What's in a Genome? A Look at Invasive Plant Species
Cassandra Hash

Faculty Mentor: Dr. Tara Pelletier

Presented: 2023 Fall Winter CARD

Abstract

Invasive plant species thrive to the detriment of the ecosystems in which they are introduced (Zhang et al., 2021). They are difficult to manage due to how prolific they grow and how costly they are to remove. I studied the genomic characteristics of invasive vs. non-invasive plant species to determine what—if any—aspects of plant genomes contribute to whether or not a species was invasive. I gathered data from genome assemblies within open-source genome repositories and research papers for 15 non-invasive plant species and 19 invasive plant species. I compared the size and the components of the genomes (e.g., the number of coding genes, the number of chromosomes, the percentage of repetitive and transposable elements, etc.) of the invasive plant species to those of the non-invasive plant species. My results indicate that these characteristics are not statistically different for invasive vs. non-invasive plant species. These findings indicate that the aspects of plant genomes that impact invasiveness may be more specific. Comparing these characteristics in invasive and non-invasive plant species found in the same location or climate may garner some significant results; however, the scope of a study like that would not reflect invasive plant species broadly. Comparing specific genes found in invasive vs. non-invasive plant species will likely reveal specific candidate genes or gene-gene interactions that may contribute to invasiveness. If we can identify these genes—or any other aspect of plant genomes—that contribute to a plant's invasiveness, we may eventually be able to alter these characteristics to prevent invasive plant species from successfully establishing themselves in the ecosystems they harm. My project is an initial step towards that outcome.

Methods

      I gathered genomic data from online open-source genome repositories and research papers for 15 non-invasive plant species and 19 invasive plant species, spanning 13 families (Table 1). I created a dataset of nine characteristics for plant species and analyzed five of these genome-specific characteristics for which there was complete data—the size of the genome, number of chromosomes, number of coding genes, percent of transposable elements (TE) in the genome, and percent of repetitive elements (RE) in the genome. I performed T-tests to test my hypothesis that these characteristics would be significantly different for invasive and non-invasive plant species.

Table 1: # Species sampled per family.		
Family	**Non-Invasive Species**	**Invasive Species**
Amborellaceae	1	0
Asteraceae	1	1
Brassicaceae	0	3
Cannabaceae	1	0
Caprifoliaceae	0	2
Fabaceae	1	0
Lauraceae	1	0
Paulowniaceae	1	1
Poaceae	5	8
Polygonaceae	0	1
Rosaceae	2	1
Solanaceae	2	1
Typhaceae	0	1
Total:	15	19

Species Selection

      In short, I built my sample of plant species by determining what species were invasive or non-invasive and had accessible genome assemblies. I used the standard definition for invasive species: a non-native species that is able to thrive and reproduce in areas they are introduced to, altering natural processes and killing the native wildlife at rates that are difficult to manage (Zhang et al., 2021). I predominantly identified plant species by referencing lists of species within various plant families and by searching state listings of invasive plant species. Once I found a species, I checked to see if I could access its genome assembly; if a genome assembly was available, I then reviewed the literature and cited papers that confirmed the species was invasive or non-invasive (Table 2). I excluded species that lacked sufficient research that clearly categorized them as invasive or non-invasive.

      Building a large enough sample was challenging because primary literature that clearly categorizes plant species as invasive or non-invasive was difficult to obtain for many species—even when species were listed as invasive in certain locations—and the lack of clarity was partly caused by inconsistent definitions of invasive plant species, with some calling any weedy or non-native plant species invasive. Despite this challenge, I obtained a satisfactory sample size (Table 1). Once I had selected 19 invasive plant species from various families, I focused on matching that number for non-invasive species within those families. Ultimately, I managed to find 15

non-invasive plant species. The large sample size made for a stronger statistical analysis, and finding species from various families helped me account for any genomic similarities that species from the same families may possess.

Data Sourcing

My goal in the beginning was to derive my data solely from plant genome databases. I made a Google search to identify open-source genome databases that contain large numbers of genome assemblies for various plant species. Many databases were ruled out because the type of data they contained (e.g., c-values) is irrelevant to this study, and others were ruled out because they contained a limited number of plant species and/or data. My final selection of databases included Ensembl Plants (https://plants.ensembl.org/index.html) and Brassica.info (http://www.brassica.info/genome/genomes.html). However, these databases ultimately contained too few plant species—especially invasive plant species—and I expanded my sourcing to published nuclear genome assemblies (Table 2). To accomplish this, I identified invasive and non-invasive plant species and searched for their genome assemblies. For example, to find a genome assembly for the invasive species *Lonicera japonica*, I Googled "*Lonicera japonica* genome assembly" and found a published paper containing the statistics for the species' assembled nuclear genome. I excluded any species that only possessed chloroplast or transcriptome assemblies, and I included chromosome-level genome assemblies when available.

