**Introduction** - Polyploidy, caused by whole-genome duplication (WGD), is the presence of more than two genome copies per somatic cell. These duplications dramatically change an organism’s genome, and are important drivers of lineage divergence and evolutionary novelty. Near-ubiquitous in the plant kingdom, WGD is increasingly recognized throughout eukaryotes. Despite a century of study and high ecological, evolutionary, and economic importance, there remain substantial unresolved questions in our understanding of WGD. My research program brings together organismal, genomic, cytometric, and bioinformatic approaches to investigate three of these major questions, detailed below:

**How does polyploidy affect speciation?** - Whole-genome duplication can rapidly induce strong reproductive isolation between plants of different genome copy numbers (i.e. different cytotypes) and until recently was widely assumed to be an “instant” speciation process. My previous work examined this assumption using the *Campanula rotundifolia* polyploid complex, which comprises three dominant cytotypes: diploids, tetraploids, and hexaploids. I demonstrated that reproductive isolation is nearly complete between diploids and tetraploids, but significantly lower between tetraploid and hexaploid cytotypes, and that germination drives this differential isolation. In field studies, I found that pollinators do not exhibit specific cytotype preferences; instead, they prefer rare plants over common ones regardless of cytotype. Using backcrosses, flow cytometry, and surveys of natural mixed-ploidy populations, I found that pentaploids and putative aneuploids are common when tetraploids and hexaploids occur in sympatry. This work challenges the long-standing perception that change in cytotype incurs de facto reproductive isolation.

**How has ancient whole-genome duplication affected eukaryote evolution? –** AfterWGD, genomes begin losing, modifying, and rearranging duplicate genes via diploidization. This results in extant diploid organisms that have polyploid evolutionary histories. The biological processes that affect polyploidization and diploidization rates across eukaryotes are not well understood, nor are the mechanisms which mediate development and maintenance of higher polyploids (i.e. above tetraploidy). Diploidization obscures evidence of ancient WGD, making these mechanisms difficult to study. Ancient WGD can be inferred from genomic or transcriptomic data in extant diploids, but current inference methods are often difficult to replicate or interpret. To address this difficulty, I developed the Supervised Learning Estimation of Duplicated GEnomes (SLEDGE) toolkit, a set of simulated genomes and classification algorithms that screen transcriptome data for evidence of ancient WGD. This tool streamlines identification of ancient polyploids across eukaryotes.

**How does polyploidy facilitate long-range establishment and invasion? –** Whole-genome duplication often induces traits amenable to long-range dispersal and establishment, including increased selfing, greater phenotypic plasticity, and more adaptive potential—all of which also promote invasiveness. By characterizing adaptations which facilitate dispersal and establishment in polyploids, we can better understand the evolutionary histories of polyploid taxa, and better predict and manage the invasive plant emergences.

Selfing often arises during population bottlenecks experienced by newly-formed polyploids, but some work shows WGD itself can increase selfing. I investigated how mating system has evolved in *C. rotundifolia*. Prior literature claimed the *C. rotundifolia* aggregate were self-incompatible, but some accounts note some autogamous seed set. Using self- and outcrosses in populations spanning both the geographic and cytotypic diversity of the complex, we determined self-incompatibility decreases both with WGD and long-distance dispersal.

***Current research*** – My current work investigates the role of polyploidy on dispersal and establishment of invasive plants as research scientist for the Consortium for Plant Invasion Genomics (CPING). Using five invasive species across green plants, we assess colonization dynamics, roles of hybridization and polyploidy, and adaptive mechanisms. My primary focus is the pentaploid aquatic fern, *Salvinia molesta,* a noxious freshwater weed in Asia, Africa, and North America. *Salvinia molesta* is considered obligately asexual; however, the lineage has surprisingly high genetic diversity in the southeastern U.S. Along with CPING collaborators, I have sampled contemporary populations of *S. molesta* along the U.S. Gulf Coast as well as a time-series of herbarium specimens spanning the invasion in the U.S. I have developed probes for reduced-representation sequencing via hybridization (hyRAD) to assess genetic diversity across the U.S. invaded range. These probes are designed to work across multiple CPING focal species, allowing later comparison of evolutionary and adaptive histories across invasive taxa.

