R_Assignment

Zheyuan Zhang

3/14/2025

Part I

Data Inspection

Check the enviomental frame

```
library("tidyverse")
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✓ dplyr     1.1.4     ✓ readr     2.1.5
## ✓ forcats   1.0.0     ✓ stringr   1.5.1
## ✓ ggplot2   3.5.1     ✓ tibble    3.2.1
## ✓ lubridate 1.9.3     ✓ tidyr     1.3.1
## ✓ purrr     1.0.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becom
e errors
```

```
library(ggplot2)
```

Read files

```
raw_genotypes <- read_tsv("fang_et_al_genotypes.txt")
```

```
## Rows: 2782 Columns: 986
## ── Column specification ────────────────────────────────────────────────────────
## Delimiter: "\t"
## chr (986): Sample_ID, JG_OTU, Group, abph1.20, abph1.22, ae1.3, ae1.4, ae1.5...
##
## ℹ Use `spec()` to retrieve the full column specification for this data.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
raw_snp_position <- read_tsv("snp_position.txt")
```

```
## Rows: 983 Columns: 15
## ── Column specification ────────────────────────────────────────────────────────
## Delimiter: "\t"
## chr (9): SNP_ID, Chromosome, Position, alt_pos, mult_positions, amplicon, cd...
## dbl (6): cdv_marker_id, Genaissance_daa_id, Sequenom_daa_id, count_amplicons...
##
## ℹ Use `spec()` to retrieve the full column specification for this data.
## ℹ Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Inspect:

```
head(raw_genotypes)
```

```
## # A tibble: 6 × 986
##   Sample_ID JG_OTU   Group abph1.20 abph1.22 ae1.3 ae1.4 ae1.5 an1.4 ba1.6 ba1.9
##   <chr>     <chr>    <chr> <chr>    <chr>    <chr> <chr> <chr> <chr> <chr> <chr>
## 1 SL-15     T-aust-1 TRIPS ?/?      ?/?      T/T   G/G   T/T   C/C   ?/?   G/G
## 2 SL-16     T-aust-2 TRIPS ?/?      ?/?      T/T   ?/?   T/T   C/C   A/G   G/G
## 3 SL-11     T-brav-1 TRIPS ?/?      ?/?      T/T   G/G   T/T   ?/?   G/G   G/G
## 4 SL-12     T-brav-2 TRIPS ?/?      ?/?      T/T   G/G   T/T   C/C   G/G   G/G
## 5 SL-18     T-cund   TRIPS ?/?      ?/?      T/T   G/G   T/T   C/C   ?/?   G/G
## 6 SL-2      T-dact-1 TRIPS ?/?      ?/?      T/T   G/G   T/T   C/C   A/G   G/G
## # ℹ 975 more variables: bt2.5 <chr>, bt2.7 <chr>, bt2.8 <chr>, Fea2.1 <chr>,
## #   Fea2.5 <chr>, id1.3 <chr>, lg2.11 <chr>, lg2.2 <chr>, pbf1.1 <chr>,
## #   pbf1.2 <chr>, pbf1.3 <chr>, pbf1.5 <chr>, pbf1.6 <chr>, pbf1.7 <chr>,
## #   pbf1.8 <chr>, PZA00003.11 <chr>, PZA00004.2 <chr>, PZA00005.8 <chr>,
## #   PZA00005.9 <chr>, PZA00006.13 <chr>, PZA00006.14 <chr>, PZA00008.1 <chr>,
## #   PZA00010.5 <chr>, PZA00013.10 <chr>, PZA00013.11 <chr>, PZA00013.9 <chr>,
## #   PZA00015.4 <chr>, PZA00017.1 <chr>, PZA00018.5 <chr>, PZA00029.11 <chr>, …
```

```
head(raw_snp_position)
```

```
## # A tibble: 6 × 15
##   SNP_ID   cdv_marker_id Chromosome Position  alt_pos mult_positions amplicon
##   <chr>            <dbl> <chr>      <chr>     <chr>   <chr>          <chr>
## 1 abph1.20          5976 2          27403404  <NA>    <NA>           abph1
## 2 abph1.22          5978 2          27403892  <NA>    <NA>           abph1
## 3 ae1.3             6605 5          167889790 <NA>    <NA>           ae1
## 4 ae1.4             6606 5          167889682 <NA>    <NA>           ae1
## 5 ae1.5             6607 5          167889821 <NA>    <NA>           ae1
## 6 an1.4             5982 1          240498509 <NA>    <NA>           an1
## # ℹ 8 more variables: cdv_map_feature.name <chr>, gene <chr>,
## #   `candidate/random` <chr>, Genaissance_daa_id <dbl>, Sequenom_daa_id <dbl>,
## #   count_amplicons <dbl>, count_cmf <dbl>, count_gene <dbl>
```

```
dim(raw_genotypes)
```

```
## [1] 2782  986
```

```
dim(raw_snp_position)
```

```
## [1] 983  15
```

```
str(raw_genotypes)
```

```
## spc_tbl_ [2,782 × 986] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ Sample_ID     : chr [1:2782] "SL-15" "SL-16" "SL-11" "SL-12" ...
##  $ JG_OTU        : chr [1:2782] "T-aust-1" "T-aust-2" "T-brav-1" "T-brav-2" ...
##  $ Group         : chr [1:2782] "TRIPS" "TRIPS" "TRIPS" "TRIPS" ...
##  $ abph1.20      : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ abph1.22      : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ ae1.3         : chr [1:2782] "T/T" "T/T" "T/T" "T/T" ...
##  $ ae1.4         : chr [1:2782] "G/G" "?/?" "G/G" "G/G" ...
##  $ ae1.5         : chr [1:2782] "T/T" "T/T" "T/T" "T/T" ...
##  $ an1.4         : chr [1:2782] "C/C" "C/C" "?/?" "C/C" ...
##  $ ba1.6         : chr [1:2782] "?/?" "A/G" "G/G" "G/G" ...
##  $ ba1.9         : chr [1:2782] "G/G" "G/G" "G/G" "G/G" ...
##  $ bt2.5         : chr [1:2782] "?/?" "?/?" "C/C" "C/C" ...
##  $ bt2.7         : chr [1:2782] "A/A" "A/A" "A/A" "A/A" ...
##  $ bt2.8         : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ Fea2.1        : chr [1:2782] "C/C" "C/C" "?/?" "?/?" ...
##  $ Fea2.5        : chr [1:2782] "A/A" "A/A" "A/A" "A/A" ...
##  $ id1.3         : chr [1:2782] "T/T" "T/T" "T/T" "T/T" ...
##  $ lg2.11        : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ lg2.2         : chr [1:2782] "A/A" "A/A" "A/A" "A/A" ...
##  $ pbf1.1        : chr [1:2782] "?/?" "T/T" "T/T" "T/T" ...
##  $ pbf1.2        : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ pbf1.3        : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ pbf1.5        : chr [1:2782] "?/?" "?/?" "A/A" "A/A" ...
##  $ pbf1.6        : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ pbf1.7        : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ pbf1.8        : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00003.11   : chr [1:2782] "?/?" "?/?" "C/C" "?/?" ...
##  $ PZA00004.2    : chr [1:2782] "T/T" "T/T" "?/?" "T/T" ...
##  $ PZA00005.8    : chr [1:2782] "G/G" "G/G" "G/G" "G/G" ...
##  $ PZA00005.9    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00006.13   : chr [1:2782] "A/A" "A/A" "A/A" "A/A" ...
##  $ PZA00006.14   : chr [1:2782] "?/?" "G/G" "G/G" "G/G" ...
##  $ PZA00008.1    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00010.5    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00013.10   : chr [1:2782] "A/A" "A/A" "A/A" "A/A" ...
##  $ PZA00013.11   : chr [1:2782] "C/C" "C/C" "C/T" "C/T" ...
##  $ PZA00013.9    : chr [1:2782] "T/T" "T/T" "T/T" "T/T" ...
##  $ PZA00015.4    : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ PZA00017.1    : chr [1:2782] "G/G" "G/G" "G/G" "G/G" ...
##  $ PZA00018.5    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00029.11   : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00029.12   : chr [1:2782] "T/T" "T/T" "T/T" "T/T" ...
##  $ PZA00030.11   : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ PZA00031.5    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00041.3    : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ PZA00042.2    : chr [1:2782] "?/?" "T/T" "?/?" "?/?" ...
##  $ PZA00042.5    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00043.7    : chr [1:2782] "T/T" "T/T" "T/T" "T/T" ...
##  $ PZA00045.1    : chr [1:2782] "G/G" "G/G" "G/G" "G/G" ...
##  $ PZA00047.2    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00049.12   : chr [1:2782] "?/?" "?/?" "T/T" "T/T" ...
```

```
##  $ PZA00050.9    : chr [1:2782] "A/A" "A/A" "?/?" "?/?" ...
##  $ PZA00051.2    : chr [1:2782] "A/A" "A/A" "A/A" "A/A" ...
##  $ PZA00058.5    : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ PZA00058.6    : chr [1:2782] "T/T" "?/?" "T/T" "T/T" ...
##  $ PZA00060.2    : chr [1:2782] "?/?" "?/?" "C/C" "C/C" ...
##  $ PZA00061.1    : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ PZA00065.2    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00069.4    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00070.5    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00078.2    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00079.1    : chr [1:2782] "C/C" "?/?" "C/C" "C/C" ...
##  $ PZA00081.17   : chr [1:2782] "?/?" "T/T" "T/T" "T/T" ...
##  $ PZA00084.2    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00084.3    : chr [1:2782] "T/T" "T/T" "T/T" "T/T" ...
##  $ PZA00086.8    : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ PZA00088.3    : chr [1:2782] "G/G" "G/G" "G/G" "G/G" ...
##  $ PZA00090.2    : chr [1:2782] "A/A" "A/A" "?/?" "?/?" ...
##  $ PZA00092.1    : chr [1:2782] "?/?" "?/?" "T/T" "?/?" ...
##  $ PZA00092.5    : chr [1:2782] "?/?" "C/C" "C/C" "?/?" ...
##  $ PZA00093.2    : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ PZA00096.26   : chr [1:2782] "T/T" "T/T" "T/T" "T/T" ...
##  $ PZA00097.13   : chr [1:2782] "G/G" "G/G" "G/G" "G/G" ...
##  $ PZA00098.14   : chr [1:2782] "?/?" "?/?" "C/C" "?/?" ...
##  $ PZA00100.10   : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ PZA00100.12   : chr [1:2782] "T/T" "?/?" "T/T" "T/T" ...
##  $ PZA00100.14   : chr [1:2782] "?/?" "A/A" "A/A" "A/A" ...
##  $ PZA00100.9    : chr [1:2782] "C/C" "C/C" "C/C" "?/?" ...
##  $ PZA00103.20   : chr [1:2782] "A/A" "A/A" "A/A" "A/A" ...
##  $ PZA00106.9    : chr [1:2782] "G/G" "G/G" "G/G" "G/G" ...
##  $ PZA00107.18   : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ PZA00108.12   : chr [1:2782] "?/?" "C/C" "C/C" "C/C" ...
##  $ PZA00108.14   : chr [1:2782] "G/G" "G/G" "G/G" "G/G" ...
##  $ PZA00108.15   : chr [1:2782] "A/A" "A/A" "A/A" "A/A" ...
##  $ PZA00109.3    : chr [1:2782] "A/A" "A/A" "?/?" "?/?" ...
##  $ PZA00109.5    : chr [1:2782] "A/A" "?/?" "A/A" "?/?" ...
##  $ PZA00111.2    : chr [1:2782] "T/T" "T/T" "T/T" "T/T" ...
##  $ PZA00111.4    : chr [1:2782] "C/C" "?/?" "C/C" "C/C" ...
##  $ PZA00111.5    : chr [1:2782] "?/?" "?/?" "A/A" "A/A" ...
##  $ PZA00111.6    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00111.8    : chr [1:2782] "T/T" "T/T" "T/T" "T/T" ...
##  $ PZA00114.3    : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00116.2    : chr [1:2782] "C/T" "C/T" "C/T" "C/T" ...
##  $ PZA00119.4    : chr [1:2782] "G/G" "G/G" "G/G" "G/G" ...
##  $ PZA00120.4    : chr [1:2782] "G/G" "G/G" "G/G" "G/G" ...
##  $ PZA00123.1    : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ PZA00125.2    : chr [1:2782] "?/?" "?/?" "?/?" "?/?" ...
##  $ PZA00131.14   : chr [1:2782] "C/C" "C/C" "C/C" "C/C" ...
##  $ PZA00132.17   : chr [1:2782] "T/T" "T/T" "T/T" "T/T" ...
##   [list output truncated]
##  - attr(*, "spec")=
##   .. cols(
##   ..   Sample_ID = col_character(),
```

```
##   ..   JG_OTU = col_character(),
##   ..   Group = col_character(),
##   ..   abph1.20 = col_character(),
##   ..   abph1.22 = col_character(),
##   ..   ae1.3 = col_character(),
##   ..   ae1.4 = col_character(),
##   ..   ae1.5 = col_character(),
##   ..   an1.4 = col_character(),
##   ..   ba1.6 = col_character(),
##   ..   ba1.9 = col_character(),
##   ..   bt2.5 = col_character(),
##   ..   bt2.7 = col_character(),
##   ..   bt2.8 = col_character(),
##   ..   Fea2.1 = col_character(),
##   ..   Fea2.5 = col_character(),
##   ..   id1.3 = col_character(),
##   ..   lg2.11 = col_character(),
##   ..   lg2.2 = col_character(),
##   ..   pbf1.1 = col_character(),
##   ..   pbf1.2 = col_character(),
##   ..   pbf1.3 = col_character(),
##   ..   pbf1.5 = col_character(),
##   ..   pbf1.6 = col_character(),
##   ..   pbf1.7 = col_character(),
##   ..   pbf1.8 = col_character(),
##   ..   PZA00003.11 = col_character(),
##   ..   PZA00004.2 = col_character(),
##   ..   PZA00005.8 = col_character(),
##   ..   PZA00005.9 = col_character(),
##   ..   PZA00006.13 = col_character(),
##   ..   PZA00006.14 = col_character(),
##   ..   PZA00008.1 = col_character(),
##   ..   PZA00010.5 = col_character(),
##   ..   PZA00013.10 = col_character(),
##   ..   PZA00013.11 = col_character(),
##   ..   PZA00013.9 = col_character(),
##   ..   PZA00015.4 = col_character(),
##   ..   PZA00017.1 = col_character(),
##   ..   PZA00018.5 = col_character(),
##   ..   PZA00029.11 = col_character(),
##   ..   PZA00029.12 = col_character(),
##   ..   PZA00030.11 = col_character(),
##   ..   PZA00031.5 = col_character(),
##   ..   PZA00041.3 = col_character(),
##   ..   PZA00042.2 = col_character(),
##   ..   PZA00042.5 = col_character(),
##   ..   PZA00043.7 = col_character(),
##   ..   PZA00045.1 = col_character(),
##   ..   PZA00047.2 = col_character(),
##   ..   PZA00049.12 = col_character(),
##   ..   PZA00050.9 = col_character(),
##   ..   PZA00051.2 = col_character(),
```