Dataset: Characteristics and Data Collection

The initial goal of my study was to compare the genomes of invasive vs. non-invasive plant species to identify some candidate genes that may contribute to a plant's invasiveness. I intended to use genome databases to identify these potential candidate genes. However, the complexity of such a task paired with the amount of time I had to work on this project led me to revisit my research goal.

After some troubleshooting and experimenting with different methods and research questions, I decided to study how certain characteristics of plant genomes present in invasive vs. non-invasive plant species. I noted what statistics were available throughout the genome assemblies, then recorded data for nine characteristics of plant species (including native geographical distribution, non-native geographical distribution, number of non-coding genes, number of gene transcripts, genome size, number of coding genes, number of chromosomes, percent of TE in genomes, and percent of RE in genomes) and analyzed the last five

characteristics because they specifically relate to species' genomes. I chose these five characteristics because they were recorded for most genome assemblies within my sample and because they impact growth, reproduction, and trait expression (Comai, 2005; Muñoz-López & García-Pérez, 2010; Suda et al., 2014). For example, genome size greatly impacts growth, reproductive success, and dispersal, and some researchers hypothesize that smaller genomes result in more trait states than larger genomes, which can also impact a plant's ability to succeed in various environments (Suda et al. 2014). Thus, genome size may impact a plant's ability to become invasive in their non-native habitats.

I pulled the data for these characteristics from the genome assemblies I gathered for each species and reported the numbers in an Excel datatable (Table 3). Locating this data in the text was challenging for several species because some studies either excluded the statistics for some characteristics or reported vague numbers. For example, one paper included the number of genes found in the genome, but it was unclear if those were protein-coding genes specifically (Kesel et al., 2022). Another paper used transposable and repetitive elements interchangeably (Rajewski et al., 2021). When I couldn't verify that the numbers corresponded to the specific characteristics I studied, I excluded the numbers from the dataset. I was still able to include about 15 samples per characteristic for both invasive and non-invasive plant species.

Once I had obtained a complete dataset, I performed two-sample T-tests in Excel for each characteristic and recorded the mean, standard deviation, and p-values. Then, I tried out multiple plots to determine what would best represent my results, including bar charts, histograms, and box-and-whisker plots. Ultimately, I made five box-and-whisker plots for the data.

Table 2: Papers referenced for the dataset.	
Species	**References**
Amborella trichopoda	(Albert et al., 2013; Chamala et al., 2013)
Ambrosia artemisiifolia	(Bieker et al., 2022)
Erigeron breviscaupus	(Yang et al., 2017)
Alliaria petiolata	(Alabi et al., 2021)
Brassica juncea	(Kang et al., 2021; Semenya & Maroyi, 2020)
Brassica nigra	2015)
Humulus lupulus	(Jeong et al., 2018; Natsume et al., 2015)
Lonicera japonica	(Pu et al., 2020; Wang et al., 2012)
Lonicera maackii	(Kesel et al., 2022)
Sophora flavescens	(Qu et al., 2023)
Lindera glauca	(Xiong et al., 2022)
Paulownia fortunei	(Cao et al., 2021)
Paulownia tomentosa	(Zhao et al., 2017)
Aegilops tauschii	(Luo et al., 2017; Yang et al., 2023)
Avena sativa	(Kamal et al., 2022)
Brachypodium distachyon	(Bakker et al., 2009; The International Brachypodium Initiative, 2010)
Coix lacryma-jobi	(Kang et al., 2020)
Dactylis glomerata	(Huang et al., 2019)
Echinochloa crus-galli	(Gu et al., 2023; Guo et al., 2017)
Eragrostis curvula	(Carballo et al., 2019; Firn et al., 2012)
Eragrostis tef	(Cannarozzi et al., 2014)
Miscanthus sinensis	(Mitros et al., 2020)
Stipa capillata	(Baiakhmetov et al., 2021)
Triticum turgidum	(Avni et al., 2017)
Zizania palustris	(Haas et al., 2021)
Zoysia japonica	(Tanaka et al., 2016)
Polygonium cuspidatum	(Quinty et al., 2022; Zhang et al., 2019)
Malus domestica	(Daccord et al., 2017)
Prunus persica	(The International Peach Genome Initiative et al., 2013)
Rosa multiflora	(Chung et al., 2014; Nakamura et al., 2018)
Capsicum annuum	(Hulse-Kemp et al., 2018)
Datura stramonium	(Chadha et al., 2020; Rajewski et al., 2021)
Solanum lycopersicum	(Su et al., 2021)
Typha latifolia	(Widangama et al., 2022)

Results

I obtained enough data to analyze five characteristics of plant genomes. The other characteristics remained in the original dataset, but there wasn't enough data to include them in my statistical analyses. The dataset for the main five characteristics is shown below (Table 3). Genome size was collected for 14 non-invasive plant species and 18 invasive plant species. The percent of TE within genomes was collected for 15 non-invasive plant species and 14 invasive plant species. The percent of RE within genomes was collected for 13 non-invasive plant species and 15 invasive plant species. The number of coding genes was collected for 15 non-invasive plant species and 18 invasive plant species. Lastly, the number of chromosomes was collected for 13 non-invasive plant species and 13 invasive plant species.