**Short and long-term research goals at George Mason University**

***Campanula rotundifolia* -** The questions that comprise my research program are varied and allow for a flexible, scaffolded approach that can respond to changes in lab personnel and funding. I have maintained seed stocks from over 20 well-characterized *Campanula rotundifolia* populations throughout North America that can be used to quickly begin collecting phenotype and ploidy level data at George Mason, and existing field sites in southwestern Virginia, Maryland, and Pennsylvania present tractable, low-cost locations for graduate and undergraduate field work. Longer-term, confirming and mapping cytotype differences across known mixed-ploidy field sites in Canada and the United Kingdom would be a tractable project for students at multiple levels. Screening and cytotyping these populations would work well as a Research Semester projects, while characterizing the frequently observed aneuploidy and linking chromosome dosage to fitness is a longer-term project suitable for one or more PhD students—this would necessitate fine-scale collections, flow cytometry, and field observations to fully map euploid and aneuploid plants, growth of seeds in common garden to establish viability and fertility of aneuploids, and ongoing genotyping to track interploidy geneflow in situ.

Planned longer-term work includes investigation of the self-incompatibility mechanism within the *C. rotundifolia* complex. While work in congeners suggests that *C. rotundifolia* uses gametophytic self-incompatibility, the S-locus has not been characterized in this system. This would require generation of whole-genome sequence data from all three dominant cytotypes—this is aided by diploid specimens I have already collected and the presence of a draft genome for the congener *C. gentilis*. This project will entail generating, assembling, and annotating a draft genome, developing S-locus markers, genotyping individuals throughout the North American and European ranges, and generating neotetraploid and neohexaploid individuals. This work likely spans 5-7 years and represents 2-3 PhD-level projects.

**Machine Learning Assessment of Polyploids and Aneuploids -** The variability of polyploid incidence across plants and, more broadly, eukaryotes has been a long-standing evolutionary question. Differences in whole-genome duplication and diploidization rates may explain why polyploidy is common in some lineages and absent in others.My lab will make two big expansions of the SLEDGE toolkit—enabling it to screen for extant as well as ancient polyploidy, and enabling screening for extant aneuploidy. This work is ongoing and will continue in my lab at George Mason. Although the computing requirements for SLEDGe are comparatively modest, I am excited about using GMU’s Argo and Hopper clusters to accelerate this work and to train students to work in a computing cluster environment. Because all of my research aims involve bioinformatic approaches, I will adapt modules I helped develop as part of CPING to teach undergraduate and graduate students basic command-line computing and common bioinformatic analyses.

**Invasion Biology -** At GMU, I will continue the work I began with CPING on invasive ferns *Salvinia molesta* and *S. minima*. Depending on ongoing COVID restrictions and University policies, I would include GMU students in plant collections or alternatively import dried leaf tissue for further study. Additionally, I plan to begin preliminary investigations into the weedy creeping bellflower, *Campanula rapunculoides*. Although this species is not nationally-listed as an invasive in Canada or the U.S., is it listed as a state or provincial invasive in Minnesota and Alberta. This species is well-suited for work on invasions in my lab because it is a close analog of *Campanula rotundifolia*. Both species are polyploid complexes with dominant diploid, tetraploid, and hexaploid cytotypes, both have gametophytic self-incompatibility systems, and both are inferred to have originated in southeastern Europe. Despite these similarities, these species have very different histories in North America. Preliminary work on this system would include regional surveys of cytotype distribution and the development of reduced-representation markers using the hyRAD system similar to that I have applied in Salvinia. These early investigations would make ideal undergraduate or masters-level projects.