```
##   ..   PZA00058.5 = col_character(),
##   ..   PZA00058.6 = col_character(),
##   ..   PZA00060.2 = col_character(),
##   ..   PZA00061.1 = col_character(),
##   ..   PZA00065.2 = col_character(),
##   ..   PZA00069.4 = col_character(),
##   ..   PZA00070.5 = col_character(),
##   ..   PZA00078.2 = col_character(),
##   ..   PZA00079.1 = col_character(),
##   ..   PZA00081.17 = col_character(),
##   ..   PZA00084.2 = col_character(),
##   ..   PZA00084.3 = col_character(),
##   ..   PZA00086.8 = col_character(),
##   ..   PZA00088.3 = col_character(),
##   ..   PZA00090.2 = col_character(),
##   ..   PZA00092.1 = col_character(),
##   ..   PZA00092.5 = col_character(),
##   ..   PZA00093.2 = col_character(),
##   ..   PZA00096.26 = col_character(),
##   ..   PZA00097.13 = col_character(),
##   ..   PZA00098.14 = col_character(),
##   ..   PZA00100.10 = col_character(),
##   ..   PZA00100.12 = col_character(),
##   ..   PZA00100.14 = col_character(),
##   ..   PZA00100.9 = col_character(),
##   ..   PZA00103.20 = col_character(),
##   ..   PZA00106.9 = col_character(),
##   ..   PZA00107.18 = col_character(),
##   ..   PZA00108.12 = col_character(),
##   ..   PZA00108.14 = col_character(),
##   ..   PZA00108.15 = col_character(),
##   ..   PZA00109.3 = col_character(),
##   ..   PZA00109.5 = col_character(),
##   ..   PZA00111.2 = col_character(),
##   ..   PZA00111.4 = col_character(),
##   ..   PZA00111.5 = col_character(),
##   ..   PZA00111.6 = col_character(),
##   ..   PZA00111.8 = col_character(),
##   ..   PZA00114.3 = col_character(),
##   ..   PZA00116.2 = col_character(),
##   ..   PZA00119.4 = col_character(),
##   ..   PZA00120.4 = col_character(),
##   ..   PZA00123.1 = col_character(),
##   ..   PZA00125.2 = col_character(),
##   ..   PZA00131.14 = col_character(),
##   ..   PZA00132.17 = col_character(),
##   ..   PZA00132.18 = col_character(),
##   ..   PZA00132.3 = col_character(),
##   ..   PZA00135.6 = col_character(),
##   ..   PZA00137.2 = col_character(),
##   ..   PZA00139.14 = col_character(),
##   ..   PZA00140.10 = col_character(),
```

```
##   ..   PZA00140.6 = col_character(),
##   ..   PZA00140.9 = col_character(),
##   ..   PZA00142.6 = col_character(),
##   ..   PZA00148.2 = col_character(),
##   ..   PZA00153.3 = col_character(),
##   ..   PZA00153.6 = col_character(),
##   ..   PZA00163.4 = col_character(),
##   ..   PZA00164.1 = col_character(),
##   ..   PZA00164.2 = col_character(),
##   ..   PZA00164.3 = col_character(),
##   ..   PZA00166.1 = col_character(),
##   ..   PZA00166.3 = col_character(),
##   ..   PZA00170.1 = col_character(),
##   ..   PZA00170.3 = col_character(),
##   ..   PZA00170.4 = col_character(),
##   ..   PZA00174.1 = col_character(),
##   ..   PZA00174.2 = col_character(),
##   ..   PZA00175.2 = col_character(),
##   ..   PZA00176.8 = col_character(),
##   ..   PZA00177.4 = col_character(),
##   ..   PZA00178.3 = col_character(),
##   ..   PZA00182.3 = col_character(),
##   ..   PZA00182.4 = col_character(),
##   ..   PZA00184.1 = col_character(),
##   ..   PZA00184.4 = col_character(),
##   ..   PZA00188.1 = col_character(),
##   ..   PZA00188.3 = col_character(),
##   ..   PZA00191.5 = col_character(),
##   ..   PZA00192.6 = col_character(),
##   ..   PZA00192.7 = col_character(),
##   ..   PZA00193.2 = col_character(),
##   ..   PZA00198.39 = col_character(),
##   ..   PZA00200.11 = col_character(),
##   ..   PZA00200.17 = col_character(),
##   ..   PZA00200.9 = col_character(),
##   ..   PZA00201.2 = col_character(),
##   ..   PZA00204.1 = col_character(),
##   ..   PZA00210.1 = col_character(),
##   ..   PZA00210.6 = col_character(),
##   ..   PZA00211.7 = col_character(),
##   ..   PZA00212.1 = col_character(),
##   ..   PZA00213.19 = col_character(),
##   ..   PZA00214.1 = col_character(),
##   ..   PZA00216.9 = col_character(),
##   ..   PZA00218.1 = col_character(),
##   ..   PZA00218.6 = col_character(),
##   ..   PZA00219.7 = col_character(),
##   ..   PZA00220.11 = col_character(),
##   ..   PZA00220.12 = col_character(),
##   ..   PZA00221.7 = col_character(),
##   ..   PZA00225.8 = col_character(),
##   ..   PZA00226.7 = col_character(),
```

```
##   ..   PZA00227.8 = col_character(),
##   ..   PZA00230.5 = col_character(),
##   ..   PZA00232.24 = col_character(),
##   ..   PZA00234.21 = col_character(),
##   ..   PZA00235.6 = col_character(),
##   ..   PZA00235.8 = col_character(),
##   ..   PZA00237.2 = col_character(),
##   ..   PZA00237.7 = col_character(),
##   ..   PZA00237.8 = col_character(),
##   ..   PZA00238.3 = col_character(),
##   ..   PZA00240.9 = col_character(),
##   ..   PZA00241.6 = col_character(),
##   ..   PZA00243.27 = col_character(),
##   ..   PZA00245.14 = col_character(),
##   ..   PZA00245.16 = col_character(),
##   ..   PZA00245.17 = col_character(),
##   ..   PZA00245.18 = col_character(),
##   ..   PZA00245.19 = col_character(),
##   ..   PZA00249.2 = col_character(),
##   ..   PZA00250.1 = col_character(),
##   ..   PZA00251.1 = col_character(),
##   ..   PZA00254.3 = col_character(),
##   ..   PZA00255.15 = col_character(),
##   ..   PZA00255.17 = col_character(),
##   ..   PZA00256.16 = col_character(),
##   ..   PZA00256.21 = col_character(),
##   ..   PZA00256.23 = col_character(),
##   ..   PZA00257.11 = col_character(),
##   ..   PZA00257.22 = col_character(),
##   ..   PZA00261.6 = col_character(),
##   ..   PZA00263.14 = col_character(),
##   ..   PZA00266.5 = col_character(),
##   ..   PZA00270.3 = col_character(),
##   ..   PZA00273.1 = col_character(),
##   ..   PZA00274.7 = col_character(),
##   ..   PZA00277.17 = col_character(),
##   ..   PZA00277.9 = col_character(),
##   ..   PZA00280.14 = col_character(),
##   ..   PZA00287.1 = col_character(),
##   ..   PZA00289.11 = col_character(),
##   ..   PZA00294.20 = col_character(),
##   ..   PZA00296.6 = col_character(),
##   ..   PZA00297.2 = col_character(),
##   ..   PZA00297.3 = col_character(),
##   ..   PZA00297.4 = col_character(),
##   ..   PZA00298.4 = col_character(),
##   ..   PZA00298.5 = col_character(),
##   ..   PZA00299.2 = col_character(),
##   ..   PZA00300.12 = col_character(),
##   ..   PZA00300.13 = col_character(),
##   ..   PZA00300.14 = col_character(),
##   ..   PZA00300.16 = col_character(),
```

```
##   ..   PZA00301.3 = col_character(),
##   ..   PZA00303.19 = col_character(),
##   ..   PZA00303.21 = col_character(),
##   ..   PZA00307.12 = col_character(),
##   ..   PZA00307.14 = col_character(),
##   ..   PZA00307.17 = col_character(),
##   ..   PZA00309.2 = col_character(),
##   ..   PZA00310.5 = col_character(),
##   ..   PZA00314.6 = col_character(),
##   ..   PZA00314.8 = col_character(),
##   ..   PZA00315.1 = col_character(),
##   ..   PZA00315.6 = col_character(),
##   ..   PZA00318.2 = col_character(),
##   ..   PZA00323.3 = col_character(),
##   ..   PZA00323.4 = col_character(),
##   ..   PZA00326.16 = col_character(),
##   ..   PZA00326.18 = col_character(),
##   ..   PZA00326.19 = col_character(),
##   ..   PZA00332.8 = col_character(),
##   ..   PZA00332.9 = col_character(),
##   ..   PZA00334.2 = col_character(),
##   ..   PZA00335.12 = col_character(),
##   ..   PZA00337.3 = col_character(),
##   ..   PZA00337.4 = col_character(),
##   ..   PZA00337.5 = col_character(),
##   ..   PZA00342.9 = col_character(),
##   ..   PZA00344.10 = col_character(),
##   ..   PZA00345.15 = col_character(),
##   ..   PZA00346.1 = col_character(),
##   ..   PZA00346.2 = col_character(),
##   ..   PZA00346.3 = col_character(),
##   ..   PZA00349.3 = col_character(),
##   ..   PZA00349.5 = col_character(),
##   ..   PZA00350.2 = col_character(),
##   ..   PZA00352.22 = col_character(),
##   ..   PZA00355.1 = col_character(),
##   ..   PZA00355.2 = col_character(),
##   ..   PZA00356.9 = col_character(),
##   ..   PZA00364.5 = col_character(),
##   ..   PZA00364.6 = col_character(),
##   ..   PZA00367.2 = col_character(),
##   ..   PZA00369.1 = col_character(),
##   ..   PZA00370.1 = col_character(),
##   ..   PZA00370.5 = col_character(),
##   ..   PZA00380.5 = col_character(),
##   ..   PZA00380.7 = col_character(),
##   ..   PZA00381.3 = col_character(),
##   ..   PZA00381.4 = col_character(),
##   ..   PZA00381.5 = col_character(),
##   ..   PZA00382.17 = col_character(),
##   ..   PZA00385.3 = col_character(),
##   ..   PZA00386.3 = col_character(),
```

```
##   ..   PZA00390.6 = col_character(),
##   ..   PZA00391.2 = col_character(),
##   ..   PZA00392.3 = col_character(),
##   ..   PZA00392.4 = col_character(),
##   ..   PZA00393.1 = col_character(),
##   ..   PZA00393.4 = col_character(),
##   ..   PZA00394.11 = col_character(),
##   ..   PZA00395.1 = col_character(),
##   ..   PZA00395.2 = col_character(),
##   ..   PZA00396.12 = col_character(),
##   ..   PZA00401.11 = col_character(),
##   ..   PZA00401.6 = col_character(),
##   ..   PZA00406.1 = col_character(),
##   ..   PZA00407.9 = col_character(),
##   ..   PZA00408.7 = col_character(),
##   ..   PZA00409.3 = col_character(),
##   ..   PZA00411.1 = col_character(),
##   ..   PZA00411.4 = col_character(),
##   ..   PZA00411.5 = col_character(),
##   ..   PZA00413.17 = col_character(),
##   ..   PZA00413.18 = col_character(),
##   ..   PZA00413.21 = col_character(),
##   ..   PZA00417.2 = col_character(),
##   ..   PZA00417.3 = col_character(),
##   ..   PZA00419.1 = col_character(),
##   ..   PZA00420.4 = col_character(),
##   ..   PZA00422.2 = col_character(),
##   ..   PZA00422.5 = col_character(),
##   ..   PZA00422.6 = col_character(),
##   ..   PZA00423.16 = col_character(),
##   ..   PZA00423.17 = col_character(),
##   ..   PZA00424.1 = col_character(),
##   ..   PZA00425.4 = col_character(),
##   ..   PZA00425.9 = col_character(),
##   ..   PZA00429.1 = col_character(),
##   ..   PZA00433.5 = col_character(),
##   ..   PZA00436.7 = col_character(),
##   ..   PZA00439.6 = col_character(),
##   ..   PZA00440.1 = col_character(),
##   ..   PZA00442.3 = col_character(),
##   ..   PZA00442.4 = col_character(),
##   ..   PZA00442.5 = col_character(),
##   ..   PZA00442.6 = col_character(),
##   ..   PZA00444.1 = col_character(),
##   ..   PZA00444.5 = col_character(),
##   ..   PZA00445.18 = col_character(),
##   ..   PZA00449.2 = col_character(),
##   ..   PZA00452.4 = col_character(),
##   ..   PZA00458.6 = col_character(),
##   ..   PZA00459.5 = col_character(),
##   ..   PZA00460.3 = col_character(),
##   ..   PZA00460.5 = col_character(),
```

```
##   ..   PZA00460.7 = col_character(),
##   ..   PZA00462.2 = col_character(),
##   ..   PZA00463.3 = col_character(),
##   ..   PZA00466.1 = col_character(),
##   ..   PZA00468.11 = col_character(),
##   ..   PZA00468.7 = col_character(),
##   ..   PZA00470.1 = col_character(),
##   ..   PZA00471.2 = col_character(),
##   ..   PZA00471.3 = col_character(),
##   ..   PZA00471.4 = col_character(),
##   ..   PZA00472.2 = col_character(),
##   ..   PZA00477.10 = col_character(),
##   ..   PZA00477.11 = col_character(),
##   ..   PZA00477.5 = col_character(),
##   ..   PZA00477.9 = col_character(),
##   ..   PZA00478.10 = col_character(),
##   ..   PZA00478.11 = col_character(),
##   ..   PZA00478.7 = col_character(),
##   ..   PZA00478.9 = col_character(),
##   ..   PZA00480.10 = col_character(),
##   ..   PZA00481.7 = col_character(),
##   ..   PZA00484.5 = col_character(),
##   ..   PZA00485.2 = col_character(),
##   ..   PZA00486.2 = col_character(),
##   ..   PZA00487.16 = col_character(),
##   ..   PZA00487.24 = col_character(),
##   ..   PZA00487.26 = col_character(),
##   ..   PZA00489.1 = col_character(),
##   ..   PZA00493.1 = col_character(),
##   ..   PZA00493.2 = col_character(),
##   ..   PZA00493.5 = col_character(),
##   ..   PZA00495.3 = col_character(),
##   ..   PZA00495.4 = col_character(),
##   ..   PZA00495.6 = col_character(),
##   ..   PZA00496.1 = col_character(),
##   ..   PZA00497.1 = col_character(),
##   ..   PZA00497.4 = col_character(),
##   ..   PZA00498.4 = col_character(),
##   ..   PZA00499.10 = col_character(),
##   ..   PZA00499.12 = col_character(),
##   ..   PZA00499.3 = col_character(),
##   ..   PZA00501.12 = col_character(),
##   ..   PZA00501.14 = col_character(),
##   ..   PZA00502.5 = col_character(),
##   ..   PZA00503.5 = col_character(),
##   ..   PZA00504.1 = col_character(),
##   ..   PZA00504.2 = col_character(),
##   ..   PZA00505.4 = col_character(),
##   ..   PZA00505.8 = col_character(),
##   ..   PZA00510.2 = col_character(),
##   ..   PZA00510.3 = col_character(),
##   ..   PZA00514.1 = col_character(),
```