Table 3: Simplified dataset for the five genome characteristics studied. The blue cells represent the data that wasn't obtainable.

Family	Species	Status	Size (Mb)	Coding Genes	% TE	% RE	Chromosomes
Amborellaceae	*Amborella trichopoda*	Non	748	27313	57.20%	57.20%	7
Asteraceae	*Ambrosia artemisiifolia*	Inv	1579.1	34066	13.97%	30.76%	
Asteraceae	*Erigeron breviscapus*	Non	1520	37504	52.22%	54.58%	
Brassicaceae	*Alliaria petiolata*	Inv	1350	64770			
Brassicaceae	*Brassica juncea*	Inv	933.5	92887	50.36%	~50.36%	18
Brassicaceae	*Brassica nigra*	Inv	607.8	67030	53.73%	~53.73%	8
Cannabaceae	*Humulus lupulus*	Non	2570	41228	34.64%	34.68%	10
Caprifoliaceae	*Lonicera japonica*	Inv	887.15	33961	58.21%	58.21%	9
Caprifoliaceae	*Lonicera maackii*	Inv	2270				9
Fabaceae	*Sophora flavescens*	Non	2100	60485	83.06%	80.00%	9
Lauraceae	*Lindera glauca*	Non	2215.47	65145	62.78%	76.75%	12
Paulowniaceae	*Paulownia fortunei*	Non	528.24	31985	29.50%	50.34%	20
Paulowniaceae	*Paulownia tomentosa*	Inv		15873			
Poaceae	*Aegilops tauschii*	Inv	4300	39630	84.40%	84.40%	7
Poaceae	*Avena sativa*	Non		80608	64.00%		21
Poaceae	*Brachypodium distachyon*	Inv	271.2	34310	28.10%	42.20%	5
Poaceae	*Coix lacryma-jobi*	Inv	1560	39574	64.80%	77.00%	5
Poaceae	*Dactylis glomerata*	Inv	1940	40088	69.00%		7
Poaceae	*Echinochloa crus-galli*	Inv	1400.22	108771	32.40%	40.70%	27
Poaceae	*Eragrostis curvula*	Inv	603.1	55182		28.70%	10
Poaceae	*Eragrostis tef*	Non	772	3800	6.00%	7.47%	10
Poaceae	*Miscanthus sinensis*	Inv	2500	67789	72.40%	72.40%	19
Poaceae	*Stipa capillata*	Non	2355	81224	23.34%	57.68%	11
Poaceae	*Triticum turgidum*	Non	12000	65012	82.20%	80.00%	14
Poaceae	*Zizania palustris*	Non	1800	46421	59.95%	76.00%	15
Poaceae	*Zoysia japonica*	Inv	334	59271	15.20%	40.90%	20
Polygonaceae	*Polygonium cuspidatum*	Inv	2600	55075	71.54%	71.54%	
Rosaceae	*Malus domestica*	Non	651	42140	59.50%		17
Rosaceae	*Prunus persica*	Non	265	27852	29.60%	37.14%	8
Rosaceae	*Rosa multiflora*	Inv	750	67380		56.40%	
Solanaceae	*Capsicum annuum*	Non	3500	32323	70.00%	77.50%	12
Solanaceae	*Datura stramonium*	Inv	2217.53	52149	60.00%	61.00%	12
Solanaceae	*Solanum lycopersicum*	Non	875.21	34384	50.25%	65.66%	12
Typhaceae	*Typha latifolia*	Inv	287	27432	16.65%	43.84%	

