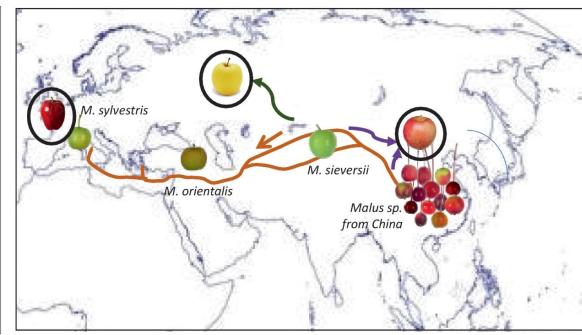
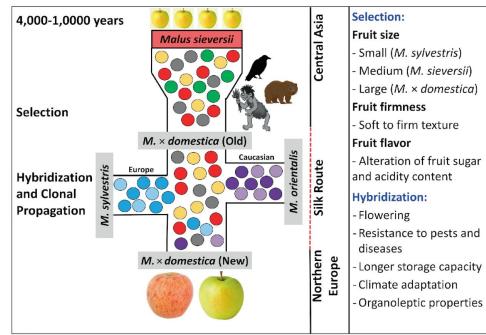


How ‘Bout ‘Dem Apples?

Has Modern Agriculture Made Apple Cultivars More Related
And Less Genetically Diverse?



Malus Domestica Originated In Central Asia And Spread To The Rest Of The World Via The Silk Road Through Vegetative Propagation

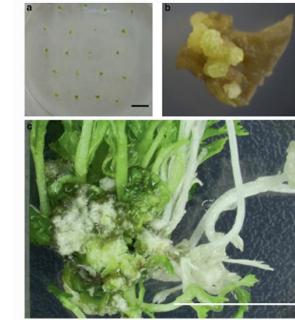
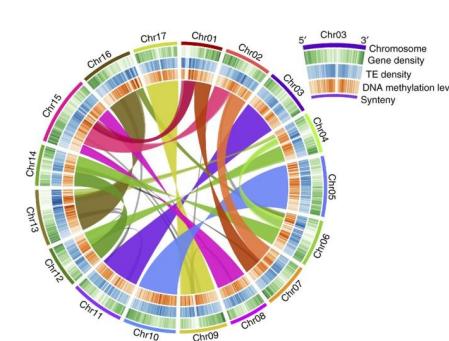
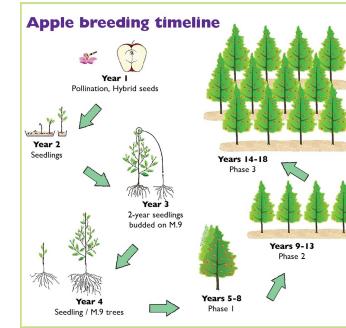
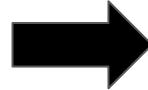


M. Domestica is typically:

- Diploid, 17 chromosomes
- Self-infertile
- Highly heterozygous
- Propagated through grafting to maintain desirable traits



Starting In The 19th Century, The Apple Industry Moved From Small Farms Selecting Chance Cultivars To Large Vertically Integrated Firms Using Sophisticated Breeding Techniques



Research Has Shown That Today Many Cultivars Are Clonally Related To A Small Number Of Commercially Dominant Cultivars

nature > horticulture research > articles > article

Article | Open access | Published: 01 January 2021

Genomic consequences of apple improvement

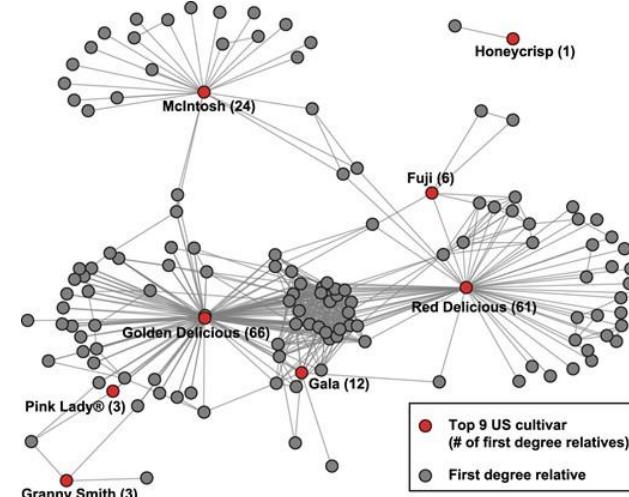
Zoë Migicovsky, Kyle M. Gardner, Christopher Richards, C. Thomas Chao, Heidi R. Schwaninger, Gennaro Fazio, Gan-Yuan Zhong & Sean Myles