```
##   ..   PZA00514.6 = col_character(),
##   ..   PZA00514.7 = col_character(),
##   ..   PZA00515.14 = col_character(),
##   ..   PZA00516.3 = col_character(),
##   ..   PZA00517.6 = col_character(),
##   ..   PZA00522.12 = col_character(),
##   ..   PZA00523.2 = col_character(),
##   ..   PZA00525.16 = col_character(),
##   ..   PZA00525.2 = col_character(),
##   ..   PZA00527.6 = col_character(),
##   ..   PZA00527.9 = col_character(),
##   ..   PZA00529.3 = col_character(),
##   ..   PZA00531.1 = col_character(),
##   ..   PZA00533.3 = col_character(),
##   ..   PZA00533.4 = col_character(),
##   ..   PZA00533.5 = col_character(),
##   ..   PZA00533.6 = col_character(),
##   ..   PZA00534.2 = col_character(),
##   ..   PZA00536.2 = col_character(),
##   ..   PZA00538.12 = col_character(),
##   ..   PZA00538.16 = col_character(),
##   ..   PZA00538.8 = col_character(),
##   ..   PZA00543.2 = col_character(),
##   ..   PZA00543.4 = col_character(),
##   ..   PZA00543.5 = col_character(),
##   ..   PZA00545.21 = col_character(),
##   ..   PZA00545.22 = col_character(),
##   ..   PZA00545.4 = col_character(),
##   ..   PZA00547.13 = col_character(),
##   ..   PZA00547.18 = col_character(),
##   ..   PZA00552.4 = col_character(),
##   ..   PZA00560.1 = col_character(),
##   ..   PZA00560.2 = col_character(),
##   ..   PZA00562.4 = col_character(),
##   ..   PZA00565.3 = col_character(),
##   ..   PZA00566.5 = col_character(),
##   ..   PZA00568.19 = col_character(),
##   ..   PZA00573.3 = col_character(),
##   ..   PZA00578.1 = col_character(),
##   ..   PZA00579.6 = col_character(),
##   ..   PZA00582.4 = col_character(),
##   ..   PZA00586.1 = col_character(),
##   ..   PZA00587.3 = col_character(),
##   ..   PZA00587.6 = col_character(),
##   ..   PZA00588.2 = col_character(),
##   ..   PZA00588.4 = col_character(),
##   ..   PZA00589.10 = col_character(),
##   ..   PZA00589.8 = col_character(),
##   ..   PZA00589.9 = col_character(),
##   ..   PZA00593.2 = col_character(),
##   ..   PZA00595.3 = col_character(),
##   ..   PZA00600.11 = col_character(),
```

```
##   ..   PZA00603.1 = col_character(),
##   ..   PZA00608.1 = col_character(),
##   ..   PZA00608.5 = col_character(),
##   ..   PZA00610.18 = col_character(),
##   ..   PZA00610.9 = col_character(),
##   ..   PZA00613.22 = col_character(),
##   ..   PZA00614.12 = col_character(),
##   ..   PZA00615.3 = col_character(),
##   ..   PZA00615.6 = col_character(),
##   ..   PZA00615.8 = col_character(),
##   ..   PZA00617.16 = col_character(),
##   ..   PZA00618.22 = col_character(),
##   ..   PZA00620.2 = col_character(),
##   ..   PZA00621.2 = col_character(),
##   ..   PZA00622.1 = col_character(),
##   ..   PZA00622.2 = col_character(),
##   ..   PZA00623.2 = col_character(),
##   ..   PZA00626.3 = col_character(),
##   ..   PZA00626.4 = col_character(),
##   ..   PZA00630.9 = col_character(),
##   ..   PZA00636.5 = col_character(),
##   ..   PZA00636.6 = col_character(),
##   ..   PZA00637.4 = col_character(),
##   ..   PZA00639.12 = col_character(),
##   ..   PZA00639.13 = col_character(),
##   ..   PZA00639.15 = col_character(),
##   ..   PZA00641.7 = col_character(),
##   ..   PZA00641.8 = col_character(),
##   ..   PZA00644.11 = col_character(),
##   ..   PZA00647.9 = col_character(),
##   ..   PZA00650.8 = col_character(),
##   ..   PZA00654.10 = col_character(),
##   ..   PZA00654.12 = col_character(),
##   ..   PZA00655.1 = col_character(),
##   ..   PZA00656.15 = col_character(),
##   ..   PZA00656.16 = col_character(),
##   ..   PZA00656.18 = col_character(),
##   ..   PZA00656.4 = col_character(),
##   ..   PZA00658.19 = col_character(),
##   ..   PZA00658.23 = col_character(),
##   ..   PZA00662.3 = col_character(),
##   ..   PZA00665.6 = col_character(),
##   ..   PZA00667.1 = col_character(),
##   ..   PZA00672.6 = col_character(),
##   ..   PZA00672.8 = col_character(),
##   ..   PZA00673.2 = col_character(),
##   ..   PZA00674.3 = col_character(),
##   ..   PZA00676.2 = col_character(),
##   ..   PZA00680.1 = col_character(),
##   ..   PZA00680.3 = col_character(),
##   ..   PZA00682.2 = col_character(),
##   ..   PZA00684.12 = col_character(),
```

```
##   ..   PZA00686.8 = col_character(),
##   ..   PZA00692.5 = col_character(),
##   ..   PZA00693.3 = col_character(),
##   ..   PZA00695.1 = col_character(),
##   ..   PZA00698.4 = col_character(),
##   ..   PZA00700.3 = col_character(),
##   ..   PZA00704.11 = col_character(),
##   ..   PZA00705.5 = col_character(),
##   ..   PZA00706.16 = col_character(),
##   ..   PZA00710.1 = col_character(),
##   ..   PZA00710.16 = col_character(),
##   ..   PZA00712.4 = col_character(),
##   ..   PZA00715.3 = col_character(),
##   ..   PZA00717.14 = col_character(),
##   ..   PZA00719.1 = col_character(),
##   ..   PZA00719.2 = col_character(),
##   ..   PZA00719.3 = col_character(),
##   ..   PZA00720.2 = col_character(),
##   ..   PZA00720.3 = col_character(),
##   ..   PZA00721.4 = col_character(),
##   ..   PZA00721.5 = col_character(),
##   ..   PZA00725.4 = col_character(),
##   ..   PZA00726.6 = col_character(),
##   ..   PZA00726.7 = col_character(),
##   ..   PZA00726.9 = col_character(),
##   ..   PZA00727.11 = col_character(),
##   ..   PZA00727.12 = col_character(),
##   ..   PZA00729.18 = col_character(),
##   ..   PZA00729.19 = col_character(),
##   ..   PZA00730.2 = col_character(),
##   ..   PZA00731.6 = col_character(),
##   ..   PZA00731.7 = col_character(),
##   ..   PZA01104.1 = col_character(),
##   ..   PZA01149.1 = col_character(),
##   ..   PZA01149.3 = col_character(),
##   ..   PZA01182.1 = col_character(),
##   ..   PZA01240.1 = col_character(),
##   ..   PZA01240.2 = col_character(),
##   ..   PZA01420.1 = col_character(),
##   ..   PZA01420.2 = col_character(),
##   ..   PZA01420.3 = col_character(),
##   ..   PZA01474.2 = col_character(),
##   ..   PZA01637.2 = col_character(),
##   ..   PZA01637.3 = col_character(),
##   ..   PZA01637.4 = col_character(),
##   ..   PZA01725.1 = col_character(),
##   ..   PZA01725.2 = col_character(),
##   ..   PZA01782.2 = col_character(),
##   ..   PZA01782.3 = col_character(),
##   ..   PZA01782.4 = col_character(),
##   ..   PZA02789.31 = col_character(),
##   ..   PZA02789.36 = col_character(),
```

```
##   ..   PZA02791.6 = col_character(),
##   ..   PZA02792.16 = col_character(),
##   ..   PZA02792.9 = col_character(),
##   ..   PZA02806.4 = col_character(),
##   ..   PZA02806.9 = col_character(),
##   ..   PZA02807.5 = col_character(),
##   ..   PZA02808.12 = col_character(),
##   ..   PZA02808.16 = col_character(),
##   ..   PZA02819.35 = col_character(),
##   ..   PZA02820.6 = col_character(),
##   ..   PZA02822.2 = col_character(),
##   ..   PZA02824.1 = col_character(),
##   ..   PZA02824.3 = col_character(),
##   ..   PZA02825.8 = col_character(),
##   ..   PZA02831.5 = col_character(),
##   ..   PZA02837.5 = col_character(),
##   ..   PZA02844.1 = col_character(),
##   ..   PZA02850.18 = col_character(),
##   ..   PZA02850.4 = col_character(),
##   ..   PZA02853.10 = col_character(),
##   ..   PZA02853.7 = col_character(),
##   ..   PZA02856.1 = col_character(),
##   ..   PZA02862.3 = col_character(),
##   ..   PZA02865.11 = col_character(),
##   ..   PZA02869.2 = col_character(),
##   ..   PZA02869.8 = col_character(),
##   ..   PZA02872.1 = col_character(),
##   ..   PZA02872.3 = col_character(),
##   ..   PZA02878.12 = col_character(),
##   ..   PZA02888.3 = col_character(),
##   ..   PZA02890.3 = col_character(),
##   ..   PZA02890.4 = col_character(),
##   ..   PZA02890.5 = col_character(),
##   ..   PZA02894.1 = col_character(),
##   ..   PZA02897.12 = col_character(),
##   ..   PZA02906.12 = col_character(),
##   ..   PZA02906.7 = col_character(),
##   ..   PZA02921.9 = col_character(),
##   ..   PZA02923.7 = col_character(),
##   ..   PZA02927.1 = col_character(),
##   ..   PZA02938.5 = col_character(),
##   ..   PZA02939.6 = col_character(),
##   ..   PZA02940.3 = col_character(),
##   ..   PZA02941.3 = col_character(),
##   ..   PZA02941.6 = col_character(),
##   ..   PZA02941.8 = col_character(),
##   ..   PZA02947.2 = col_character(),
##   ..   PZA02948.19 = col_character(),
##   ..   PZA02948.21 = col_character(),
##   ..   PZA02948.22 = col_character(),
##   ..   PZA02949.22 = col_character(),
##   ..   PZA02949.26 = col_character(),
```

```
##   ..   PZA02952.10 = col_character(),
##   ..   PZA02954.2 = col_character(),
##   ..   PZA02955.3 = col_character(),
##   ..   PZA02958.17 = col_character(),
##   ..   PZA02959.7 = col_character(),
##   ..   PZA02961.1 = col_character(),
##   ..   PZA02962.13 = col_character(),
##   ..   PZA02963.5 = col_character(),
##   ..   PZA02966.11 = col_character(),
##   ..   PZA02968.4 = col_character(),
##   ..   PZA02969.11 = col_character(),
##   ..   PZA02970.9 = col_character(),
##   ..   PZA02972.1 = col_character(),
##   ..   PZA02982.5 = col_character(),
##   ..   PZA02982.6 = col_character(),
##   ..   PZA02983.38 = col_character(),
##   ..   PZA02984.7 = col_character(),
##   ..   PZA02988.2 = col_character(),
##   ..   PZA02993.5 = col_character(),
##   ..   PZA02997.16 = col_character(),
##   ..   PZA02997.19 = col_character(),
##   ..   PZA03001.15 = col_character(),
##   ..   PZA03001.18 = col_character(),
##   ..   PZA03001.9 = col_character(),
##   ..   PZA03009.5 = col_character(),
##   ..   PZA03009.6 = col_character(),
##   ..   PZA03009.7 = col_character(),
##   ..   PZA03009.8 = col_character(),
##   ..   PZA03011.6 = col_character(),
##   ..   PZA03012.10 = col_character(),
##   ..   PZA03013.7 = col_character(),
##   ..   PZA03013.8 = col_character(),
##   ..   PZA03014.10 = col_character(),
##   ..   PZA03014.21 = col_character(),
##   ..   PZA03014.24 = col_character(),
##   ..   PZA03017.10 = col_character(),
##   ..   PZA03017.11 = col_character(),
##   ..   PZA03024.16 = col_character(),
##   ..   PZA03024.18 = col_character(),
##   ..   PZA03024.7 = col_character(),
##   ..   PZA03028.5 = col_character(),
##   ..   PZA03032.16 = col_character(),
##   ..   PZA03034.1 = col_character(),
##   ..   PZA03035.5 = col_character(),
##   ..   PZA03037.8 = col_character(),
##   ..   PZA03037.9 = col_character(),
##   ..   PZA03041.8 = col_character(),
##   ..   PZA03042.1 = col_character(),
##   ..   PZA03042.5 = col_character(),
##   ..   PZA03046.2 = col_character(),
##   ..   PZA03046.3 = col_character(),
##   ..   PZA03047.12 = col_character(),
```

```
##   ..   PZA03047.20 = col_character(),
##   ..   PZA03047.22 = col_character(),
##   ..   PZA03048.16 = col_character(),
##   ..   PZA03048.17 = col_character(),
##   ..   PZA03049.23 = col_character(),
##   ..   PZA03051.1 = col_character(),
##   ..   PZA03051.3 = col_character(),
##   ..   PZA03052.15 = col_character(),
##   ..   PZA03054.3 = col_character(),
##   ..   PZA03054.5 = col_character(),
##   ..   PZA03058.17 = col_character(),
##   ..   PZA03062.15 = col_character(),
##   ..   PZA03062.7 = col_character(),
##   ..   PZA03063.17 = col_character(),
##   ..   PZA03063.18 = col_character(),
##   ..   PZA03064.6 = col_character(),
##   ..   PZA03067.17 = col_character(),
##   ..   PZA03067.20 = col_character(),
##   ..   PZA03068.11 = col_character(),
##   ..   PZA03068.13 = col_character(),
##   ..   PZA03069.6 = col_character(),
##   ..   PZA03073.23 = col_character(),
##   ..   PZA03073.24 = col_character(),
##   ..   PZA03074.24 = col_character(),
##   ..   PZA03078.33 = col_character(),
##   ..   PZA03081.1 = col_character(),
##   ..   PZA03081.10 = col_character(),
##   ..   PZA03081.11 = col_character(),
##   ..   PZA03081.13 = col_character(),
##   ..   PZA03081.6 = col_character(),
##   ..   PZA03083.7 = col_character(),
##   ..   PZA03089.12 = col_character(),
##   ..   PZA03090.31 = col_character(),
##   ..   PZA03092.7 = col_character(),
##   ..   PZA03094.18 = col_character(),
##   ..   PZA03094.6 = col_character(),
##   ..   PZA03095.1 = col_character(),
##   ..   PZA03095.2 = col_character(),
##   ..   PZA03095.3 = col_character(),
##   ..   PZA03097.4 = col_character(),
##   ..   PZA03097.7 = col_character(),
##   ..   PZA03097.9 = col_character(),
##   ..   PZA03102.10 = col_character(),
##   ..   PZA03102.2 = col_character(),
##   ..   PZA03102.9 = col_character(),
##   ..   PZA03137.1 = col_character(),
##   ..   PZA03172.2 = col_character(),
##   ..   PZA03223.3 = col_character(),
##   ..   PZA03258.2 = col_character(),
##   ..   PZA03283.2 = col_character(),
##   ..   PZA03284.3 = col_character(),
##   ..   PZA03290.1 = col_character(),
```