The T-tests revealed that this data was insignificant. The p-values were high for all five characteristics, ranging from 0.2855 for genome size to 0.8514 for the percent of TE within genomes (Figure 1). Genome size, while having the lowest p-value, also had an outlier in its data. This may be due to faulty data recording or the paper reporting its statistics differently. Even without an outlier, the data indicates the results would be insignificant for each characteristic. Ultimately, I failed to reject the null hypothesis.
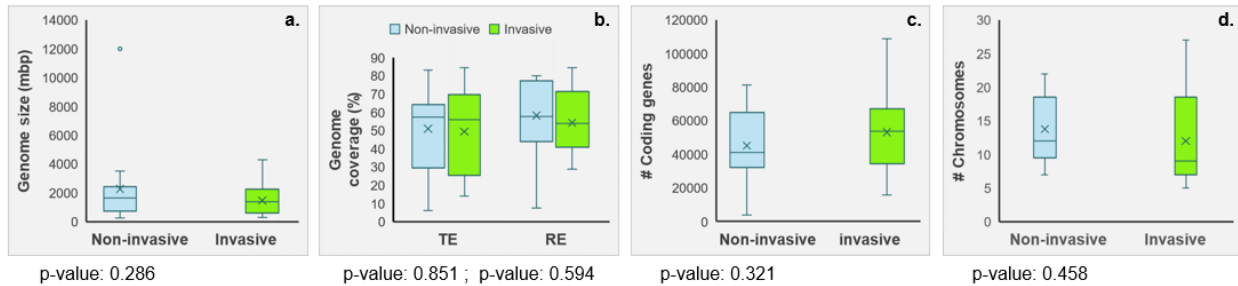

Figure 1: Box-and-whisker plots for five variables comparing invasive and non-invasive plant genomes. **a)** genome size in mega-base pairs (mbp); **b)** percent of genomes consisting of transposable elements (TE) and repetitive elements (RE); **c)** number of coding genes; and **d)** number of chromosomes. Results are insignificant.

Discussion/Reflection

None of the five characteristics that were studied—genome size, the percent of the genomes consisting of transposable elements (TE), the percent of the genomes consisting of repetitive elements (RE), the number of protein-coding genes, and the number of chromosomes—was statistically different for invasive vs. non-invasive plant species. These results indicate that these particular aspects of plant genomes may not impact the potential for a plant to be invasive in their non-native habitats.

Thus, what aspects of plant genomes impact the invasive potential of plants remains unknown. There were some challenges with this study, particularly with obtaining a large enough sample of invasive and non-invasive plant species and with obtaining enough data for all species. There are inconsistencies in how people define invasive species, with some using the standard definition and some incorrectly calling any non-native or weedy plant invasive. Some plants are listed on state websites as invasive but lack supporting research, and we can't truly know that a plant is truly non-invasive or if it simply hasn't been introduced into other ecosystems it can invade. These factors limited the number of species that were able to be included in the study, and it is possible that this may have impacted the results. Regarding the data itself, some papers reported the data from their genome assemblies in an inconsistent manner (e.g., some papers excluded the percentage of transposable elements, and some used

terms like transposable elements and repetitive elements interchangeably). To avoid testing inaccurate data, unclear numbers and definitions were excluded from the dataset, which led to smaller sample sizes when testing certain variables. Repeating this study with a larger sample may yield different results, although it is likely that these variables are indeed the same for both invasive and non-invasive plant species, and the answer to this research question may be more specific.

There are many paths to explore regarding the genetic variables that impact plants' invasive potential, but some will be more limited in what can be concluded about invasive plant species more broadly. Comparing the genes found in invasive vs. non-invasive plant species may reveal specific candidate genes or gene-gene interactions that might contribute to invasiveness. Studying this is the natural progression to this current study because the focus remains on invasive plant species, and a study of this nature could provide further knowledge that can be applied in conservation efforts. For example, if we can identify these genes or gene-gene interactions—or any other aspect of plant genomes—that contribute to a plant's invasiveness, we may eventually discover a way to (safely and ethically) alter the genomes of invasive plants to prevent them from harming ecosystems they invade.

References

Alabi, N., Wu, Y., Bossdorf, O., Rieseberg, L. H., & Colautti, R. I. (2021). Genome report: a draft genome of *Alliaria petiolata* (garlic mustard) as a model system for invasion genetics. *G3 (Bethesda)*, *11*(12), Article jkab339. https://doi.org/10.1093/g3journal/jkab339

Albert, V. A., Barbazuk, W. B., dePamphilis, C. W., Der, J. P., Leebens-Mack, J., Ma, H., Palmer, J. D., Rounsley, S., Sankoff, D., Schuster, S. C., Soltis, D. E., Soltis, P. S., Wessler, S. R., Wing, R. A., Ammiraju, J. S. S., Chamala, S., Chanderbali, A. S., Detemann, R., Ralph, P., … Tomsho, L. (2013). The *Amborella* genome and the evolution of flowering plants. *Science*, *342*, Article 6165. https://doi.org./10.1126/science.1241089

Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O., Gundlach, H., Hale, I., Mascher, M., Spannagl, M., Wiebe, K., Jordan, K. W., Golan, G., Deek, J., Ben-Zvi, B., Ben-Zvi, G., Himmelbach, A., MacLachlan, R. P., Sharpe, A. G., Fritz, A., … Distelfeld, A. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science*, *357*, 93-97. https://doi.org/10.1126/science.aan0032