Horticulture Research 8, Article number: 9 (2021) | [Cite this article](#)

8884 Accesses | 42 Citations | 117 Altmetric | [Metrics](#)

Abstract

The apple (*Malus domestica*) is one of the world's most commercially important perennial crops and its improvement has been the focus of human effort for thousands of years. Here, we genetically characterise over 1000 apple accessions from the United States Department of Agriculture (USDA) germplasm collection using over 30,000 single-nucleotide polymorphisms (SNPs). We confirm the close genetic relationship between modern apple cultivars and their primary progenitor species, *Malus sieversii* from Central Asia, and find that cider apples derive more of their ancestry from the European crabapple, *Malus sylvestris*, than do dessert apples. We determine that most of the USDA collection is a large complex pedigree: over half of the collection is interconnected by a series of first-degree relationships. In addition, 15% of the accessions have a first-degree relationship with one of the top 8 cultivars produced in the USA. With the exception of 'Honeycrisp', the top 8



Hypothesis: Newly developed apple cultivars have become more related and less genetically diverse over time.

Used Phenomic And SNP Data From Canada's Apple Biodiversity Collection To Evaluate This Hypothesis

Quantifying apple diversity: A phenomic characterization of Canada's Apple Biodiversity Collection

Sophie Watts, Zoë Migicovsky, Kendra A. McClure, Cindy H. J. Yu, Beatrice Amyotte, Thomas Baker, David Bowlby, Karen Burgher-MacLellan, Laura Butler, Richard Donald, Lihua Fan ... See all authors ▾

First published: 23 June 2021 | <https://doi.org/10.1002/ppp3.10211> | Citations: 17

```
# A tibble: 446 × 7
#> #>   apple_id Cultivar      Year Country     Use    type Period
#> #>   <dbl> <chr>       <dbl> <chr>      <chr>  <chr>  <dbl>
#> 1     559 Adersleber Kalvill 1870 Germany   Eating  NA     1
#> 2     411 Aivania        1950 Bulgaria   Eating  cider   2
#> 3     245 Akane          1937 Japan     Eating  dessert  2
#> 4     505 Alkmene         1962 Germany   Eating  NA     3
#> 5     414 Alnarp 2       1944 Sweden    Rootstock NA     2
#> 6     395 Alton          1938 United States Eating  NA     2
#> 7     1216 Ambrosia       1993 Canada   Eating  dessert  3
#> 8     394 American Summer Pearmain 1817 United States Eating  dessert  1
#> 9     468 Ananas Berzenicki 1900 Lithuania Dessert dessert  2
#> 10    418 Anisim          1870 Former Soviet Union Dessert NA     1
#> # i 436 more rows
#> # i Use `print(n = ...)` to see more rows
```

Genotyping-by-sequencing of Canada's apple biodiversity collection

Zoë Migicovsky¹, Gavin M. Douglas², Sean Myles^{1*}

¹ Plant, Food, and Environmental Sciences, Faculty of Agriculture, Dalhousie University, Truro, NS, Canada