```
##   ..   PZA03290.2 = col_character(),
##   ..   PZA03295.4 = col_character(),
##   ..   PZA03296.6 = col_character(),
##   ..   PZA03296.7 = col_character(),
##   ..   PZA03298.1 = col_character(),
##   ..   PZA03298.2 = col_character(),
##   ..   PZA03301.2 = col_character(),
##   ..   PZA03301.4 = col_character(),
##   ..   PZA03302.1 = col_character(),
##   ..   PZA03305.6 = col_character(),
##   ..   PZA03305.7 = col_character(),
##   ..   PZA03311.2 = col_character(),
##   ..   PZA03311.3 = col_character(),
##   ..   PZA03311.4 = col_character(),
##   ..   PZA03311.5 = col_character(),
##   ..   PZA03312.1 = col_character(),
##   ..   PZA03312.2 = col_character(),
##   ..   PZA03316.2 = col_character(),
##   ..   PZA03317.1 = col_character(),
##   ..   PZA03319.3 = col_character(),
##   ..   PZA03319.4 = col_character(),
##   ..   PZA03320.3 = col_character(),
##   ..   PZA03320.4 = col_character(),
##   ..   PZA03328.5 = col_character(),
##   ..   PZA03329.1 = col_character(),
##   ..   PZA03329.2 = col_character(),
##   ..   PZA03333.3 = col_character(),
##   ..   PZA03335.2 = col_character(),
##   ..   PZA03335.3 = col_character(),
##   ..   PZA03337.1 = col_character(),
##   ..   PZA03338.5 = col_character(),
##   ..   PZA03340.2 = col_character(),
##   ..   PZA03342.2 = col_character(),
##   ..   PZA03344.4 = col_character(),
##   ..   PZA03344.5 = col_character(),
##   ..   PZA03344.6 = col_character(),
##   ..   PZA03345.1 = col_character(),
##   ..   PZA03345.2 = col_character(),
##   ..   PZA03345.4 = col_character(),
##   ..   PZA03347.1 = col_character(),
##   ..   PZA03348.1 = col_character(),
##   ..   PZA03349.1 = col_character(),
##   ..   PZA03349.9 = col_character(),
##   ..   PZA03767.1 = col_character(),
##   ..   PZA03767.4 = col_character(),
##   ..   PZA03767.5 = col_character(),
##   ..   PZA03773.2 = col_character(),
##   ..   PZA03773.3 = col_character(),
##   ..   PZA03774.1 = col_character(),
##   ..   PZA03774.10 = col_character(),
##   ..   PZA03774.2 = col_character(),
##   ..   PZA03774.4 = col_character(),
```

```
##   ..   PZA03774.5 = col_character(),
##   ..   PZA03774.6 = col_character(),
##   ..   PZA03774.8 = col_character(),
##   ..   PZA03774.9 = col_character(),
##   ..   PZA03775.1 = col_character(),
##   ..   PZA03775.11 = col_character(),
##   ..   PZA03775.2 = col_character(),
##   ..   PZA03775.3 = col_character(),
##   ..   PZA03775.4 = col_character(),
##   ..   PZA03775.6 = col_character(),
##   ..   PZA03775.7 = col_character(),
##   ..   PZA03775.8 = col_character(),
##   ..   PZA03775.9 = col_character(),
##   ..   PZA03781.1 = col_character(),
##   ..   PZA03781.2 = col_character(),
##   ..   PZA03781.3 = col_character(),
##   ..   PZA03781.4 = col_character(),
##   ..   PZA03781.5 = col_character(),
##   ..   PZA03781.6 = col_character(),
##   ..   PZA03781.7 = col_character(),
##   ..   PZA03781.8 = col_character(),
##   ..   PZA03782.1 = col_character(),
##   ..   PZA03782.3 = col_character(),
##   ..   PZA03786.1 = col_character(),
##   ..   PZA03786.2 = col_character(),
##   ..   PZA03789.1 = col_character(),
##   ..   PZA03789.2 = col_character(),
##   ..   PZA03789.4 = col_character(),
##   ..   PZB00011.4 = col_character(),
##   ..   PZB00011.5 = col_character(),
##   ..   PZB00041.2 = col_character(),
##   ..   PZB00041.4 = col_character(),
##   ..   PZB00049.2 = col_character(),
##   ..   PZB00049.4 = col_character(),
##   ..   PZB00049.7 = col_character(),
##   ..   PZB00055.1 = col_character(),
##   ..   PZB00060.4 = col_character(),
##   ..   PZB00062.6 = col_character(),
##   ..   PZB00062.7 = col_character(),
##   ..   PZB00062.8 = col_character(),
##   ..   PZB00067.2 = col_character(),
##   ..   PZB00067.3 = col_character(),
##   ..   PZB00067.4 = col_character(),
##   ..   PZB00067.5 = col_character(),
##   ..   PZB00078.1 = col_character(),
##   ..   PZB00081.2 = col_character(),
##   ..   PZB00081.4 = col_character(),
##   ..   PZB00081.5 = col_character(),
##   ..   PZB00081.7 = col_character(),
##   ..   PZB00092.1 = col_character(),
##   ..   PZB00092.4 = col_character(),
##   ..   PZB00093.3 = col_character(),
```

```
##   ..   PZB00093.4 = col_character(),
##   ..   PZB00093.6 = col_character(),
##   ..   PZB00096.2 = col_character(),
##   ..   PZB00096.3 = col_character(),
##   ..   PZB00136.3 = col_character(),
##   ..   PZB00140.1 = col_character(),
##   ..   PZB00145.2 = col_character(),
##   ..   PZB00149.2 = col_character(),
##   ..   PZB00149.4 = col_character(),
##   ..   PZB00153.1 = col_character(),
##   ..   PZB00153.2 = col_character(),
##   ..   PZB00153.3 = col_character(),
##   ..   PZB00153.5 = col_character(),
##   ..   PZB00160.1 = col_character(),
##   ..   PZB00160.2 = col_character(),
##   ..   PZB00160.4 = col_character(),
##   ..   PZB00165.2 = col_character(),
##   ..   PZB00165.6 = col_character(),
##   ..   PZB00169.4 = col_character(),
##   ..   PZB00169.6 = col_character(),
##   ..   PZB00175.1 = col_character(),
##   ..   PZB00175.2 = col_character(),
##   ..   PZB00175.3 = col_character(),
##   ..   PZB00175.4 = col_character(),
##   ..   PZB00175.5 = col_character(),
##   ..   PZB00180.1 = col_character(),
##   ..   PZB00180.2 = col_character(),
##   ..   PZB00183.3 = col_character(),
##   ..   PZB00188.6 = col_character(),
##   ..   PZB00207.3 = col_character(),
##   ..   PZB00221.3 = col_character(),
##   ..   PZB00221.8 = col_character(),
##   ..   PZB00229.3 = col_character(),
##   ..   PZB00232.1 = col_character(),
##   ..   PZB00232.2 = col_character(),
##   ..   PZB00232.4 = col_character(),
##   ..   PZB00232.5 = col_character(),
##   ..   PZB00379.3 = col_character(),
##   ..   PZB00379.4 = col_character(),
##   ..   PZB00379.5 = col_character(),
##   ..   PZB00393.7 = col_character(),
##   ..   PZB00409.3 = col_character(),
##   ..   PZB00416.2 = col_character(),
##   ..   PZB00416.5 = col_character(),
##   ..   PZB00454.2 = col_character(),
##   ..   PZB00454.3 = col_character(),
##   ..   PZB00454.4 = col_character(),
##   ..   PZB00454.5 = col_character(),
##   ..   PZB00498.2 = col_character(),
##   ..   PZB00498.4 = col_character(),
##   ..   PZB00598.1 = col_character(),
##   ..   PZB00598.2 = col_character(),
```

```
##   ..   PZB00603.3 = col_character(),
##   ..   PZB00603.4 = col_character(),
##   ..   PZB00603.5 = col_character(),
##   ..   PZB00607.2 = col_character(),
##   ..   PZB00761.1 = col_character(),
##   ..   PZB00761.2 = col_character(),
##   ..   PZB00849.2 = col_character(),
##   ..   PZB00849.3 = col_character(),
##   ..   PZB00849.4 = col_character(),
##   ..   PZB00859.1 = col_character(),
##   ..   PZB01109.2 = col_character(),
##   ..   PZB01109.3 = col_character(),
##   ..   PZB01110.1 = col_character(),
##   ..   PZB01110.2 = col_character(),
##   ..   PZB01110.3 = col_character(),
##   ..   PZB01111.6 = col_character(),
##   ..   PZB01111.7 = col_character(),
##   ..   PZB01111.8 = col_character(),
##   ..   PZB01112.3 = col_character(),
##   ..   PZB01112.4 = col_character(),
##   ..   PZB01112.5 = col_character(),
##   ..   PZB01112.6 = col_character(),
##   ..   PZB01113.4 = col_character(),
##   ..   PZB01114.1 = col_character(),
##   ..   PZB01114.3 = col_character(),
##   ..   PZB01115.1 = col_character(),
##   ..   PZB01115.5 = col_character(),
##   ..   PZB01115.6 = col_character(),
##   ..   PZB01116.2 = col_character(),
##   ..   PZB01221.1 = col_character(),
##   ..   PZB01222.1 = col_character(),
##   ..   PZB01222.3 = col_character(),
##   ..   PZB01223.3 = col_character(),
##   ..   PZB01223.4 = col_character(),
##   ..   PZB01223.7 = col_character(),
##   ..   PZB01225.1 = col_character(),
##   ..   PZB01225.2 = col_character(),
##   ..   PZB01225.4 = col_character(),
##   ..   PZB01228.1 = col_character(),
##   ..   PZB01228.3 = col_character(),
##   ..   PZB01228.4 = col_character(),
##   ..   PZB01233.2 = col_character(),
##   ..   PZB01233.3 = col_character(),
##   ..   PZB01238.5 = col_character(),
##   ..   PZB01238.6 = col_character(),
##   ..   PZB01427.1 = col_character(),
##   ..   PZB01427.3 = col_character(),
##   ..   PZB01463.2 = col_character(),
##   ..   PZB01463.3 = col_character(),
##   ..   PZB01463.4 = col_character(),
##   ..   PZD00003.1 = col_character(),
##   ..   PZD00003.3 = col_character(),
```

```
##   ..   PZD00007.1 = col_character(),
##   ..   PZD00008.3 = col_character(),
##   ..   PZD00011.1 = col_character(),
##   ..   PZD00011.3 = col_character(),
##   ..   PZD00011.4 = col_character(),
##   ..   PZD00012.1 = col_character(),
##   ..   PZD00012.2 = col_character(),
##   ..   PZD00012.3 = col_character(),
##   ..   PZD00012.4 = col_character(),
##   ..   PZD00012.5 = col_character(),
##   ..   PZD00013.3 = col_character(),
##   ..   PZD00013.4 = col_character(),
##   ..   PZD00014.3 = col_character(),
##   ..   PZD00017.1 = col_character(),
##   ..   PZD00019.1 = col_character(),
##   ..   PZD00020.2 = col_character(),
##   ..   PZD00020.3 = col_character(),
##   ..   PZD00020.4 = col_character(),
##   ..   PZD00020.6 = col_character(),
##   ..   PZD00021.2 = col_character(),
##   ..   PZD00021.4 = col_character(),
##   ..   PZD00021.5 = col_character(),
##   ..   PZD00022.1 = col_character(),
##   ..   PZD00022.3 = col_character(),
##   ..   PZD00022.4 = col_character(),
##   ..   PZD00024.2 = col_character(),
##   ..   PZD00025.1 = col_character(),
##   ..   PZD00025.2 = col_character(),
##   ..   PZD00030.1 = col_character(),
##   ..   PZD00030.4 = col_character(),
##   ..   PZD00030.5 = col_character(),
##   ..   PZD00030.6 = col_character(),
##   ..   PZD00034.3 = col_character(),
##   ..   PZD00043.1 = col_character(),
##   ..   PZD00043.2 = col_character(),
##   ..   PZD00043.3 = col_character(),
##   ..   PZD00043.4 = col_character(),
##   ..   PZD00044.2 = col_character(),
##   ..   PZD00044.3 = col_character(),
##   ..   PZD00044.4 = col_character(),
##   ..   PZD00045.1 = col_character(),
##   ..   PZD00045.2 = col_character(),
##   ..   PZD00045.3 = col_character(),
##   ..   PZD00045.4 = col_character(),
##   ..   PZD00049.3 = col_character(),
##   ..   PZD00049.4 = col_character(),
##   ..   PZD00049.5 = col_character(),
##   ..   PZD00051.1 = col_character(),
##   ..   PZD00052.3 = col_character(),
##   ..   PZD00052.4 = col_character(),
##   ..   PZD00062.2 = col_character(),
##   ..   PZD00066.1 = col_character(),
```

```
##   ..   PZD00067.1 = col_character(),
##   ..   PZD00067.2 = col_character(),
##   ..   PZD00067.3 = col_character(),
##   ..   PZD00068.1 = col_character(),
##   ..   PZD00069.2 = col_character(),
##   ..   PZD00069.3 = col_character(),
##   ..   PZD00069.4 = col_character(),
##   ..   PZD00069.5 = col_character(),
##   ..   PZD00073.1 = col_character(),
##   ..   PZD00073.2 = col_character(),
##   ..   PZD00073.6 = col_character(),
##   ..   PZD00074.1 = col_character(),
##   ..   PZD00075.1 = col_character(),
##   ..   PZD00075.2 = col_character(),
##   ..   PZD00076.1 = col_character(),
##   ..   PZD00076.2 = col_character(),
##   ..   PZD00076.4 = col_character(),
##   ..   PZD00077.10 = col_character(),
##   ..   PZD00077.5 = col_character(),
##   ..   PZD00077.7 = col_character(),
##   ..   PZD00077.8 = col_character(),
##   ..   PZD00078.2 = col_character(),
##   ..   Ra2_ORF.1 = col_character(),
##   ..   Ra2_ORF.2 = col_character(),
##   ..   Ra2_ORF.4 = col_character(),
##   ..   Ra2_promoter.1 = col_character(),
##   ..   Ra2_promoter.2 = col_character(),
##   ..   Ra2_promoter.3 = col_character(),
##   ..   sh2.5 = col_character(),
##   ..   sh2.6 = col_character(),
##   ..   sh2.7 = col_character(),
##   ..   sh2.9 = col_character(),
##   ..   su1.4 = col_character(),
##   ..   su1.5 = col_character(),
##   ..   su1.7 = col_character(),
##   ..   tb1.17 = col_character(),
##   ..   tb1.18 = col_character(),
##   ..   tb1.19 = col_character(),
##   ..   tb1.5 = col_character(),
##   ..   te1.3 = col_character(),
##   ..   te1.4 = col_character(),
##   ..   zagl1.1 = col_character(),
##   ..   zagl1.6 = col_character(),
##   ..   zap1.2 = col_character(),
##   ..   zen1.1 = col_character(),
##   ..   zen1.2 = col_character(),
##   ..   zen1.4 = col_character(),
##   ..   zfl2.6 = col_character(),
##   ..   zmm3.4 = col_character()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

```
str(raw_snp_position)
```

```
## spc_tbl_ [983 × 15] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ SNP_ID              : chr [1:983] "abph1.20" "abph1.22" "ae1.3" "ae1.4" ...
##  $ cdv_marker_id       : num [1:983] 5976 5978 6605 6606 6607 ...
##  $ Chromosome          : chr [1:983] "2" "2" "5" "5" ...
##  $ Position            : chr [1:983] "27403404" "27403892" "167889790" "167889682" ...
##  $ alt_pos             : chr [1:983] NA NA NA NA ...
##  $ mult_positions      : chr [1:983] NA NA NA NA ...
##  $ amplicon            : chr [1:983] "abph1" "abph1" "ae1" "ae1" ...
##  $ cdv_map_feature.name: chr [1:983] "AB042260" "AB042260" "ae1" "ae1" ...
##  $ gene                : chr [1:983] "abph1" "abph1" "ae1" "ae1" ...
##  $ candidate/random    : chr [1:983] "candidate" "candidate" "candidate" "candidate" ...
##  $ Genaissance_daa_id  : num [1:983] 8393 8394 8395 8396 8397 ...
##  $ Sequenom_daa_id     : num [1:983] 10474 10475 10477 10478 10479 ...
##  $ count_amplicons     : num [1:983] 1 0 1 0 0 1 1 0 1 0 ...
##  $ count_cmf           : num [1:983] 1 0 1 0 0 1 0 0 1 0 ...
##  $ count_gene          : num [1:983] 1 0 1 0 0 1 1 0 1 0 ...
##  - attr(*, "spec")=
##   .. cols(
##   ..   SNP_ID = col_character(),
##   ..   cdv_marker_id = col_double(),
##   ..   Chromosome = col_character(),
##   ..   Position = col_character(),
##   ..   alt_pos = col_character(),
##   ..   mult_positions = col_character(),
##   ..   amplicon = col_character(),
##   ..   cdv_map_feature.name = col_character(),
##   ..   gene = col_character(),
##   ..   `candidate/random` = col_character(),
##   ..   Genaissance_daa_id = col_double(),
##   ..   Sequenom_daa_id = col_double(),
##   ..   count_amplicons = col_double(),
##   ..   count_cmf = col_double(),
##   ..   count_gene = col_double()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

Data Processing

Delete irrelevant columns