Baiakhmetov, E., Guyomar, C., Shelest, E., Nobis, M., & Gudkova, P. D. (2021). The first draft genome of feather grasses using SMRT sequencing and its implications in molecular studies of *Stipa*. *Sci Rep*, *11*, Article 15345. https://doi.org/10.1038/s41598-021-94068-w

Bakker, E. G., Montgomery, B., Nguyen, T., Eide, K., Chang, J., Mockler, T. C., Liston, A., Seabloom, E. W., Borer, E. T. (2009). Strong population structure characterizes weediness gene evolution in the invasive grass species Brachypodium distachyon. *Mol Ecol*, *18*(12), 2588-2601. https://doi.org/10.1111/j.1365-294X.2009.04225.x

Bieker, V. C., Battlay, P., Petersen, B., Sun, X., Wilson, J., Brealey, J. C., Bretagnoille, F., Nurkowski, K., Lee, C., Barreiro, F. S., Owens, G. L., Lee, J. Y., Kellner, F. L., Boheemen, L. V., Gopalakrishnan, S., Gaudeul, M., Mueller-Schaerer, H., Lommen, S., Karrer, G., … Martin, M. D. (2022). Uncovering the genomic basis of an extraordinary plant invasion. *Sci Adv*, *8*(34), Article eabo5115. https://doi.org/10.1126/sciadv.abo5115

Cannarozzi, G., Plaza-Wüthrich, S., Esfeld, K. Larti, S., Wilson, Y. S., Girma, D., de Castro, E. Chanyalew, S., Blösch, R., Farinelli, L., Lyons, E., Schneider, M., Falquet, L., Kuhlemeier, C., Assefa, K., & Tadele, Z. (2014). Genome and transcriptome sequencing

identifies breeding targets in the orphan crop tef (Eragrostis tef). *BMC Gen*, *15*, Article 581. https://doi.org/10.1186/1471-2164-15-581

Cao, Y., Sun, G., Zhai, X., Xu, P., Ma, L., Deng, M., Zhao, Z., Yang, H., Dong, Y., Shang, Z., Lv, Y., Yan, L., Liu, H., Cao, X., Li, B., Wang, Z., Zhao, X., Yu, H., Wang, F., … Fan, G. (2021). Genomic insights into the fast growth of paulownias and the formation of *Paulownia* witches' broom. *Molec Plant*, *14*(10), 1668-1682. https://doi.org/10.1016/j.molp.2021.06.021

Carballo, J., Santos, B. A. C. M., Zappacosta, D. Garbus, I., Selva, J. P., Gallo, C. A., Diaz, A., Albertini, E., Caccamo, M., & Echenique, V. (2019). A high-quality genome of *Eragrostis curvula* grass provides insights into Poaceae evolution and supports new strategies to enhance forage quality. *Sci Rep*, *9*, Article 10250. https://doi.org/10.1038/s41598-019-46610-0

Chadha, A. Florentine, S., Javaid, M., Welgama, A., & Turville, C. (2020). Influence of elements of climate change on the growth and fecundity of Datura stramonium. *Environ Sci Pollut Inst*, *27*(28), 35859-35869. https://doi.org/10.1007/s11356-020-10251-y

Chamala, S., Chanderbali, A. S., Der, J. P., Lan, T., Walts, B., Albert, V. A., Depamphilis, C. W., Leebens-Mack, J., Rounsley, S., Schuster, S. C., Wing, R. A., Xiao, N., Moore, R., Soltis, P. S., Soltis, D. E., & Barbazuk, W. B. (2013). Assembly and validation of the genome of the nonmodel basal angiosperm *Amborella*. *Science*, *342*(6165), 1516-1517. https://doi.org/10.1126/science.1241130

Chung, Y. A., Burkle, L. A., & Knight, T. M. (2014). Minimal effects of an invasive flowering shrub on the pollinator community of native forbs. *PLoS One*, *9*(10), Article e109088. https://doi.org/10.1371/journal.pone.0109088

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nature*, *6*, 836-846. https://doi.org/10.1038/nrg1711

Daccord, N. Celton, J. M., Linsmith, G., Becker, C., Choisne, N., Schijlen, E., van de Geest, H., Bianco, L., Micheletti, D., Velasco, R., di Pierro, E. A., Gouzy, J., Rees, D. J. G., Guérif, P., Muranty, H., Durel, C. E., Laurens, F., Lespinasse, Y., Gaillard, S., … Bucher, E. (2017). High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. *Nat genet*, *49*(7), 1099-1106. https://doi.org/10.1038/ng.3886

