**Aneuvis: Reproducible web-based visualization of aneuploidy in single cells**

Daniel G. Piqué1,2, Samuel E. Zimmerman1, Grasiella A. Andriani2, Elaine Maggi2, John M. Greally2, Cristina Montagna2,3, and Jessica C. Mar1,4,5\*

1Department of Systems and Computational Biology

2Department of Genetics

3Department of Pathology

4Department of Epidemiology and Population Health

Albert Einstein College of Medicine, Bronx, NY, 10461, USA

5Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, QLD 4072, Australia

**Abstract**

Aneuploidy is a predisposing factor for cancer. There are traditionally two main methods for detecting aneuploidy in research and clinical practice. Molecular cytogenetic approaches, often used in clinical practice, involve direct visualization of interphase or metaphase preparations of chromosomes using fluorescent probes. More recently, whole genome sequencing has been utilized to discern chromosomal copy number at high resolution. Each of these approaches has their advantages and disadvantages, and an integrated analysis would aid in the better understanding of the strengths and limitations of these methods. Here, we develop a web-based tool, termed Aneuvis, that allows users to upload molecular cytogenetic or processed whole-genome sequencing data. Aneuvis automatically constructs visualizations and summary statistics and generates a reproducible report that can be used to customize output. We present the first web-based tool that incorporates both molecular cytogenetic and whole genome sequencing data. Aneuvis will highlight molecular heterogeneity both within and between different approaches and help distinguish cell populations.

**Introduction**

In humans, alterations in chromosome number are a source of congenital disease and a predisposing factor for carcinogenesis. Increased variability in the karyotypes of cancer cell populations is also associated with resistance to chemotherapy and poor clinical outcomes. Alterations in chromosome number is also a sign of chromosomal instability, which can lead to cancer therapeutic resistance. Alterations can take place as departures from an integer-valued set, known as polyploidy. There are two major methods for detecting alterations in chromosome number --

In a cell, there are three possible states that a set of chromosomes can have. Diploidy refers to the presence of two copies of each chromosome. Polyploidy refers to an integer-values increase in the number of chromosomes, often resulting from whole-genome duplication. The presence of WGD alters sensitivity of cells to potential cancer therapeutics and is thus important to detect1,2.

Instavility

Invisibility

SUMCIN CIN-VIS

*Whole chromosome instability*

Whole chromosome instability is thought to arise from the missegregation of chromosomes during mitosis.

*Drivers of aneuploidy*

*Measuring aneuploidy*

Whole genome sequencing

FISH

Fluorescence in situ hybridization is an approach that uses fluorescently-labeled DNA-based probes to

Chromosomal copy number variation variation Aneuploidy is The drivers of aneuploidy are unclear. Understanding how molecular stressors are

In a cell, there are three possible states

**Results**

Aneuvis overview.

The idea behind this web application is to take in copy number data from different treatment groups and to output a summary of the relationship between the treatment groups with each other.

There are three types of single-cell chromosomal data that can be uploaded into aneuvis. First, fluorescence in situ hybridization (FISH) data, where the chromosome number is inferred from the number of distinct fluorescent probes, can be represented as a matrix where each column is a chromosome and each row is a separate biological cell. The numbers in the matrix represent the number of copies of a given chromosome for a given cell. The second type of data is single cell whole genome sequencing data. There are currently user-friendly programs, such as Ginkgo, for converting .bam files containing aligned reads obtained from DNA-sequencing to copy number data. Aneuvis will convert output in .bed format to a summarized copy number state using a weighted mean, where the copy number at each bin contributes proportionally to its overall size.

We acknowledge several limitations of aneuvis. First is the ability to incorporate sub-chromosomal events.

Analyze chromosomal counts from 2 to 4 chromosomes. Single cell whole genome sequencing (SC-WGS) analyze chromosome counts from single cell sequencing data." Spectral karyotyping (SKY)"), "- analyze chromosome counts and structural variation from all chromosomes.")

),

these datatypes. visualization and summary statistics three kinds of copy number data – single cell,

Aneuvis highlights previously unseen variation in chromosome counts

Comparing the degree of aneuploidy between predefined treatment groups

Summary statistics

Anevis compares the degree of aneuploidy between predefined treatment groups

displays new findings