```
genotypes <- raw_genotypes %>% select(-JG_OTU)
snp_position <- raw_snp_position %>% select(SNP_ID, Chromosome, Position)
```

Separate groups maize and teosinte

```
genotypes_maize <- genotypes %>%
  filter(Group %in% c("ZMMIL", "ZMMLR", "ZMMMR"))

genotypes_teosinte <- genotypes %>%
  filter(Group %in% c("ZMPBA", "ZMPIL", "ZMPJA"))
```

Transpose data set and adjust rows and headers

```
genotypes_maize_t <- genotypes_maize %>%
  t() %>%
  as.data.frame() %>%
  slice(-2) %>%
  mutate(Sample_ID = rownames(.)) %>%
  relocate(Sample_ID) %>%
  { colnames(.) <- .[1, ]; . } %>%
  slice(-1) %>%
  mutate(across(everything(), as.character)) %>%
  mutate(across(-1, as.factor))

rownames(genotypes_maize_t) <- NULL

genotypes_teosinte_t <- genotypes_teosinte %>%
  t() %>%
  as.data.frame() %>%
  slice(-2) %>%
  mutate(Sample_ID = rownames(.)) %>%
  relocate(Sample_ID) %>%
  { colnames(.) <- .[1, ]; . } %>%
  slice(-1) %>%
  mutate(across(everything(), as.character)) %>%
  mutate(across(-1, as.factor))

rownames(genotypes_teosinte_t) <- NULL
```

Joint snp and genotypes

```
joint_snp_maize <- snp_position %>%
  bind_cols(genotypes_maize_t[, -1])

joint_snp_teosinte <- snp_position %>%
  bind_cols(genotypes_teosinte_t[, -1])
```

Generate 40 files

```
# maize increase file generate
maize_chromosome_increase <- function(chr) {
  joint_snp_maize %>%
    filter(Chromosome == chr) %>%
    mutate(Numeric_Position = suppressWarnings(as.numeric(Position))) %>%
    arrange(is.na(Numeric_Position), Numeric_Position) %>%
    select(-Numeric_Position) %>%
    write_tsv(paste0("maize_data/maize_chrom", chr, "_increase.txt"))
}

lapply(1:10, maize_chromosome_increase)
```

```
## [[1]]
## # A tibble: 155 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <fct>     <fct>     <fct>     <fct>     <fct>
##  1 PZB008… 1          157104   C/C       C/C       A/C       C/C       C/C
##  2 PZA029… 1          3205252  T/T       T/T       T/T       T/T       T/T
##  3 PZA003… 1          4175293  C/T       C/C       T/T       T/T       T/T
##  4 PZA003… 1          4175573  G/G       G/G       G/G       G/G       G/G
##  5 PZA028… 1          4429897  C/C       C/C       ?/?       ?/?       C/C
##  6 PZA028… 1          4430055  T/T       T/T       C/T       ?/?       T/T
##  7 PZD000… 1          4835472  G/G       G/G       G/G       G/G       G/G
##  8 PZD000… 1          4835540  T/T       ?/?       T/T       T/T       T/T
##  9 PZD000… 1          4835596  A/A       A/A       A/A       A/A       A/A
## 10 zagl1.6 1          4835658  T/T       T/T       T/T       T/T       T/T
## # ℹ 145 more rows
## # ℹ 1,568 more variables: ZDP_0552a <fct>, ZDP_0543a <fct>, ZDP_0042a <fct>,
## #   ZDP_0121a <fct>, ZDP_0101b <fct>, ZDP_0087a <fct>, ZDP_0033a <fct>,
## #   ZDP_0373b <fct>, ZDP_0168a <fct>, ZDP_0169a <fct>, ZDP_0517b <fct>,
## #   ZDP_1098a <fct>, ZDP_0591a <fct>, ZDP_0258a <fct>, ZDP_0009a <fct>,
## #   ZDP_0016a <fct>, ZDP_0010a <fct>, ZDP_0018a <fct>, ZDP_0647b <fct>,
## #   ZDP_1064b <fct>, ZDP_1053a <fct>, ZDP_1445b <fct>, ZDP_0163a <fct>, …
##
## [[2]]
## # A tibble: 127 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <fct>     <fct>     <fct>     <fct>     <fct>
##  1 PZA006… 2          1081781  G/G       G/G       G/G       G/G       A/G
##  2 PZA006… 2          1081791  A/C       C/C       C/C       C/C       C/C
##  3 PZA005… 2          2527199  G/G       G/G       G/G       C/G       C/C
##  4 PZA005… 2          2527379  G/G       G/G       G/T       G/T       G/G
##  5 PZD000… 2          3375254  C/G       C/G       C/G       ?/?       C/C
##  6 PZB012… 2          3375911  C/G       G/G       C/C       C/C       C/C
##  7 PZB012… 2          3376330  C/C       C/C       C/C       A/C       C/C
##  8 PZA006… 2          4180789  C/C       C/T       T/T       C/T       C/T
##  9 PZA003… 2          4679467  A/G       A/A       G/G       A/G       G/G
## 10 PZA002… 2          5821018  A/A       A/A       A/A       A/A       A/G
## # ℹ 117 more rows
## # ℹ 1,568 more variables: ZDP_0552a <fct>, ZDP_0543a <fct>, ZDP_0042a <fct>,
## #   ZDP_0121a <fct>, ZDP_0101b <fct>, ZDP_0087a <fct>, ZDP_0033a <fct>,
## #   ZDP_0373b <fct>, ZDP_0168a <fct>, ZDP_0169a <fct>, ZDP_0517b <fct>,
## #   ZDP_1098a <fct>, ZDP_0591a <fct>, ZDP_0258a <fct>, ZDP_0009a <fct>,
## #   ZDP_0016a <fct>, ZDP_0010a <fct>, ZDP_0018a <fct>, ZDP_0647b <fct>,
## #   ZDP_1064b <fct>, ZDP_1053a <fct>, ZDP_1445b <fct>, ZDP_0163a <fct>, …
##
## [[3]]
## # A tibble: 107 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <fct>     <fct>     <fct>     <fct>     <fct>
##  1 PZA003… 3          1240409  C/C       C/C       C/C       A/C       C/C
##  2 PZA006… 3          1882683  C/C       C/T       C/C       C/C       C/C
##  3 PZA006… 3          1882706  G/G       A/G       G/G       G/G       G/G
##  4 PZA006… 3          1882847  G/G       G/G       ?/?       ?/?       G/G
```

```
##  5 PZA001… 3          5479958  C/C       C/C       C/T       C/T       C/T
##  6 PZA001… 3          5479991  C/T       C/T       T/T       C/T       T/T
##  7 PZA001… 3          5480277  C/T       C/T       T/T       T/T       C/T
##  8 PZA001… 3          5480312  A/C       A/C       A/A       A/C       A/A
##  9 PZA004… 3          5480980  C/T       C/C       C/C       C/C       C/C
## 10 PZA032… 3          7602160  C/T       T/T       T/T       T/T       ?/?
## # ℹ 97 more rows
## # ℹ 1,568 more variables: ZDP_0552a <fct>, ZDP_0543a <fct>, ZDP_0042a <fct>,
## #   ZDP_0121a <fct>, ZDP_0101b <fct>, ZDP_0087a <fct>, ZDP_0033a <fct>,
## #   ZDP_0373b <fct>, ZDP_0168a <fct>, ZDP_0169a <fct>, ZDP_0517b <fct>,
## #   ZDP_1098a <fct>, ZDP_0591a <fct>, ZDP_0258a <fct>, ZDP_0009a <fct>,
## #   ZDP_0016a <fct>, ZDP_0010a <fct>, ZDP_0018a <fct>, ZDP_0647b <fct>,
## #   ZDP_1064b <fct>, ZDP_1053a <fct>, ZDP_1445b <fct>, ZDP_0163a <fct>, …
##
## [[4]]
## # A tibble: 91 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <fct>     <fct>     <fct>     <fct>     <fct>
##  1 PZA030… 4          139753   G/G       A/G       G/G       G/G       G/G
##  2 PZA030… 4          139810   C/C       ?/?       ?/?       ?/?       C/C
##  3 PZA002… 4          2589400  A/A       A/A       A/A       A/G       A/A
##  4 PZA029… 4          2848110  C/C       C/C       C/C       C/C       C/C
##  5 PZA007… 4          2852810  G/G       G/G       ?/?       G/G       A/G
##  6 PZA004… 4          6401332  T/T       T/T       T/T       C/T       C/T
##  7 PZA000… 4          18502179 C/C       A/C       C/C       C/C       A/A
##  8 PZA012… 4          19461323 A/A       G/G       G/G       G/G       G/G
##  9 PZA012… 4          19461458 G/G       G/G       G/G       G/G       G/G
## 10 PZA001… 4          20065054 A/A       A/A       A/A       A/A       A/A
## # ℹ 81 more rows
## # ℹ 1,568 more variables: ZDP_0552a <fct>, ZDP_0543a <fct>, ZDP_0042a <fct>,
## #   ZDP_0121a <fct>, ZDP_0101b <fct>, ZDP_0087a <fct>, ZDP_0033a <fct>,
## #   ZDP_0373b <fct>, ZDP_0168a <fct>, ZDP_0169a <fct>, ZDP_0517b <fct>,
## #   ZDP_1098a <fct>, ZDP_0591a <fct>, ZDP_0258a <fct>, ZDP_0009a <fct>,
## #   ZDP_0016a <fct>, ZDP_0010a <fct>, ZDP_0018a <fct>, ZDP_0647b <fct>,
## #   ZDP_1064b <fct>, ZDP_1053a <fct>, ZDP_1445b <fct>, ZDP_0163a <fct>, …
##
## [[5]]
## # A tibble: 122 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <fct>     <fct>     <fct>     <fct>     <fct>
##  1 PZA006… 5          920922   C/C       C/C       ?/?       ?/?       C/C
##  2 PZA028… 5          945545   C/C       C/C       G/G       C/G       G/G
##  3 PZA001… 5          2123914  ?/?       T/T       C/T       T/T       C/C
##  4 PZA030… 5          3194670  ?/?       C/C       C/C       C/C       C/T
##  5 PZA030… 5          3194673  T/T       T/T       ?/?       ?/?       C/T
##  6 PZA030… 5          6784084  C/C       C/G       C/G       C/G       C/C
##  7 PZB000… 5          6960673  A/A       G/G       A/G       A/G       G/G
##  8 PZB000… 5          6960836  T/T       C/T       C/C       C/C       T/T
##  9 PZB000… 5          6960889  C/C       C/T       C/C       C/C       T/T
## 10 PZB000… 5          6961472  A/A       A/A       A/T       A/T       A/A
## # ℹ 112 more rows
## # ℹ 1,568 more variables: ZDP_0552a <fct>, ZDP_0543a <fct>, ZDP_0042a <fct>,
```

```
## #   ZDP_0121a <fct>, ZDP_0101b <fct>, ZDP_0087a <fct>, ZDP_0033a <fct>,
## #   ZDP_0373b <fct>, ZDP_0168a <fct>, ZDP_0169a <fct>, ZDP_0517b <fct>,
## #   ZDP_1098a <fct>, ZDP_0591a <fct>, ZDP_0258a <fct>, ZDP_0009a <fct>,
## #   ZDP_0016a <fct>, ZDP_0010a <fct>, ZDP_0018a <fct>, ZDP_0647b <fct>,
## #   ZDP_1064b <fct>, ZDP_1053a <fct>, ZDP_1445b <fct>, ZDP_0163a <fct>, …
##
## [[6]]
## # A tibble: 76 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <fct>     <fct>     <fct>     <fct>     <fct>
##  1 PZA032… 6          9243581  G/G       A/A       A/A       A/A       A/A
##  2 PZA032… 6          9243719  T/T       C/C       C/C       C/C       C/C
##  3 PZA005… 6          12075777 C/C       C/C       C/C       C/T       C/C
##  4 PZA030… 6          20837292 G/G       G/G       G/G       G/G       A/G
##  5 PZA030… 6          20837423 T/T       C/C       C/C       C/C       C/T
##  6 PZA030… 6          20837504 G/G       G/G       A/A       G/G       G/G
##  7 PZA004… 6          22404308 G/G       C/G       C/C       C/C       G/G
##  8 PZA005… 6          23874393 T/T       G/G       G/G       G/G       G/G
##  9 PZA005… 6          23874474 ?/?       C/T       T/T       C/C       C/C
## 10 PZA030… 6          25335322 T/T       T/T       T/T       T/T       A/A
## # ℹ 66 more rows
## # ℹ 1,568 more variables: ZDP_0552a <fct>, ZDP_0543a <fct>, ZDP_0042a <fct>,
## #   ZDP_0121a <fct>, ZDP_0101b <fct>, ZDP_0087a <fct>, ZDP_0033a <fct>,
## #   ZDP_0373b <fct>, ZDP_0168a <fct>, ZDP_0169a <fct>, ZDP_0517b <fct>,
## #   ZDP_1098a <fct>, ZDP_0591a <fct>, ZDP_0258a <fct>, ZDP_0009a <fct>,
## #   ZDP_0016a <fct>, ZDP_0010a <fct>, ZDP_0018a <fct>, ZDP_0647b <fct>,
## #   ZDP_1064b <fct>, ZDP_1053a <fct>, ZDP_1445b <fct>, ZDP_0163a <fct>, …
##
## [[7]]
## # A tibble: 97 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <fct>     <fct>     <fct>     <fct>     <fct>
##  1 PZB001… 7          2081333  C/C       T/T       C/T       C/T       T/T
##  2 PZA001… 7          2998280  A/A       A/A       ?/?       ?/?       A/T
##  3 PZA001… 7          2998619  A/A       A/A       A/A       ?/?       A/A
##  4 PZA029… 7          4244157  G/G       A/A       ?/?       ?/?       A/A
##  5 PZA028… 7          13058709 A/G       G/G       G/G       G/G       A/A
##  6 PZA028… 7          13058813 A/G       ?/?       A/A       A/A       A/A
##  7 PZA030… 7          14354166 C/T       C/C       C/T       C/T       C/T
##  8 PZA030… 7          14354237 A/G       G/G       A/G       A/G       G/G
##  9 PZB012… 7          15335449 A/A       A/A       A/G       A/G       A/G
## 10 PZB012… 7          15335531 G/G       A/A       A/A       A/G       A/A
## # ℹ 87 more rows
## # ℹ 1,568 more variables: ZDP_0552a <fct>, ZDP_0543a <fct>, ZDP_0042a <fct>,
## #   ZDP_0121a <fct>, ZDP_0101b <fct>, ZDP_0087a <fct>, ZDP_0033a <fct>,
## #   ZDP_0373b <fct>, ZDP_0168a <fct>, ZDP_0169a <fct>, ZDP_0517b <fct>,
## #   ZDP_1098a <fct>, ZDP_0591a <fct>, ZDP_0258a <fct>, ZDP_0009a <fct>,
## #   ZDP_0016a <fct>, ZDP_0010a <fct>, ZDP_0018a <fct>, ZDP_0647b <fct>,
## #   ZDP_1064b <fct>, ZDP_1053a <fct>, ZDP_1445b <fct>, ZDP_0163a <fct>, …
##
## [[8]]
## # A tibble: 62 × 1,576
```