Firn, J., Prober, S. M., & Buckley, Y. M. (2012). Plastic traits of an exotic grass contribute to its abundance but are not always favourable. *PLoS One*, *7*(4), Article e35870. https://doi.org/10.1371/journal.pone.0035870

Gu, Q., Chu, S., Huang, Q., Chen, A., Li, L., & Li, R. (2023). *Colletotrichum echinochloae:* a potential bioherbicide agent for control of barnyardgrass (Echinochloa crus-galli (L.) Beauv.). *Plants (Basel)*, *12*(3), Article 421. https://doi.org/10.3390/plants12030421

Guo, L., Qiu, J., Ye, C., Jin, G., Mao, L., Zhang, H., Yang, X., Peng, Q., Wang, Y., Jia, L., Lin, Z., Li, G., Fu, F., Liu, C., Chen, L., Shen, E., Wang, W., Chu, Q., Wu, D., … Fan, L. (2017). *Echinochloa-crus-galli* genome analysis provides insight into its adaptation and invasiveness as a weed. *Nat Comm*, *8*, Article 1031. https://doi.org/10.1038/s41467-017-01067-5

Haas, M., Kono, T., Macchietto, M., Millas, R., McGilp, L., Shao, M., Duquette, J., Qiu, Y., Hirsch, C. N., & Kimball, J. (2021). Whole-genome assembly and annotation of northern wild rice, *Zizania palustris* L., supports a whole-genome duplication in the *Zizania* genus. *Plant J*, *107*, 1802-1818. https://doi.org/10.1111/tpj.15419

Huang, L., Feng, G., Yan, H., Zhang, Z., Bushman, B. S., Wang, J., Bombarely, A., Li, M., Yang, Z., Nie, G., Xie, W., Xu, L., Chen, P., Zhao, X., Jiang, W., & Zhang, X. (2019). Genome assembly provides insights into the genome evolution and flowering regulations of orchardgrass. *Plant Biotech*, *18*(2), 373-388. https://doi.org/10.1111/pbi.13205

Hulse-Kemp, A. M., Maheshwari, S., Stoffel, K., Hill, T. A., Jaffe, D., Williams, S. R., Weisenfeld, N., Ramakrishnan, S., Kumar, V., Shah, P., Schatz, M. C., Church, D. M., & van Deynze, A. (2018). Reference quality assembly of the 3.5-Gb genome of *Capsicum annuum* from a single linked-read library. *Hortic Res*, *5*, Article 4. https://doi.org/10.1038/s41438-017-0011-0

Jeong, K. Y., Lee, J. Mistrello, G., Park, K. H., & Park, J. W. (2018). IgE cross-reactivity between *Humulus japonicus* and *Humulus lupulus*. *Yonsei Med*, *59*(7), 852-856. https://doi.org/10.3349/ymj.2018.59.7.852

Kamal, N., Renhuldt, N. T., Bentzer, J., Gundlach, H., Haberer, G., Juhász, A., Lux, T., Bose, U., Tye-Din, J. A., Lang, D., van Gessel, N., Reski, R., Fu, Y. B., Spégel, P., Ceplitis, A., Himmelbach, A., Waters, A. J., Bekele, W. A., Colgrave, M. L., … Sirijovski, N. (2022). The mosaic oat genome gives insights into a uniquely healthy cereal crop. *Nature*, *606*, 113-119. https://doi.org/10.1038/s41586-022-04732-y

Kang, L., Qian, L., Zheng, M., Chen, L., Chen, H., Yang, L., You, L., Yang, B., Yan, M., Gu, Y., Wang, T., Schiessl, S. V., An, H., Blischak, P., Liu, X., Lu, H., Zhang, D., Rao, Y., Jia, D., … Zhongsong, L. (2021). Genomic insights into the origin, domestication and diversification of *Brassica juncea*. *Nat Gen*, *53*, 1392-1402. https://doi.org/10.1038/s41588-021-00922-y

Kang, S. H., Kim, B., Choi, B. S., Lee, H. O., Kim, N. H., Lee, S. J., Kim, H. S., Shin, M. J., Kim, H. W., Nam, K., Kang, K. D., Kwon, S. J., Oh, T. J., Lee, S. C., & Kim, C. K. (2020). Genome assembly and annotation of soft-shelled adlay (*Coix lacryma-jobi* variety *ma-yuen*), a cereal and medicinal crop in the Poaceae family. *Plant Sci*, *11*, Article 630. https://doi.org/10.3389/fpls.2020.00630

Kesel, E., Hudson, A. O., & Osier, M. V. (2022). Whole-genome sequence, assembly, and annotation of an invasive plant, *Lonicera maackii* (amur honeysuckle). *Plants (Basel)*, *11*(23), Article 3253. https://doi.org/10.3390/plants11233253