² Genomics

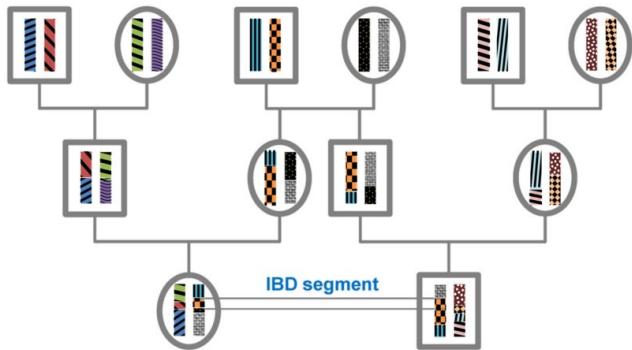


skirnir314@skirnir314-XPS-9320: ~/EECMSc/AppleProject/Dat... ↵

1	0	0	0	-9	A	T	T	C	C	G	C	C	A	A	T	G	T	G	G	A	A	G	G	A	A	G	G				
2	2	0	0	-9	C	C	G	G	T	T	T	C	C	C	G	T	T	G	G	G	A	A	G	G	A	A	C	A	G		
3	3	0	0	-9	A	A	T	T	C	T	T	G	G	C	A	A	A	G	G	G	A	A	G	G	A	A	G	G			
4	4	0	0	-9	A	A	T	T	C	T	C	G	G	C	A	A	A	G	T	G	G	A	A	G	G	A	A	G	G		
5	5	0	0	-9	C	A	G	T	T	T	T	G	C	C	G	A	T	G	G	G	A	A	G	G	A	A	G	G			
6	6	0	0	-9	C	C	G	G	T	T	T	C	C	C	G	A	T	G	G	G	A	A	G	G	A	A	C	A	G		
7	7	0	0	-9	C	A	G	T	C	T	T	G	C	C	G	A	T	G	G	G	A	A	G	G	A	A	C	A	G		
8	8	0	0	-9	C	A	G	T	C	T	T	G	C	C	G	A	T	G	G	G	A	A	G	G	A	A	G	C	A	G	
9	9	0	0	-9	C	A	G	T	C	T	T	G	C	C	G	A	T	G	G	G	A	A	G	G	A	A	G	G			
10	10	0	0	-9	C	A	G	T	C	T	T	G	G	C	G	A	T	G	G	G	A	A	G	G	A	A	G	C	A	G	
11	11	0	0	-9	C	C	G	G	T	T	T	G	C	C	G	A	T	G	G	G	A	A	G	G	A	A	G	C	A	G	
12	12	0	0	-9	C	A	G	T	C	T	T	G	C	C	G	G	T	G	G	G	A	A	G	G	A	A	G	C	A	G	
13	13	0	0	-9	A	A	T	T	C	T	T	G	G	C	C	A	A	G	G	G	A	A	G	G	T	T	G	G	A	G	
14	14	0	0	-9	C	A	G	T	C	T	T	G	C	C	G	A	T	G	G	G	A	A	G	G	A	A	G	A	A	G	
15	15	0	0	-9	C	C	G	G	T	T	T	G	G	C	G	T	T	G	G	G	A	A	G	G	A	A	A	C	A	G	
16	16	0	0	-9	C	A	G	T	C	T	T	G	G	C	C	A	A	A	G	G	G	A	A	G	G	A	A	A	A	A	
17	17	0	0	-9	C	A	G	T	C	T	T	G	G	C	C	G	A	T	G	G	G	A	A	G	G	A	A	G	C	A	G
18	18	0	0	-9	C	C	G	G	T	T	T	G	G	C	C	G	A	T	G	G	G	A	A	G	G	A	A	G	C	A	G
19	19	0	0	-9	C	A	G	T	C	T	T	G	C	C	G	A	T	G	G	G	A	A	G	G	A	A	G	C	A	G	
20	20	0	0	-9	C	A	G	T	C	T	T	G	G	C	C	C	A	A	G	G	G	A	A	G	G	A	A	G	C	A	G
21	21	0	0	-9	C	A	G	T	C	T	T	G	G	C	C	G	A	T	G	G	G	A	A	G	G	A	A	G	C	A	G
22	22	0	0	-9	A	A	T	T	C	T	T	G	G	C	C	A	A	A	G	G	G	A	A	G	G	A	A	G	G	A	A
23	23	0	0	-9	C	A	G	T	C	T	T	G	G	C	C	G	A	T	G	G	G	A	A	G	G	A	A	G	G	A	A

278,231 SNPs across 1,175 accessions
(i.e., cultivars)

Examined Relatedness Using Identity By Descent



- IBD measures relatedness between individuals.
 $\hat{\pi}$ (proportion of inheritance):
 - >0.125 = 3rd degree relatives (cousin)
 - >0.25 = 2nd degree (half siblings)
 - >0.5 = 1st degree (full siblings)
 - ~ 1 = twins/clones.
- SNPs that are the same are **identical by state**. If they are passed down from a common ancestor without recombination, they are **identical by descent**.