```
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <fct>     <fct>     <fct>     <fct>     <fct>
##  1 PZA033… 8          4766694  C/T       T/T       C/T       ?/?       C/T
##  2 PZA000… 8          5966698  C/C       C/C       C/C       C/C       C/C
##  3 PZA000… 8          5966738  G/T       T/T       G/G       G/G       G/G
##  4 PZA029… 8          14721554 C/C       A/A       A/A       C/C       A/A
##  5 PZB000… 8          21135142 C/C       C/C       C/C       C/C       C/C
##  6 PZB000… 8          21135254 C/C       C/C       C/C       C/C       C/C
##  7 PZB000… 8          21136819 G/G       G/G       G/G       G/G       G/G
##  8 PZB000… 8          21137213 G/G       G/G       A/G       A/G       A/A
##  9 PZB000… 8          21137286 T/T       T/T       C/T       ?/?       C/C
## 10 PZD000… 8          22274672 C/C       C/C       C/C       C/C       C/C
## # ℹ 52 more rows
## # ℹ 1,568 more variables: ZDP_0552a <fct>, ZDP_0543a <fct>, ZDP_0042a <fct>,
## #   ZDP_0121a <fct>, ZDP_0101b <fct>, ZDP_0087a <fct>, ZDP_0033a <fct>,
## #   ZDP_0373b <fct>, ZDP_0168a <fct>, ZDP_0169a <fct>, ZDP_0517b <fct>,
## #   ZDP_1098a <fct>, ZDP_0591a <fct>, ZDP_0258a <fct>, ZDP_0009a <fct>,
## #   ZDP_0016a <fct>, ZDP_0010a <fct>, ZDP_0018a <fct>, ZDP_0647b <fct>,
## #   ZDP_1064b <fct>, ZDP_1053a <fct>, ZDP_1445b <fct>, ZDP_0163a <fct>, …
##
## [[9]]
## # A tibble: 60 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <fct>     <fct>     <fct>     <fct>     <fct>
##  1 PZA030… 9          3873116  C/T       C/C       C/C       C/T       C/C
##  2 PZA030… 9          3873152  C/T       C/C       C/C       C/C       C/C
##  3 PZA004… 9          11972467 G/G       G/G       G/G       ?/?       G/G
##  4 PZD000… 9          16960436 G/G       A/G       G/G       G/G       G/G
##  5 PZD000… 9          16960510 T/T       G/T       G/T       G/G       G/T
##  6 PZD000… 9          16960633 G/G       A/G       G/G       G/G       G/G
##  7 zmm3.4  9          16966348 C/C       C/C       C/C       C/C       C/C
##  8 PZA028… 9          17653885 C/C       C/C       C/C       C/C       C/C
##  9 PZA017… 9          19252371 C/C       C/C       C/C       C/C       C/C
## 10 PZA017… 9          19252475 C/C       C/C       C/C       C/C       C/C
## # ℹ 50 more rows
## # ℹ 1,568 more variables: ZDP_0552a <fct>, ZDP_0543a <fct>, ZDP_0042a <fct>,
## #   ZDP_0121a <fct>, ZDP_0101b <fct>, ZDP_0087a <fct>, ZDP_0033a <fct>,
## #   ZDP_0373b <fct>, ZDP_0168a <fct>, ZDP_0169a <fct>, ZDP_0517b <fct>,
## #   ZDP_1098a <fct>, ZDP_0591a <fct>, ZDP_0258a <fct>, ZDP_0009a <fct>,
## #   ZDP_0016a <fct>, ZDP_0010a <fct>, ZDP_0018a <fct>, ZDP_0647b <fct>,
## #   ZDP_1064b <fct>, ZDP_1053a <fct>, ZDP_1445b <fct>, ZDP_0163a <fct>, …
##
## [[10]]
## # A tibble: 53 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <fct>     <fct>     <fct>     <fct>     <fct>
##  1 PZA030… 10         6121326  T/T       T/T       T/T       T/T       T/T
##  2 PZA002… 10         10432605 A/G       A/A       A/A       G/G       A/A
##  3 PZA002… 10         10432680 T/T       T/T       T/T       C/T       T/T
##  4 PZA004… 10         13546326 T/T       C/T       C/C       C/T       C/C
##  5 PZA029… 10         16264730 C/C       C/C       A/C       ?/?       C/C
##  6 PZA006… 10         17645859 A/G       G/G       G/G       G/G       G/G
```

```
##  7 PZA006… 10         17645938 G/G       G/G       G/G       G/G       G/G
##  8 PZA006… 10         17646056 T/T       C/T       C/C       C/T       C/T
##  9 PZA006… 10         17646117 A/A       A/A       A/A       A/A       A/A
## 10 PZA000… 10         18903460 C/C       C/C       C/C       C/C       C/C
## # ℹ 43 more rows
## # ℹ 1,568 more variables: ZDP_0552a <fct>, ZDP_0543a <fct>, ZDP_0042a <fct>,
## #   ZDP_0121a <fct>, ZDP_0101b <fct>, ZDP_0087a <fct>, ZDP_0033a <fct>,
## #   ZDP_0373b <fct>, ZDP_0168a <fct>, ZDP_0169a <fct>, ZDP_0517b <fct>,
## #   ZDP_1098a <fct>, ZDP_0591a <fct>, ZDP_0258a <fct>, ZDP_0009a <fct>,
## #   ZDP_0016a <fct>, ZDP_0010a <fct>, ZDP_0018a <fct>, ZDP_0647b <fct>,
## #   ZDP_1064b <fct>, ZDP_1053a <fct>, ZDP_1445b <fct>, ZDP_0163a <fct>, …
```

```r
# maize decrease file generate
maize_chromosome_decrease <- function(chr) {
  joint_snp_maize %>%
    filter(Chromosome == chr) %>%
    mutate(Numeric_Position = suppressWarnings(as.numeric(Position))) %>%
    arrange(is.na(Numeric_Position), desc(Numeric_Position)) %>%
    select(-Numeric_Position) %>%
    mutate(across(4:ncol(.), ~ ifelse(. == "?/?", "-/-", .))) %>%
    mutate(across(4:ncol(.), as.character)) %>%
    write_tsv(paste0("maize_data/maize_chrom", chr, "_decrease.txt"))
}

lapply(1:10, maize_chromosome_decrease)
```

```
## [[1]]
## # A tibble: 155 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <chr>     <chr>     <chr>     <chr>     <chr>
##  1 PZA002… 1          2984129… 9         9         -/-       6         6
##  2 PZA004… 1          2980826… 9         9         9         4         4
##  3 PZA004… 1          2980825… 6         6         6         6         6
##  4 PZA004… 1          2980825… 9         9         9         9         9
##  5 PZA004… 1          2980824… 11        11        11        11        11
##  6 PZA002… 1          2957711… 2         2         2         2         2
##  7 PZA000… 1          2954595… 2         2         4         2         2
##  8 PZA006… 1          2936327… 7         -/-       9         9         9
##  9 PZA002… 1          2927289… 9         9         9         4         4
## 10 PZA002… 1          2927288… 10        9         9         10        11
## # ℹ 145 more rows
## # ℹ 1,568 more variables: ZDP_0552a <chr>, ZDP_0543a <chr>, ZDP_0042a <chr>,
## #   ZDP_0121a <chr>, ZDP_0101b <chr>, ZDP_0087a <chr>, ZDP_0033a <chr>,
## #   ZDP_0373b <chr>, ZDP_0168a <chr>, ZDP_0169a <chr>, ZDP_0517b <chr>,
## #   ZDP_1098a <chr>, ZDP_0591a <chr>, ZDP_0258a <chr>, ZDP_0009a <chr>,
## #   ZDP_0016a <chr>, ZDP_0010a <chr>, ZDP_0018a <chr>, ZDP_0647b <chr>,
## #   ZDP_1064b <chr>, ZDP_1053a <chr>, ZDP_1445b <chr>, ZDP_0163a <chr>, …
##
## [[2]]
## # A tibble: 127 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <chr>     <chr>     <chr>     <chr>     <chr>
##  1 PZD000… 2          2331291… 6         6         6         6         8
##  2 PZD000… 2          2331290… 11        11        6         6         6
##  3 zap1.2  2          2331285… 4         2         2         4         9
##  4 PZD000… 2          2331285… 6         6         11        8         6
##  5 PZA006… 2          2289910… 9         9         9         9         9
##  6 PZA006… 2          2289910… 4         2         2         4         2
##  7 PZA004… 2          2221352… 9         9         9         4         2
##  8 PZA004… 2          2221349… 4         9         4         2         2
##  9 PZA004… 2          2221349… 6         9         6         7         9
## 10 PZA001… 2          2202647… 6         6         3         3         6
## # ℹ 117 more rows
## # ℹ 1,568 more variables: ZDP_0552a <chr>, ZDP_0543a <chr>, ZDP_0042a <chr>,
## #   ZDP_0121a <chr>, ZDP_0101b <chr>, ZDP_0087a <chr>, ZDP_0033a <chr>,
## #   ZDP_0373b <chr>, ZDP_0168a <chr>, ZDP_0169a <chr>, ZDP_0517b <chr>,
## #   ZDP_1098a <chr>, ZDP_0591a <chr>, ZDP_0258a <chr>, ZDP_0009a <chr>,
## #   ZDP_0016a <chr>, ZDP_0010a <chr>, ZDP_0018a <chr>, ZDP_0647b <chr>,
## #   ZDP_1064b <chr>, ZDP_1053a <chr>, ZDP_1445b <chr>, ZDP_0163a <chr>, …
##
## [[3]]
## # A tibble: 107 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <chr>     <chr>     <chr>     <chr>     <chr>
##  1 PZA000… 3          2286142… 9         6         9         9         9
##  2 PZA004… 3          2268722… 8         6         6         6         6
##  3 PZA027… 3          2221146… 3         6         6         6         2
##  4 PZA027… 3          2221145… 10        9         9         9         9
```

```
##  5 PZA002… 3          2217306… 6         2         3         6         2
##  6 PZA002… 3          2193090… 8         8         6         11        6
##  7 PZA028… 3          2173296… 9         4         9         4         9
##  8 PZA028… 3          2173296… 6         6         6         -/-       6
##  9 sh2.9   3          2148411… 2         2         2         2         2
## 10 sh2.5   3          2148410… 11        11        11        11        8
## # ℹ 97 more rows
## # ℹ 1,568 more variables: ZDP_0552a <chr>, ZDP_0543a <chr>, ZDP_0042a <chr>,
## #   ZDP_0121a <chr>, ZDP_0101b <chr>, ZDP_0087a <chr>, ZDP_0033a <chr>,
## #   ZDP_0373b <chr>, ZDP_0168a <chr>, ZDP_0169a <chr>, ZDP_0517b <chr>,
## #   ZDP_1098a <chr>, ZDP_0591a <chr>, ZDP_0258a <chr>, ZDP_0009a <chr>,
## #   ZDP_0016a <chr>, ZDP_0010a <chr>, ZDP_0018a <chr>, ZDP_0647b <chr>,
## #   ZDP_1064b <chr>, ZDP_1053a <chr>, ZDP_1445b <chr>, ZDP_0163a <chr>, …
##
## [[4]]
## # A tibble: 91 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <chr>     <chr>     <chr>     <chr>     <chr>
##  1 PZA000… 4          2450516… 2         2         9         4         4
##  2 PZA006… 4          2449928… 9         6         -/-       -/-       7
##  3 PZA030… 4          2432228… 8         6         6         6         6
##  4 PZA030… 4          2432228… 8         11        6         11        8
##  5 PZA000… 4          2430034… 6         -/-       11        8         6
##  6 PZA000… 4          2430033… 11        10        9         10        11
##  7 PZA005… 4          2407691… 9         -/-       2         4         9
##  8 PZA006… 4          2268846… 9         9         4         4         4
##  9 PZA028… 4          2208578… 4         9         9         9         9
## 10 PZA030… 4          2153934… 7         6         9         9         9
## # ℹ 81 more rows
## # ℹ 1,568 more variables: ZDP_0552a <chr>, ZDP_0543a <chr>, ZDP_0042a <chr>,
## #   ZDP_0121a <chr>, ZDP_0101b <chr>, ZDP_0087a <chr>, ZDP_0033a <chr>,
## #   ZDP_0373b <chr>, ZDP_0168a <chr>, ZDP_0169a <chr>, ZDP_0517b <chr>,
## #   ZDP_1098a <chr>, ZDP_0591a <chr>, ZDP_0258a <chr>, ZDP_0009a <chr>,
## #   ZDP_0016a <chr>, ZDP_0010a <chr>, ZDP_0018a <chr>, ZDP_0647b <chr>,
## #   ZDP_1064b <chr>, ZDP_1053a <chr>, ZDP_1445b <chr>, ZDP_0163a <chr>, …
##
## [[5]]
## # A tibble: 122 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <chr>     <chr>     <chr>     <chr>     <chr>
##  1 PZA004… 5          2144428… 5         2         5         5         5
##  2 PZA000… 5          2144249… 6         6         6         6         6
##  3 PZA004… 5          2132715… 9         9         -/-       -/-       2
##  4 PZA033… 5          2108908… 11        11        11        11        11
##  5 PZA033… 5          2108907… 8         11        11        11        11
##  6 PZA005… 5          2066650… 9         9         9         9         9
##  7 PZA005… 5          2066649… 6         6         6         6         6
##  8 PZA005… 5          2066647… 6         6         6         8         6
##  9 PZA003… 5          2028472… 8         8         -/-       8         8
## 10 PZA003… 5          2028471… 3         3         2         3         3
## # ℹ 112 more rows
## # ℹ 1,568 more variables: ZDP_0552a <chr>, ZDP_0543a <chr>, ZDP_0042a <chr>,
```

