# Next Generation Genomic Analysis of Achondroplasia in White Pekin Ducks

Paige Richards & Dr. G.F. Barbato
Biology Program of Stockton University, Galloway NJ

## Abstract

Last year, an autosomal dominant mutation, resulting in achondroplasia, was identified in White Pekin ducks. In an attempt to identify the molecular origin of the mutation genome sequence analysis was performed. Genomic DNA was collected and pooled from 2 pureline families (1 male and 6 females) of both achondroplastic and wild-type adult ducks. The genomes were sequenced at North Carolina State University, using the Illumina NextSeq 75 PE. Afterward, they were assembled using Galaxy at Penn State against both builds of the Pekin Duck genome (BGI_duck_1.0/anaPla1 and GCA_0002743455.1_CAU_duck1.0). Further analyses were performed using both Integrative Genomics Viewer (IGV) and Genomic Workbench, where the first three chromosomes were investigated. Chromosome 2 (PEDO01021441.1) displayed differences between families at the following location of base pairs: 1574-1890, 1469-1544, and 730-860. While chromosome 3 (PEDO01021440.1) displayed differences between families at the location of base pairs: 838-865, 1345-1368, 3205-3280 and 3937-3972. Another approach was to identify any FGF4 homologies (KB743402 and KB742617) and FGF3 homologies (KB743402) in the sequences, as dominant mutations in the FGF genes are known to cause achondroplasia in both human and canine families. Chromosome 1 (PEDO01021442.1) has homology to mammalian and chicken FGF4, and the duck genome exhibits variation at contigs KB743402 at the 1021bp and KB742617 at the 236bp. Having identified these differences, the next step is to analyze whether they play a role in the resulting achondroplasia variant.

## Introduction

In working with commercial stocks of White Pekin ducks, Dr. Guy F. Barbato noticed the appearance of a shortened shank length amongst them. Further investigation of this 'dwarf" phenotype led to the discovery of an autosomal dominant mutation, that results in achondroplasia, in White Pekin ducks. Prior research shows no relationship between the gene and body weight or sex (Barbato, G. F. and Z. Lowman 2018). Therefore, this project sought to sequence the genomes of two pure line families and use genomic analytical software to identify any genomic differences between achondroplastic and wild-type adult ducks. The genomes were searched for any FGF4 and FGF3 homologies because dominant FGF mutations have been known to cause chondrodystrophy in both humans and canines (Nat Biotechnol 2011).

## Methods & Materials

Genomic DNA was collected and pooled from 2 pureline families (1 male and 6 females) of both achondroplastic and wild-type adult ducks. The genomes were sequenced at North Carolina State University using the Illumina NextSeq 75 PE. Upon receipt of the data, Galaxy (https://usegalaxy.org/) was used to assemble the genomic data against both builds of the Pekin Duck genome (BGI_duck_1.0/anaPla1 and GCA_0002743455.1_CAU_duck1.0). Integrative Genomics Viewer (IGV) and Genomic Workbench were the genomic analytical software used to compare the first three chromosomes from all families. Ducks have 80 chromosomes. Due to time constraints and limitation in operating the software only the first three chromosomes were analyzed for differences (Nat Biotechnol 2011).

Differences were identified using three techniques. The overview of coverage graphs, which plots how many times a sequence was read at a specific location along the genomic scaffold. The analysis of graphical sequence views, which displays the ordering of base pairs and their location within that chromosome. The identification of any FGF4 homologies (KB743402 and KB742617) and FGF3 homologies (KB743402) in comparison to the (GCA_0002743455.1_CAU_duck1.0) reference genome provided by the NCBI database.

For identification purposes, the wild type genomes were referred to as Long 1 and Long 2 while the achondroplastic genomes were referred to as Short 1 and Short 2. Among the four genomes sequences, two types of differences were identified. One difference noted was when genomic data for a specific location was present in both short sequences, but missing in the long sequences or vice versa. While the second type of difference occurred when all four genomes had genomic information at a specific location. However, both sequences of either short or long variants displayed fewer reads in comparison to the other phenotype.