Examined Genetic Diversity Using PCoA of Genetic Distance

```
[ ,1]      [ ,2]      [ ,3]      [ ,4]      [ ,5]      [ ,6]
[1,] 0.000000 0.176054 0.152459 0.150258 0.168394 0.188489
[2,] 0.176054 0.000000 0.178288 0.186578 0.181374 0.183655
[3,] 0.152459 0.178288 0.000000 0.131805 0.160806 0.190888
[4,] 0.150258 0.186578 0.131805 0.000000 0.167971 0.190529
[5,] 0.168394 0.181374 0.160806 0.167971 0.000000 0.191575
[6,] 0.188489 0.183655 0.190888 0.190529 0.191575 0.000000
> |
```

- Genetic distance measured as the Hamming distance, i.e. the percentage of SNPs that differ between accessions.
- Summarized in square distance matrix.

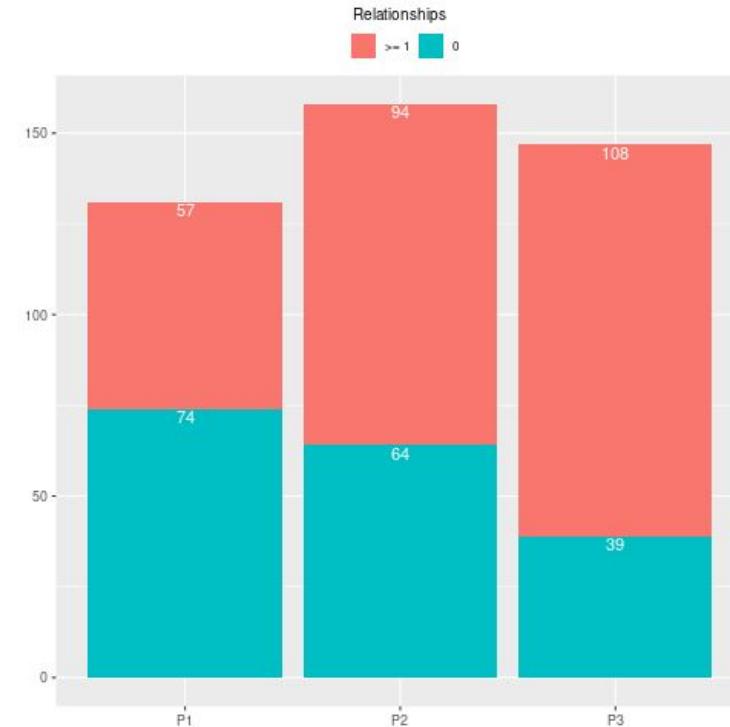
Split Data Into Three Periods; Used Plink and R For Quality Control And Analysis

- Accessions split into three periods corresponding with stages of industry development:
 - **P1 (1800-1899)**: Early breeding techniques and beginning consolidation
 - **P2 (1900-1959)**: Increased sophisticated
 - **P3 (1960-Present)**: Vertical integration and modern genomics
- Filtering/cleaning:
 - SNPs filtered based on standard protocols (e.g., removed SNPs missing from many cultivars)
 - Included only *M. domestica* with a known release year.
 - Remove non-diploid cultivars
- This left **~150k SNPs** across **131** accessions in P1, **158** in P2, and **148** in P3
- IBD estimates and genetic distance were calculated using Plink, PCoA in R.