Lankau, R. A. (2012). Coevolution between invasive and native plants driven by chemical competition and soil biota. *PNAS USA*, *109*(29), 11240-11245. https://doi.org/10.1073/pnas.1201343109

Luo, M. C., Gu, Y. Q., Puiu, D., Wang, H., Twardziok, S. O., Deal, K. R., Huo, N., Zhu, T., Wang, L., Wang, Y., McGuire, P. E., Liu, S., Long, H., Ramasamy, R. K., Rodriguez, J. C., Van, S. L., Yuan, L., Wang, Z., Xia, Z., … Dvořák, J. (2017). Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. *Nature*, *551*, 498-502. https://doi.org/10.1038/nature24486

Mitros, T., Session, A. M., James, B. T. Wu, G. A., Belaffif, M. B., Clark, L. V., Shu, S., Dong, H., Barling, A., Holmes, J. R., Mattick, J. E., Bredeson, J. V., Liu, S., Farrar, K., Glowacka, K., Jeżowski, S., Barry, K., Chae, W. B., Juvik, J. A., … Rokhsar, D. S. (2020). Genome biology of the paleotetraploid perennial biomass crop *Miscanthus*. *Nat Commun*, *11*, Article 5442. https://doi.org/10.1038/s41467-020-18923-6

Muñoz-López, M., & García-Pérez, J. L. (2010). DNA transposons: nature and applications in genomics. *Curr Gen*, *11*(2), 115-128. https://doi.org/10.2174/138920210790886871

Nakamura, N., Hirakawa, H., Sato, S., Otagaki, S., Matsumoto, S., Tabata, S., & Tanaka, Y. (2018). Genome structure of *Rosa multiflora*, a wild ancestor of cultivated roses. *DNA Res*, *25*(2), 113-121. https://doi.org/10.1093/dnares/dsx042

Natsume, S., Takagi, H., Shiraishi, A., Murata, J., Toyonaga, H., Patzak, J., Takagi, M., Yaegashi, H., Uemura, A., Mitsuoka, C., Yoshida, K., Krofta, K., Satake, H., Terauchi, R., & Ono, E. (2015). The draft genome of hop (*Humulus lupulus*), an essence for brewing. *Plant and Cell Phys*, *56*(3), 428-441. https://doi.org/10.1093/pcp/pcu169

Oduor, A. M. O., Lankau, R. A., Strauss, S. Y., & Gómez, J. M. (2011). Introduced *Brassica nigra* populations exhibit greater growth and herbivore resistance but less tolerance than native populations in the native range. *New Phyt*, *191*(2), 536-544. https://doi.org/10.1111/j.1469-8137.2011.03685.x

Pakpour, S., & Klironomos, J. (2015). The invasive plant, *Brassica nigra*, degrades local mycorrhizas across a wide geographical landscape. *R Soc Open Sci*, *2*(9), Article 150300. https://doi.org/10.1098/rsos.150300

Perumal, S., Koh, C. S., Jin, L., Buchwaldt, M., Higgins, E. E., Zheng, C., Sankoff, D., Robinson, S. J., Kagale, S., Navabi, Z. K., Tang, L., Horner, K. N., He, Z., Bancroft, I., Chalhoub, B., Sharpe, A. G., & Parkin, I. A. P. (2020). A high-contiguity *Brassica nigra* genome localizes active centromeres and defines the ancestral *Brassica* genome. *Nat Plants*, *6*, 929-941. https://doi.org/10.1038/s41477-020-0735-y

Pu, X., Li, Z., Tian, Y., Gao, R., Hao, L., Hu, Y., He, C., Sun, W., Xu, M., Peters, R. J., van de Peer, Y., Xu, Z., & Song, J. (2020). The honeysuckle genome provides insight into the molecular mechanism of carotenoid metabolism underlying dynamic flower coloration. *New Phytol*, *227*(3), 930-943. https://doi.org/10.1111/nph.16552

Qu, Z., Wang, W., & Adelson, D. L. (2023). Chromosomal level genome assembly of medicinal plant *Sophora flavescens*. *Sci Data*, *10*(1), Article 572. https://doi.org/10.1038/s41597-023-02490-8

Quinty, V., Colas, C., Nasreddine, R., Nehmé, R., Pilot, C., Draye, M., Destandau, E., da Silva, D., & Chatel, G. (2022). Screening and evaluation of dermo-cosmetic activities of the invasive plant species *Polygonum cuspidatum*. *Plants (Basel)*, *12*(1), Article 83. https://doi.org/10.3390/plants12010083

Rajewski, A., Carter-House, D., Stajich, J., & Litt, A. (2021). Datura genome reveals duplications of psychoactive alkaloid biosynthetic genes and high mutation rate following tissue culture. *BMC Gen*, *22*, Article 201. https://doi.org/10.1186/s12864-021-07489-2