```
## #   ZDP_0121a <chr>, ZDP_0101b <chr>, ZDP_0087a <chr>, ZDP_0033a <chr>,
## #   ZDP_0373b <chr>, ZDP_0168a <chr>, ZDP_0169a <chr>, ZDP_0517b <chr>,
## #   ZDP_1098a <chr>, ZDP_0591a <chr>, ZDP_0258a <chr>, ZDP_0009a <chr>,
## #   ZDP_0016a <chr>, ZDP_0010a <chr>, ZDP_0018a <chr>, ZDP_0647b <chr>,
## #   ZDP_1064b <chr>, ZDP_1053a <chr>, ZDP_1445b <chr>, ZDP_0163a <chr>, …
##
## [[6]]
## # A tibble: 76 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <chr>     <chr>     <chr>     <chr>     <chr>
##  1 PZA030… 6          1657249… 7         -/-       9         7         9
##  2 PZA001… 6          1644703… 6         6         6         6         8
##  3 PZB012… 6          1644194… 7         7         6         6         6
##  4 PZB012… 6          1644193… 6         8         6         6         6
##  5 PZA002… 6          1615701… 8         6         8         6         8
##  6 PZA033… 6          1598081… 2         2         2         2         2
##  7 PZA000… 6          1571293… 11        5         5         11        11
##  8 PZA000… 6          1571291… 8         -/-       8         6         6
##  9 PZA029… 6          1519934… 8         6         6         6         8
## 10 PZA002… 6          1504706… 11        11        11        11        11
## # ℹ 66 more rows
## # ℹ 1,568 more variables: ZDP_0552a <chr>, ZDP_0543a <chr>, ZDP_0042a <chr>,
## #   ZDP_0121a <chr>, ZDP_0101b <chr>, ZDP_0087a <chr>, ZDP_0033a <chr>,
## #   ZDP_0373b <chr>, ZDP_0168a <chr>, ZDP_0169a <chr>, ZDP_0517b <chr>,
## #   ZDP_1098a <chr>, ZDP_0591a <chr>, ZDP_0258a <chr>, ZDP_0009a <chr>,
## #   ZDP_0016a <chr>, ZDP_0010a <chr>, ZDP_0018a <chr>, ZDP_0647b <chr>,
## #   ZDP_1064b <chr>, ZDP_1053a <chr>, ZDP_1445b <chr>, ZDP_0163a <chr>, …
##
## [[7]]
## # A tibble: 97 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <chr>     <chr>     <chr>     <chr>     <chr>
##  1 PZA028… 7          1690679… 2         2         2         2         2
##  2 PZA028… 7          1690677… 9         9         7         7         9
##  3 PZA004… 7          1690175… 11        11        8         8         8
##  4 PZA000… 7          1672182… 6         11        11        11        11
##  5 PZA004… 7          1670607… 6         6         6         6         6
##  6 PZA004… 7          1670606… 6         6         6         6         6
##  7 PZA032… 7          1670148… 6         6         6         6         6
##  8 PZB006… 7          1659948… 6         6         7         7         6
##  9 PZB006… 7          1659948… 6         6         7         6         6
## 10 PZB006… 7          1659948… 4         2         2         2         2
## # ℹ 87 more rows
## # ℹ 1,568 more variables: ZDP_0552a <chr>, ZDP_0543a <chr>, ZDP_0042a <chr>,
## #   ZDP_0121a <chr>, ZDP_0101b <chr>, ZDP_0087a <chr>, ZDP_0033a <chr>,
## #   ZDP_0373b <chr>, ZDP_0168a <chr>, ZDP_0169a <chr>, ZDP_0517b <chr>,
## #   ZDP_1098a <chr>, ZDP_0591a <chr>, ZDP_0258a <chr>, ZDP_0009a <chr>,
## #   ZDP_0016a <chr>, ZDP_0010a <chr>, ZDP_0018a <chr>, ZDP_0647b <chr>,
## #   ZDP_1064b <chr>, ZDP_1053a <chr>, ZDP_1445b <chr>, ZDP_0163a <chr>, …
##
## [[8]]
## # A tibble: 62 × 1,576
```

```
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <chr>     <chr>     <chr>     <chr>     <chr>
##  1 PZA002… 8          1669800… 2         4         2         2         2
##  2 PZA002… 8          1669799… 11        11        8         8         11
##  3 PZA006… 8          1661827… 2         2         2         2         2
##  4 PZA006… 8          1661827… 9         9         9         9         9
##  5 PZA005… 8          1635723… 2         9         -/-       -/-       2
##  6 PZA004… 8          1635654… 6         6         11        6         6
##  7 PZA004… 8          1635654… 6         8         6         8         8
##  8 PZA004… 8          1635653… 6         6         -/-       6         8
##  9 PZA007… 8          1598979… 8         11        11        11        11
## 10 PZB001… 8          1574730… 8         11        11        11        11
## # ℹ 52 more rows
## # ℹ 1,568 more variables: ZDP_0552a <chr>, ZDP_0543a <chr>, ZDP_0042a <chr>,
## #   ZDP_0121a <chr>, ZDP_0101b <chr>, ZDP_0087a <chr>, ZDP_0033a <chr>,
## #   ZDP_0373b <chr>, ZDP_0168a <chr>, ZDP_0169a <chr>, ZDP_0517b <chr>,
## #   ZDP_1098a <chr>, ZDP_0591a <chr>, ZDP_0258a <chr>, ZDP_0009a <chr>,
## #   ZDP_0016a <chr>, ZDP_0010a <chr>, ZDP_0018a <chr>, ZDP_0647b <chr>,
## #   ZDP_1064b <chr>, ZDP_1053a <chr>, ZDP_1445b <chr>, ZDP_0163a <chr>, …
##
## [[9]]
## # A tibble: 60 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <chr>     <chr>     <chr>     <chr>     <chr>
##  1 PZA001… 9          1512899… 7         9         -/-       -/-       -/-
##  2 PZA001… 9          1512898… 9         9         9         4         9
##  3 PZA001… 9          1493987… 6         6         6         6         6
##  4 PZA005… 9          1473849… 9         9         9         9         9
##  5 PZA029… 9          1464706… 6         6         6         6         6
##  6 PZA029… 9          1464705… 2         2         2         2         2
##  7 PZA003… 9          1425000… 6         6         6         6         6
##  8 PZA003… 9          1424998… 11        11        8         8         11
##  9 PZB002… 9          1422714… 11        11        11        11        11
## 10 PZB002… 9          1422710… 2         2         2         4         2
## # ℹ 50 more rows
## # ℹ 1,568 more variables: ZDP_0552a <chr>, ZDP_0543a <chr>, ZDP_0042a <chr>,
## #   ZDP_0121a <chr>, ZDP_0101b <chr>, ZDP_0087a <chr>, ZDP_0033a <chr>,
## #   ZDP_0373b <chr>, ZDP_0168a <chr>, ZDP_0169a <chr>, ZDP_0517b <chr>,
## #   ZDP_1098a <chr>, ZDP_0591a <chr>, ZDP_0258a <chr>, ZDP_0009a <chr>,
## #   ZDP_0016a <chr>, ZDP_0010a <chr>, ZDP_0018a <chr>, ZDP_0647b <chr>,
## #   ZDP_1064b <chr>, ZDP_1053a <chr>, ZDP_1445b <chr>, ZDP_0163a <chr>, …
##
## [[10]]
## # A tibble: 53 × 1,576
##    SNP_ID  Chromosome Position ZDP_0752a ZDP_0793a ZDP_0612a ZDP_0602a ZDP_0581a
##    <chr>   <chr>      <chr>    <chr>     <chr>     <chr>     <chr>     <chr>
##  1 PZD000… 10         1465994… 6         6         6         6         6
##  2 PZD000… 10         1465991… 7         6         7         7         9
##  3 PZA029… 10         1436761… 6         -/-       6         6         6
##  4 PZA007… 10         1428547… 2         4         9         2         -/-
##  5 PZA007… 10         1428546… 6         8         6         6         6
##  6 PZB000… 10         1404931… 6         6         7         6         7
```

```
##  7 PZB011… 10         1340348… 11        11        11        11        11
##  8 PZB011… 10         1340346… 9         4         9         9         9
##  9 PZB011… 10         1340344… 2         6         6         6         6
## 10 PZA003… 10         1301707… 9         9         9         9         9
## # ℹ 43 more rows
## # ℹ 1,568 more variables: ZDP_0552a <chr>, ZDP_0543a <chr>, ZDP_0042a <chr>,
## #   ZDP_0121a <chr>, ZDP_0101b <chr>, ZDP_0087a <chr>, ZDP_0033a <chr>,
## #   ZDP_0373b <chr>, ZDP_0168a <chr>, ZDP_0169a <chr>, ZDP_0517b <chr>,
## #   ZDP_1098a <chr>, ZDP_0591a <chr>, ZDP_0258a <chr>, ZDP_0009a <chr>,
## #   ZDP_0016a <chr>, ZDP_0010a <chr>, ZDP_0018a <chr>, ZDP_0647b <chr>,
## #   ZDP_1064b <chr>, ZDP_1053a <chr>, ZDP_1445b <chr>, ZDP_0163a <chr>, …
```

```
# teosinte increase file generate
teosinte_chromosome_increase <- function(chr) {
  joint_snp_teosinte %>%
    filter(Chromosome == chr) %>%
    mutate(Numeric_Position = suppressWarnings(as.numeric(Position))) %>%
    arrange(is.na(Numeric_Position), Numeric_Position) %>%
    select(-Numeric_Position) %>%
    write_tsv(paste0("teosinte_data/teosinte_chrom", chr, "_increase.txt"))
}

lapply(1:10, teosinte_chromosome_increase)
```

```
## [[1]]
## # A tibble: 155 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct>
##  1 PZB00859… 1          157104   A/A   A/C   A/A   C/C   A/C   C/C   C/C   A/C
##  2 PZA02962… 1          3205252  T/T   A/A   T/T   A/T   A/T   T/T   T/T   A/T
##  3 PZA00393… 1          4175293  T/T   T/T   T/T   C/T   C/T   C/T   C/T   C/C
##  4 PZA00393… 1          4175573  G/G   A/G   G/G   G/G   G/G   G/G   G/G   G/G
##  5 PZA02869… 1          4429897  C/C   C/C   C/C   C/C   C/C   C/C   C/C   C/C
##  6 PZA02869… 1          4430055  T/T   T/T   T/T   T/T   T/T   T/T   C/C   C/T
##  7 PZD00021… 1          4835472  G/G   G/G   G/G   G/G   G/G   G/G   ?/?   G/G
##  8 PZD00021… 1          4835540  C/C   C/C   T/T   C/C   C/T   T/T   C/C   ?/?
##  9 PZD00021… 1          4835596  A/T   A/A   A/A   A/A   A/T   A/A   A/A   A/T
## 10 zagl1.6   1          4835658  C/T   C/T   T/T   T/T   C/T   C/T   C/C   C/C
## # ℹ 145 more rows
## # ℹ 967 more variables: TA057 <fct>, TA137 <fct>, TA153 <fct>, TA205 <fct>,
## #   TA217 <fct>, TA242 <fct>, TA285 <fct>, TA293 <fct>, TAMex0091.1 <fct>,
## #   TAMex0153.1 <fct>, TAMex0238.1 <fct>, TAMex0396.1 <fct>, TAMex0446.1 <fct>,
## #   TAMex0455.2 <fct>, TAMex0520.1 <fct>, TAMex0608.5 <fct>, TAMex0624.1 <fct>,
## #   TAMex0643.1 <fct>, TAMex0775.2 <fct>, TAMex0930.1 <fct>, TAMex1083.1 <fct>,
## #   TAMex1111.1 <fct>, TAMex1195.2 <fct>, TAMex1222.2 <fct>, …
##
## [[2]]
## # A tibble: 127 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct>
##  1 PZA00680… 2          1081781  A/G   G/G   A/A   G/G   G/G   A/G   G/G   A/A
##  2 PZA00680… 2          1081791  C/C   C/C   C/C   C/C   C/C   C/C   A/C   C/C
##  3 PZA00525… 2          2527199  G/G   G/G   G/G   C/G   C/G   C/G   G/G   C/G
##  4 PZA00525… 2          2527379  T/T   G/T   T/T   G/T   G/T   G/G   G/T   G/G
##  5 PZD00074… 2          3375254  C/C   C/G   C/C   C/G   C/C   C/G   C/C   G/G
##  6 PZB01233… 2          3375911  C/G   G/G   G/G   C/C   C/G   C/G   G/G   C/G
##  7 PZB01233… 2          3376330  C/C   C/C   C/C   C/C   C/C   C/C   C/C   C/C
##  8 PZA00613… 2          4180789  C/T   T/T   C/C   C/T   C/T   C/T   T/T   C/T
##  9 PZA00396… 2          4679467  ?/?   A/G   ?/?   A/G   ?/?   ?/?   ?/?   ?/?
## 10 PZA00200… 2          5821018  A/A   A/A   A/A   A/G   G/G   A/G   A/G   A/G
## # ℹ 117 more rows
## # ℹ 967 more variables: TA057 <fct>, TA137 <fct>, TA153 <fct>, TA205 <fct>,
## #   TA217 <fct>, TA242 <fct>, TA285 <fct>, TA293 <fct>, TAMex0091.1 <fct>,
## #   TAMex0153.1 <fct>, TAMex0238.1 <fct>, TAMex0396.1 <fct>, TAMex0446.1 <fct>,
## #   TAMex0455.2 <fct>, TAMex0520.1 <fct>, TAMex0608.5 <fct>, TAMex0624.1 <fct>,
## #   TAMex0643.1 <fct>, TAMex0775.2 <fct>, TAMex0930.1 <fct>, TAMex1083.1 <fct>,
## #   TAMex1111.1 <fct>, TAMex1195.2 <fct>, TAMex1222.2 <fct>, …
##
## [[3]]
## # A tibble: 107 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct>
##  1 PZA00309… 3          1240409  C/C   C/C   C/C   C/C   C/C   C/C   C/C   C/C
##  2 PZA00615… 3          1882683  ?/?   ?/?   ?/?   C/C   C/C   C/T   T/T   T/T
##  3 PZA00615… 3          1882706  G/G   G/G   G/G   G/G   G/G   A/G   A/A   A/A
##  4 PZA00615… 3          1882847  A/G   G/G   G/G   G/G   G/G   G/G   G/G   G/G
```