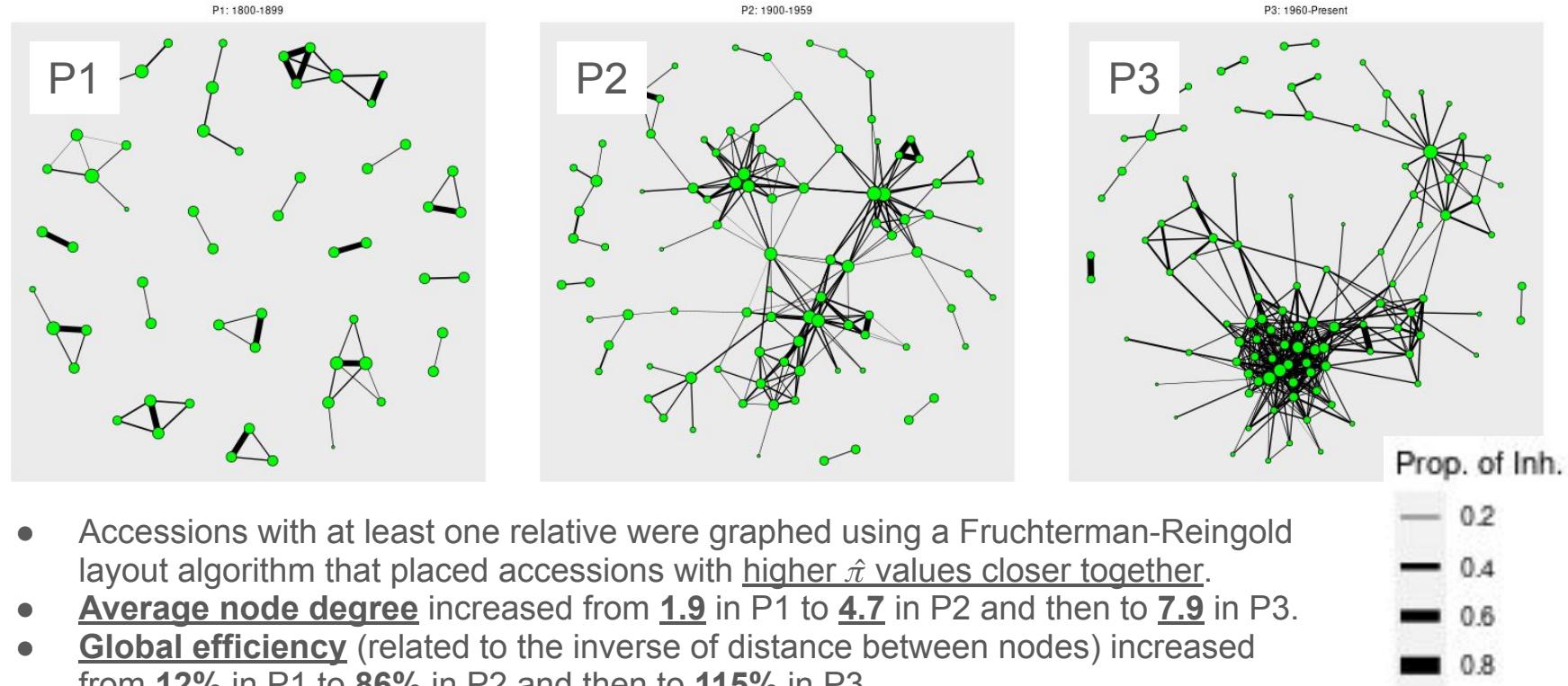
```
skirnir314@skirnir314-XPS-9320:~/EECMSc/AppleProject$ /usr/local/bin/Plink/plink  
--file ./Data/Snps/abc_combined_maf001_sort_vineland_imputed --keep ./Data/keep  
.pi.tsv --extract ./Data/Snps/plink.prune.in --allow-no-sex --geno .05 --mind .1  
.maf .01 --genome --out ./Data/Snps/p1_tbd  
PLINK v1.90b7.2 64-bit (11 Dec 2023) www.cog-genomics.org/plink/1.9/  
(C) 2005-2023 Shaun Purcell, Christopher Chang GNU General Public License v3  
Logging to ./Data/Snps/p1_tbd.log.  
options in effect:  
--allow-no-sex  
--extract ./Data/Snps/plink.prune.in  
--file ./Data/Snps/abc_combined_maf001_sort_vineland_imputed  
--geno .05  
--genome  
--keep ./Data/keep.pi.tsv  
.maf .01  
.mind .1  
.out ./Data/Snps/p1_tbd
```

IBD Estimates Showed That The Percentage Of Accessions With At Least One Relative Increased Over Time

- Used a $\hat{\pi}$ threshold of ≥ 0.125 (cousins or closer) to determine relatedness.
- Based on this threshold, the % of accessions with at least one relative increased from 44% in P1 to 59% in P2 and then to 73% in P3.