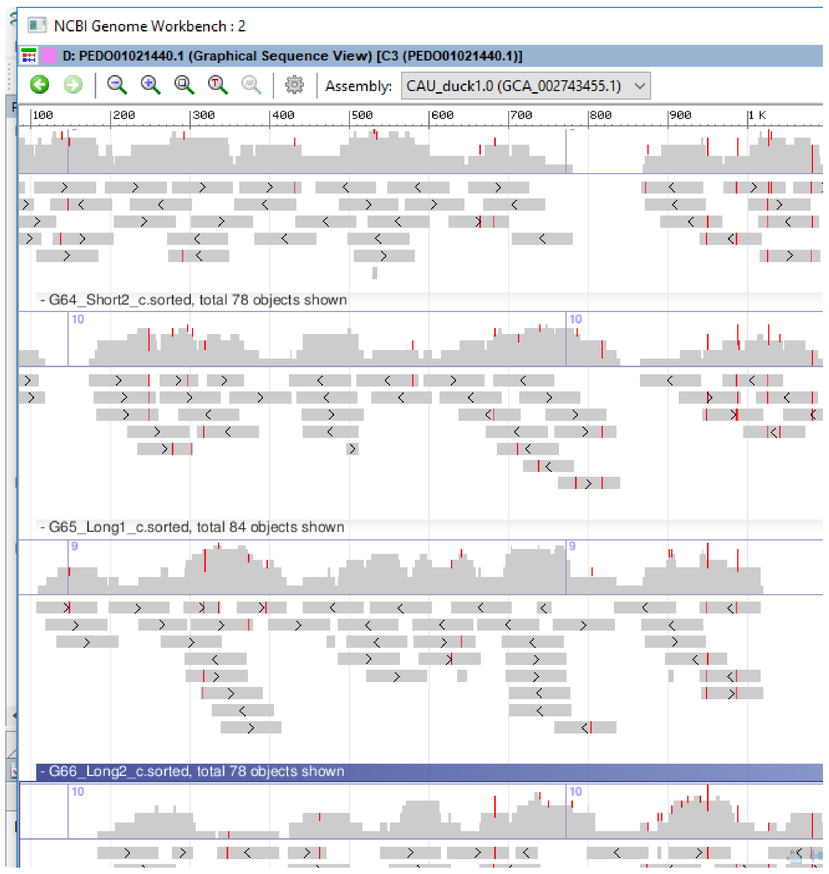


Figure 1: Displays a graphical sequence view of the third chromosome (PEDO01021440.1). This example shows a difference between the sorted short and long genomes at the 838 -865 base pair location. Short 1 and 2 are missing genetic information, while Long 1 and 2 display one read at this location.

## Results & Discussion

The graphical sequence view approach was used on Genomic Workbench. This technique revealed that Chromosome 2 (PEDO01021441.1) displayed differences between families at the following location of base pairs: 1574-1890, 1469-1544, and 730-860. While chromosome 3 (PEDO01021440.1) displayed differences between families at the location of base pairs: 838-865, 1345-1368, 3205-3280 and 3937-3972. The difference in the number of reads at each location can be found in Table 1.

Genomic Workbench was used for the identification of FGF homologies. This approach showed that Chromosome 1 (PEDO01021442.1) exhibits the FGF4 homologies KB743402 at the 1021bp and KB742617 at the 236bp. The presence of these KB sequences indicates that the duck genome has homology to mammalian and chicken FGF4.

The differences where genomic data for a specific location was present in both short sequences, but missing in both long sequences, or vice versa, are key locations to investigate in the future. This type of difference supports the idea that genomic information could either be abnormally present or missing from the genomes of Short 1 and 2 to result in an achondroplastic variant. Nevertheless, the role of differences, whose number of reads between Short and Long genomes vary for a specific location, should also be investigated. This type of difference supports the idea that more or less expression of genomic information at a specific location might contribute to the resulting achondroplastic variant.

Table 1: Displays the differences found between the short and long genomes when analyzing the graphical sequence views on Genomic Workbench. Differences are listed by which chromosome they were found on and how many reads of genomic information are present at a specific base pair location along the scaffold.

| Table 1 | | Number of Reads | | | |
|---|---|---|---|---|---|
| Chromosome | Base Pair Location (bp) | Short 1 | Short 2 | Long 1 | Long 2 |
| 2 | 1574 - 1680 | 7 | 6 | 0 | 4 |
| 2 | 1754 - 1890 | 4 | 6 | 6 | 0 |
| 2 | 1469 - 1544 | 6 | 6 | 1 | 3 |
| 2 | 730 - 860 | 8 | 7 | 1 | 0 |
| 3 | 838 - 865 | 0 | 0 | 1 | 1 |
| 3 | 1345 -1368 | 0 | 0 | 2 | 8 |
| 3 | 3205 - 3280 | 4 | 4 | 11 | 7 |
| 3 | 3937 - 3972 | 4 | 3 | 0 | 0 |

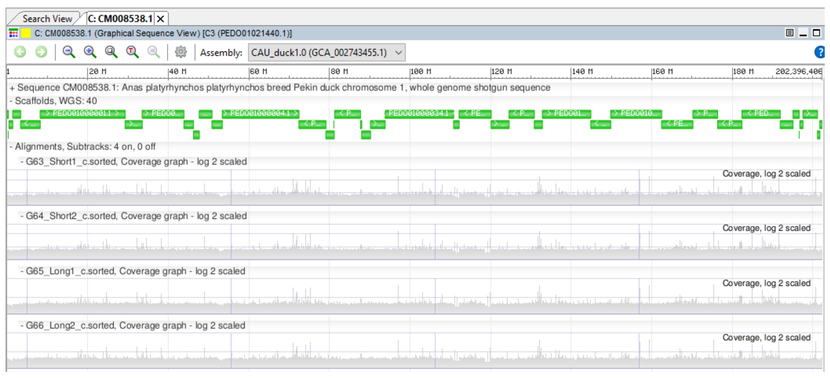


Figure 2: Displays a coverage graph of the third chromosome (PEDO01021440.1). All of the genomes within this figure have a similar shape, thus no obvious location is showing more reads on a specific genome in comparison to the others.

## Conclusion

Although there are no physical differences in body weight associated with autosomal dominant achondroplasia in White Pekin ducks, the genomic analysis had identified possible locations in which these variants differ. Given more time, further analysis would allow for the identification of more differences over a larger sample size of chromosomes. Preceding research could then investigate what role, if any, these differences play in the achondroplastic variant.

References:

Barbato, G. F. and Z. Lowman (2018) Identification of an autosomal dominant achondroplasia in White Pekin ducks. Poultry Science 97(E-Suppl.1): 152.

Nat Biotechnol. 2011 Jan; 29(1): 24–26. doi: 10.1038/nbt.1754