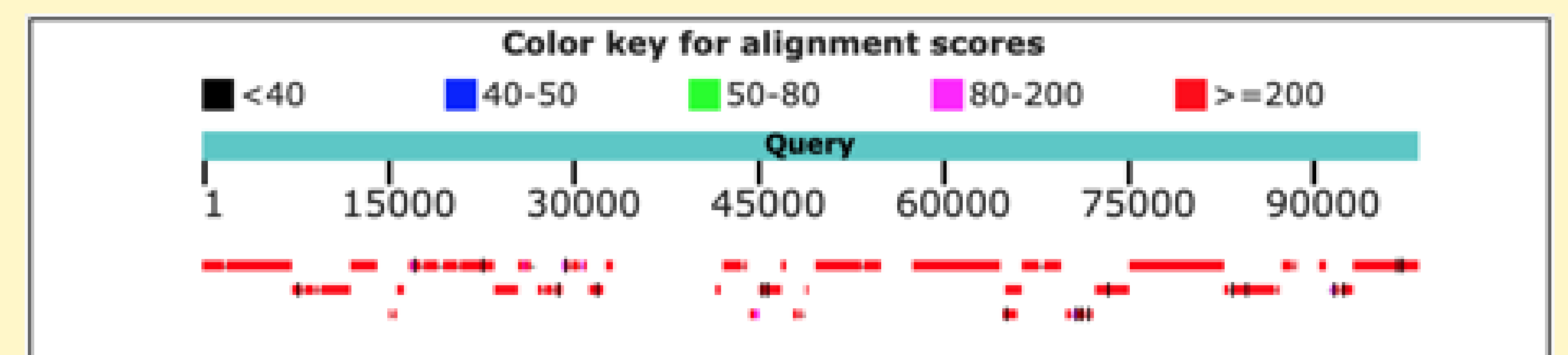
Query 61 AAFPSGVSQDLISSMKSYSATDLVRTYSEIPSVEEALVTLDRAAVALGARRYRDALKLY 120
      |||
Sbjct 1595 AAFPSGVSQDLISSMKSYSATDLVRTYSEIPSVEEALVTLDRAAVALGARRYRDALKLY 1416

```

Top contig alignment with chs\_Pinf

## Analyses vs *P. infestans*

Contigs longer than 7000bp (total: 12) were queried against corresponding parts of the *P. infestans* genome.



The best-scoring genes from *P. infestans* with known gene products which mapped with an E-value of 0 and >97% identity are:

- Dynein heavy chain (PITG\_05922)
- Sporangia induced dynein heavy chain (PITG\_12577)
- Voltage-gated ion channel superfamily (PITG\_16108)
- ATP-dependent RNA helicase (putative) (PITG\_11000)
- Endo-1,3(4)-beta-glucanase (PITG02219)
- Replication protein A 70 kDa DNA-binding subunit
- Ubiquitin activating enzyme E1 family
- Periodic tryptophan protein 2

## Outlook

Resolution of the ancestry of *P. andina* is necessary for the interpretation of its population structure, evolution, and genetics. To identify the ancestors of *P. andina*, more gene loci should be sequenced and compared. Haplotype inference can be performed for genetic recombination information. Phylogenetic analysis across the *Phytophthora* genus can provide a more detailed evolutionary history of *P. andina*.