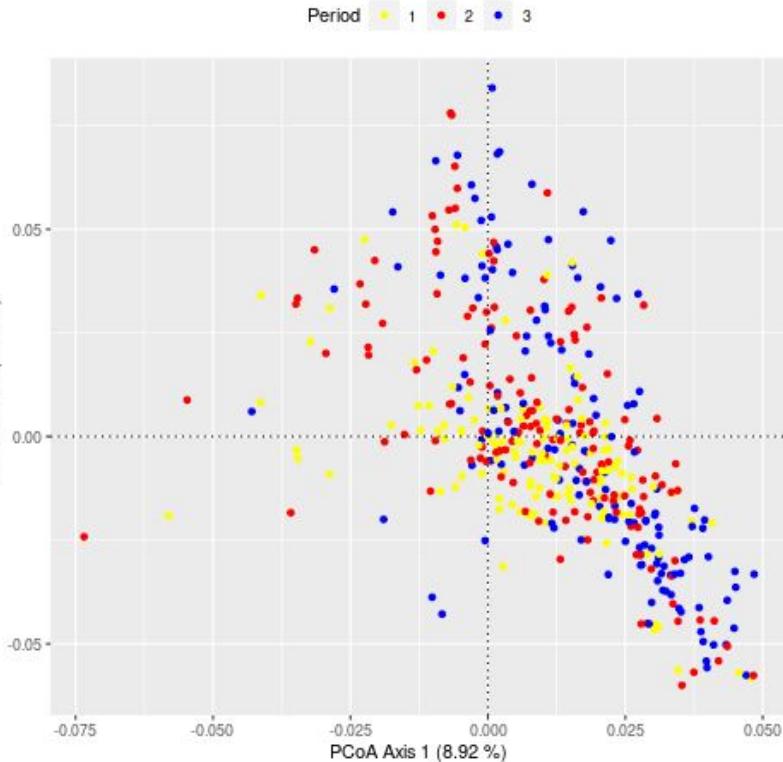
Network Graphs Of Related Accessions Also Showed Increased Relatedness Over Time



But, a PCoA of Genetic Distance Showed No Change In Genetic Diversity Over Time

- Accessions with similar genetic distance from each other will be placed closer together on the plot.
- No changes in density or different groupings evident across periods, indicating no change in genetic distance.

PCoA of apple cultivar SNPs



Discussion

Caveats:

1. Time periods were defined semi-arbitrarily
2. Accessions were not a random sample from each time period
3. Plink's IBD estimate is imperfect
4. Excluding triploid cultivars excludes some major commercially successful cultivars
5. Genetic distance is overall difference, doesn't focus on coding SNPs or SNPs that impact significant traits

That said:

1. Time periods are somewhat informed by the literature
2. Uses a reasonably high number of accessions
3. Spot check confirms that Plink's IBD estimate is reasonable
4. Most cultivars are diploid

Suggested Conclusions

- Newly developed cultivars probably are becoming increasingly related to each other.
- But, for now this does not seem to have come at the cost of decreased genetic diversity
- Regardless, we should continue to preserve heritage apple cultivars as a source of genetic diversity and a shared cultural heritage.

References

- Khan, A., Gutierrez, B., Chao, C.T., & Singh, J. (2021) Origin of the Domesticated Apples. In: Podwyszyńska, M., & Marasek-Ciołakowska, A. (eds.) (2021) *The Apple Genome*. Springer Cham, pp. 383-394.
- Migicovsky, Z., Douglas, G.M., & Myles, S. (2022) Genotyping-by-sequencing of Canada's apple biodiversity collection. *Genomics of Plants and the Phytoecosystem*. 13.
- Migicovsky, Z., Gardner, K.M., Richards, C., Chao, C.T., Schwaninger, H.R., Fazio, G., Zhong, G., & Myles, S. (2021) Genomic consequences of apple improvement. *Horticultural Research*. 8.
- O'Rourke, D. (2021) Economic Importance of the World Apple Industry. In: Podwyszyńska, M., & Marasek-Ciołakowska, A. (eds.) (2021) *The Apple Genome*. Springer Cham, pp. 1-18.
- Watts, S., et al. (2021) Quantifying apple diversity: A phenomic characterization of Canada's Apple Biodiversity Collection. *Plants People Planet*. 3 (6), 747-760.

Appendix

Yellow Bellflower



NJ, USA, mid 1700s

Winesap (Maybe)



NJ, USA, first documented 1804

Red Delicious



Iowa, USA, Jesse Hiatt, 1872