Semenya, S. S., & Maroyi, A. (2020). Assessment of useful alien plant species cultivated and managed in rural home gardens of Limpopo Province, South Africa. *Scientifica (Cairo)*, Article 3561306. https://doi.org/10.1155/2020/3561306

Su, X., Wang, B., Geng, X., Du, Y., Yang, Q., Liang, B., Meng, G., Gao, Q., Yang, W., Zhu, Y., & Lin, T. (2021). A high-continuity and annotated tomato reference genome. *BMC Gen*, *22*, Article 898. https://doi.org/10.1186/s12864-021-08212-x

Suda, J., Meyerson, L. A., Leitch, I. J., & Pyšek, P. (2014). The hidden side of plant invasions: the role of genome size. *New Phyt*, *205*(3), 994-1007. https://doi.org/10.1111/nph.13107

Tanaka, H., Hirakawa, H., Kosugi, S., Nakayama, S., Ono, A., Watanabe, A., Hashiguchi, M., Gondo, T., Ishigaki, G., Muguerza, M., Shimizu, K., Sawamura, N., Inoue, T., Shigeki, Y., Ohno, N., Tabata, S., Akashi, R., & Sato, S. (2016). Sequencing and comparative analyses of the genomes of zoysiagrasses. *DNA Res*, *23*(2), 171-180. https://doi.org/10.1093/dnares/dsw006

The International Brachypodium Initiative. (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature*, *463*, 763-768. https://doi.org/10.1038/nature08747

The International Peach Genome Initiative, Verde, I., Abbott, A. G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., Zhebentyayeva, T., Dettori, M. T., Grimwood, J., Cattonaro, F., Zuccolo, A., Rossini, L., Jenkins, J., Vendramin, E., Meisel, L. A., Decroocq, V., Sosinski, B., Prochnik, S., … Rokhsar, D. S. (2013). The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Gen*, *45*, 487-494. https://doi.org/10.1038/ng.2586

Wang, H. H., Wonkka, C. L., Grant, W. E., & Rogers, W. E. (2012). Potential range expansion of Japanese honeysuckle (*Lonicera japonica* Thunb.) in southern U.S. forestlands. *Forests*, *3*(3), 573-590. https://doi.org/10.3390/f3030573

Widanagama, S. D., Freeland, J. R., Xu, X., & Shafer, A. B. A. (2022). Genome assembly, annotation, and comparative analysis of the cattail *Typha latifolia*. *Gen Rep*, *12*(2), Article jkab401. https://doi.org/10.1093/g3journal/jkab401

Xiong, B., Zhang, L., Xie, L., Li, L. He, X., Niu, Y., Zhang, T., Liao, S., Dong, S., & Zhang, Z. (2022). Genome of *Lindera glauca* provides insights into the evolution of biosynthesis

genes for aromatic compounds. *iScience*, *25*(8), Article 104761. https://doi.org/10.1016/j.isci.2022.104761

Yang, J., Zhang, G., Zhang, J., Liu, H., Chen, W., Wang, X., Li, Y., Dong, Y., & Yang, S. (2017). Hybrid de novo genome assembly of the Chinese herbal fleabane *Erigeron breviscapus*. *GigaSci*, *6*(6), 1-7. https://doi.org/10.1093/gigascience/gix028

Yang, M, Zhao, H., Xian, X., Wang, R., Yang, N., Chen, L., & Liu, W. X. (2023). Assessing risk from invasive alien plants in China: reconstructing invasion history and estimating distribution patterns of *Lolium temulentum* and *Aegilops tauschii*. *Front Plant Sci*, *14*, Article 1113567. https://doi.org/10.3389/fpls.2023.1113567

Zhang, A., Hu, X., Yao, S., Yu, M., & Ying, Z. (2021). Alien, naturalized and invasive plants in China. *Plants (Basel)*, *10*(11), Article 2241. https://doi.org/10.3390/plants10112241

Zhang, Y., Zheng, L., Zheng, Y., Zhou, C., Huang, P., Xiao, X., Zhao, Y., Hao, X., Hu, Z., Chen, Q., Li, H., Wang, X., Fukushima, K., Wang, G., & Li, C. (2019). Assembly and annotation of a draft genome of the medicinal plant *Polygonum cuspidatum*. *Plant Sci*, *10*. https://doi.org/10.3389/fpls.2019.01274

Zhao, Z., Li, Y., Liu, H., Zhai, X., Deng, M., Dong, Y., & Fan, G. (2017). Genome-wide expression analysis of salt-stressed diploid and autotetraploid *Paulownia tomentosa*. *PLoS One*, *12*(10), Article e0185455. https://doi.org/10.1371/journal.pone.0185455