```
##  5 PZA00100… 3          5479958  C/T   C/C   C/C   C/C   C/C   C/T   C/C   C/T
##  6 PZA00100… 3          5479991  C/T   C/T   T/T   T/T   C/C   ?/?   T/T   T/T
##  7 PZA00100… 3          5480277  T/T   T/T   T/T   C/C   T/T   T/T   T/T   T/T
##  8 PZA00100… 3          5480312  A/C   A/C   A/A   A/A   C/C   A/C   A/A   A/A
##  9 PZA00458… 3          5480980  C/C   C/T   C/C   T/T   C/C   C/C   C/C   C/C
## 10 PZA03283… 3          7602160  T/T   C/T   C/C   T/T   T/T   T/T   T/T   C/T
## # ℹ 97 more rows
## # ℹ 967 more variables: TA057 <fct>, TA137 <fct>, TA153 <fct>, TA205 <fct>,
## #   TA217 <fct>, TA242 <fct>, TA285 <fct>, TA293 <fct>, TAMex0091.1 <fct>,
## #   TAMex0153.1 <fct>, TAMex0238.1 <fct>, TAMex0396.1 <fct>, TAMex0446.1 <fct>,
## #   TAMex0455.2 <fct>, TAMex0520.1 <fct>, TAMex0608.5 <fct>, TAMex0624.1 <fct>,
## #   TAMex0643.1 <fct>, TAMex0775.2 <fct>, TAMex0930.1 <fct>, TAMex1083.1 <fct>,
## #   TAMex1111.1 <fct>, TAMex1195.2 <fct>, TAMex1222.2 <fct>, …
##
## [[4]]
## # A tibble: 91 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct>
##  1 PZA03013… 4          139753   A/A   A/G   A/A   A/G   A/G   A/G   ?/?   A/G
##  2 PZA03013… 4          139810   T/T   C/C   T/T   C/T   C/T   T/T   C/T   C/T
##  3 PZA00216… 4          2589400  A/A   A/A   G/G   A/G   A/A   A/G   A/A   A/A
##  4 PZA02972… 4          2848110  T/T   C/C   ?/?   C/C   C/C   C/C   C/C   C/C
##  5 PZA00700… 4          2852810  A/A   G/G   A/A   G/G   G/G   G/G   A/A   G/G
##  6 PZA00436… 4          6401332  C/T   C/T   C/T   T/T   T/T   T/T   C/T   C/T
##  7 PZA00078… 4          18502179 A/A   C/C   A/C   A/C   A/A   A/A   A/C   C/C
##  8 PZA01240… 4          19461323 G/G   A/G   G/G   G/G   G/G   ?/?   ?/?   G/G
##  9 PZA01240… 4          19461458 G/G   A/G   G/G   G/G   G/G   G/G   G/G   G/G
## 10 PZA00139… 4          20065054 ?/?   A/A   A/A   A/A   A/A   A/A   A/G   A/A
## # ℹ 81 more rows
## # ℹ 967 more variables: TA057 <fct>, TA137 <fct>, TA153 <fct>, TA205 <fct>,
## #   TA217 <fct>, TA242 <fct>, TA285 <fct>, TA293 <fct>, TAMex0091.1 <fct>,
## #   TAMex0153.1 <fct>, TAMex0238.1 <fct>, TAMex0396.1 <fct>, TAMex0446.1 <fct>,
## #   TAMex0455.2 <fct>, TAMex0520.1 <fct>, TAMex0608.5 <fct>, TAMex0624.1 <fct>,
## #   TAMex0643.1 <fct>, TAMex0775.2 <fct>, TAMex0930.1 <fct>, TAMex1083.1 <fct>,
## #   TAMex1111.1 <fct>, TAMex1195.2 <fct>, TAMex1222.2 <fct>, …
##
## [[5]]
## # A tibble: 122 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct>
##  1 PZA00684… 5          920922   C/C   C/C   C/C   C/C   C/C   C/C   C/C   C/C
##  2 PZA02894… 5          945545   C/G   C/C   C/G   C/C   C/C   C/G   ?/?   C/G
##  3 PZA00191… 5          2123914  C/T   C/T   C/C   C/C   C/T   T/T   T/T   T/T
##  4 PZA03046… 5          3194670  C/T   C/T   C/C   C/C   C/C   C/C   ?/?   C/C
##  5 PZA03046… 5          3194673  C/T   C/T   T/T   T/T   ?/?   T/T   T/T   T/T
##  6 PZA03034… 5          6784084  C/G   C/C   C/C   C/C   C/C   C/C   C/C   C/G
##  7 PZB00081… 5          6960673  A/A   ?/?   A/A   A/A   A/A   A/G   A/A   A/A
##  8 PZB00081… 5          6960836  C/T   T/T   C/C   T/T   T/T   C/C   C/T   T/T
##  9 PZB00081… 5          6960889  C/C   T/T   C/C   T/T   C/T   C/C   C/T   ?/?
## 10 PZB00081… 5          6961472  A/A   A/A   A/A   A/A   A/A   A/T   A/A   A/A
## # ℹ 112 more rows
## # ℹ 967 more variables: TA057 <fct>, TA137 <fct>, TA153 <fct>, TA205 <fct>,
```

```
## #   TA217 <fct>, TA242 <fct>, TA285 <fct>, TA293 <fct>, TAMex0091.1 <fct>,
## #   TAMex0153.1 <fct>, TAMex0238.1 <fct>, TAMex0396.1 <fct>, TAMex0446.1 <fct>,
## #   TAMex0455.2 <fct>, TAMex0520.1 <fct>, TAMex0608.5 <fct>, TAMex0624.1 <fct>,
## #   TAMex0643.1 <fct>, TAMex0775.2 <fct>, TAMex0930.1 <fct>, TAMex1083.1 <fct>,
## #   TAMex1111.1 <fct>, TAMex1195.2 <fct>, TAMex1222.2 <fct>, …
##
## [[6]]
## # A tibble: 76 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct>
##  1 PZA03290… 6          9243581  G/G   G/G   G/G   A/G   G/G   G/G   G/G   ?/?
##  2 PZA03290… 6          9243719  C/C   C/C   C/C   C/C   C/C   C/C   C/T   C/C
##  3 PZA00503… 6          12075777 C/C   C/C   C/C   C/C   C/C   C/C   C/C   C/C
##  4 PZA03047… 6          20837292 G/G   G/G   G/G   G/G   G/G   G/G   G/G   G/G
##  5 PZA03047… 6          20837423 C/C   C/C   C/C   C/T   C/C   C/C   C/C   C/C
##  6 PZA03047… 6          20837504 ?/?   ?/?   ?/?   G/G   G/G   G/G   G/G   G/G
##  7 PZA00440… 6          22404308 C/G   C/G   C/G   C/C   G/G   G/G   C/G   G/G
##  8 PZA00501… 6          23874393 G/T   T/T   G/G   G/T   G/G   G/T   G/G   G/T
##  9 PZA00501… 6          23874474 T/T   T/T   T/T   T/T   T/T   T/T   T/T   T/T
## 10 PZA03063… 6          25335322 T/T   T/T   T/T   T/T   T/T   T/T   T/T   T/T
## # ℹ 66 more rows
## # ℹ 967 more variables: TA057 <fct>, TA137 <fct>, TA153 <fct>, TA205 <fct>,
## #   TA217 <fct>, TA242 <fct>, TA285 <fct>, TA293 <fct>, TAMex0091.1 <fct>,
## #   TAMex0153.1 <fct>, TAMex0238.1 <fct>, TAMex0396.1 <fct>, TAMex0446.1 <fct>,
## #   TAMex0455.2 <fct>, TAMex0520.1 <fct>, TAMex0608.5 <fct>, TAMex0624.1 <fct>,
## #   TAMex0643.1 <fct>, TAMex0775.2 <fct>, TAMex0930.1 <fct>, TAMex1083.1 <fct>,
## #   TAMex1111.1 <fct>, TAMex1195.2 <fct>, TAMex1222.2 <fct>, …
##
## [[7]]
## # A tibble: 97 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct>
##  1 PZB00140… 7          2081333  T/T   T/T   T/T   C/C   C/T   C/C   C/C   T/T
##  2 PZA00188… 7          2998280  A/T   T/T   T/T   T/T   T/T   T/T   A/T   A/A
##  3 PZA00188… 7          2998619  A/A   A/G   G/G   G/G   G/G   A/A   A/A   A/G
##  4 PZA02983… 7          4244157  G/G   G/G   G/G   A/G   A/G   A/A   A/A   G/G
##  5 PZA02872… 7          13058709 A/A   A/G   A/A   G/G   A/G   A/G   A/G   A/G
##  6 PZA02872… 7          13058813 G/G   A/G   G/G   G/G   A/G   A/G   G/G   G/G
##  7 PZA03067… 7          14354166 C/C   C/T   C/C   C/C   C/T   C/C   C/C   C/C
##  8 PZA03067… 7          14354237 G/G   G/G   A/A   A/G   A/G   G/G   ?/?   A/A
##  9 PZB01228… 7          15335449 A/A   A/A   A/A   A/A   A/A   A/A   A/A   A/A
## 10 PZB01228… 7          15335531 G/G   A/A   G/G   A/A   A/G   A/A   A/G   A/G
## # ℹ 87 more rows
## # ℹ 967 more variables: TA057 <fct>, TA137 <fct>, TA153 <fct>, TA205 <fct>,
## #   TA217 <fct>, TA242 <fct>, TA285 <fct>, TA293 <fct>, TAMex0091.1 <fct>,
## #   TAMex0153.1 <fct>, TAMex0238.1 <fct>, TAMex0396.1 <fct>, TAMex0446.1 <fct>,
## #   TAMex0455.2 <fct>, TAMex0520.1 <fct>, TAMex0608.5 <fct>, TAMex0624.1 <fct>,
## #   TAMex0643.1 <fct>, TAMex0775.2 <fct>, TAMex0930.1 <fct>, TAMex1083.1 <fct>,
## #   TAMex1111.1 <fct>, TAMex1195.2 <fct>, TAMex1222.2 <fct>, …
##
## [[8]]
## # A tibble: 62 × 978
```

```
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct>
##  1 PZA03316… 8          4766694  T/T   C/C   T/T   T/T   C/C   T/T   T/T   T/T
##  2 PZA00058… 8          5966698  C/C   C/C   C/C   C/C   C/C   C/C   C/C   C/C
##  3 PZA00058… 8          5966738  G/T   G/T   G/G   G/G   G/G   G/G   G/G   G/G
##  4 PZA02955… 8          14721554 A/A   A/A   A/A   A/A   C/C   A/A   C/C   A/A
##  5 PZB00096… 8          21135142 C/C   C/C   C/C   C/C   ?/?   C/C   C/C   C/C
##  6 PZB00096… 8          21135254 C/C   ?/?   C/C   C/C   C/C   C/C   C/C   C/C
##  7 PZB00049… 8          21136819 G/G   G/G   G/G   G/G   G/G   G/G   G/G   G/G
##  8 PZB00049… 8          21137213 A/A   G/G   G/G   A/A   G/G   A/A   A/G   G/G
##  9 PZB00049… 8          21137286 C/C   T/T   T/T   C/C   T/T   C/C   C/T   T/T
## 10 PZD00024… 8          22274672 C/C   C/C   C/C   C/C   C/C   C/C   C/C   C/C
## # ℹ 52 more rows
## # ℹ 967 more variables: TA057 <fct>, TA137 <fct>, TA153 <fct>, TA205 <fct>,
## #   TA217 <fct>, TA242 <fct>, TA285 <fct>, TA293 <fct>, TAMex0091.1 <fct>,
## #   TAMex0153.1 <fct>, TAMex0238.1 <fct>, TAMex0396.1 <fct>, TAMex0446.1 <fct>,
## #   TAMex0455.2 <fct>, TAMex0520.1 <fct>, TAMex0608.5 <fct>, TAMex0624.1 <fct>,
## #   TAMex0643.1 <fct>, TAMex0775.2 <fct>, TAMex0930.1 <fct>, TAMex1083.1 <fct>,
## #   TAMex1111.1 <fct>, TAMex1195.2 <fct>, TAMex1222.2 <fct>, …
##
## [[9]]
## # A tibble: 60 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct>
##  1 PZA03062… 9          3873116  C/T   C/T   C/C   C/C   ?/?   C/C   ?/?   C/T
##  2 PZA03062… 9          3873152  C/T   C/T   C/C   C/C   C/C   C/C   ?/?   C/T
##  3 PZA00466… 9          11972467 A/G   G/G   G/G   G/G   G/G   G/G   G/G   G/G
##  4 PZD00049… 9          16960436 G/G   G/G   G/G   A/G   G/G   G/G   G/G   G/G
##  5 PZD00049… 9          16960510 G/T   T/T   T/T   T/T   T/T   T/T   T/T   T/T
##  6 PZD00049… 9          16960633 A/G   G/G   G/G   G/G   ?/?   G/G   G/G   G/G
##  7 zmm3.4    9          16966348 C/C   ?/?   C/C   C/C   C/C   C/C   C/C   T/T
##  8 PZA02844… 9          17653885 C/C   C/C   C/C   C/C   C/C   C/C   C/C   C/C
##  9 PZA01782… 9          19252371 C/C   C/C   C/C   C/C   C/C   C/T   C/C   C/C
## 10 PZA01782… 9          19252475 C/C   C/C   C/C   C/C   C/C   T/T   C/T   C/C
## # ℹ 50 more rows
## # ℹ 967 more variables: TA057 <fct>, TA137 <fct>, TA153 <fct>, TA205 <fct>,
## #   TA217 <fct>, TA242 <fct>, TA285 <fct>, TA293 <fct>, TAMex0091.1 <fct>,
## #   TAMex0153.1 <fct>, TAMex0238.1 <fct>, TAMex0396.1 <fct>, TAMex0446.1 <fct>,
## #   TAMex0455.2 <fct>, TAMex0520.1 <fct>, TAMex0608.5 <fct>, TAMex0624.1 <fct>,
## #   TAMex0643.1 <fct>, TAMex0775.2 <fct>, TAMex0930.1 <fct>, TAMex1083.1 <fct>,
## #   TAMex1111.1 <fct>, TAMex1195.2 <fct>, TAMex1222.2 <fct>, …
##
## [[10]]
## # A tibble: 53 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <fct> <fct> <fct> <fct> <fct> <fct> <fct> <fct>
##  1 PZA03078… 10         6121326  T/T   T/T   C/C   T/T   T/T   C/T   T/T   T/T
##  2 PZA00257… 10         10432605 G/G   G/G   G/G   A/G   G/G   G/G   G/G   G/G
##  3 PZA00257… 10         10432680 C/T   T/T   T/T   ?/?   C/T   C/T   C/T   T/T
##  4 PZA00463… 10         13546326 C/T   T/T   T/T   T/T   T/T   T/T   ?/?   T/T
##  5 PZA02961… 10         16264730 C/C   A/C   A/A   A/C   A/C   A/A   A/C   C/C
##  6 PZA00656… 10         17645859 A/A   G/G   G/G   A/G   G/G   G/G   A/G   A/G
```

```
##  7 PZA00656… 10         17645938 A/G   G/G   G/G   G/G   G/G   G/G   A/G   A/G
##  8 PZA00656… 10         17646056 C/T   C/C   C/C   C/T   C/C   C/C   C/C   C/C
##  9 PZA00656… 10         17646117 A/G   A/A   A/A   A/G   A/A   A/A   A/G   A/G
## 10 PZA00079… 10         18903460 C/C   C/C   C/C   C/C   C/C   C/C   C/C   C/C
## # ℹ 43 more rows
## # ℹ 967 more variables: TA057 <fct>, TA137 <fct>, TA153 <fct>, TA205 <fct>,
## #   TA217 <fct>, TA242 <fct>, TA285 <fct>, TA293 <fct>, TAMex0091.1 <fct>,
## #   TAMex0153.1 <fct>, TAMex0238.1 <fct>, TAMex0396.1 <fct>, TAMex0446.1 <fct>,
## #   TAMex0455.2 <fct>, TAMex0520.1 <fct>, TAMex0608.5 <fct>, TAMex0624.1 <fct>,
## #   TAMex0643.1 <fct>, TAMex0775.2 <fct>, TAMex0930.1 <fct>, TAMex1083.1 <fct>,
## #   TAMex1111.1 <fct>, TAMex1195.2 <fct>, TAMex1222.2 <fct>, …
```

```
# teosinte decrease file generate
teosinte_chromosome_decrease <- function(chr) {
  joint_snp_teosinte %>%
    filter(Chromosome == chr) %>%
    mutate(Numeric_Position = suppressWarnings(as.numeric(Position))) %>%
    arrange(is.na(Numeric_Position), desc(Numeric_Position)) %>%
    select(-Numeric_Position) %>%
    mutate(across(4:ncol(.), ~ ifelse(. == "?/?", "-/-", .))) %>%
    mutate(across(4:ncol(.), as.character)) %>%
    write_tsv(paste0("teosinte_data/teosinte_chrom", chr, "_decrease.txt"))
}

lapply(1:10, teosinte_chromosome_decrease)
```

```
## [[1]]
## # A tibble: 155 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
##  1 PZA00230… 1          2984129… -/-   7     9     9     7     9     9     9
##  2 PZA00477… 1          2980826… 9     9     9     9     9     9     9     9
##  3 PZA00477… 1          2980825… 8     6     8     8     11    8     8     6
##  4 PZA00477… 1          2980825… 9     7     9     9     9     9     9     9
##  5 PZA00477… 1          2980824… 10    11    11    10    -/-   10    11    11
##  6 PZA00243… 1          2957711… 2     2     2     2     4     2     2     2
##  7 PZA00050… 1          2954595… 2     2     2     2     2     2     2     2
##  8 PZA00623… 1          2936327… 9     9     -/-   9     6     9     9     6
##  9 PZA00235… 1          2927289… 4     9     9     4     4     9     -/-   9
## 10 PZA00235… 1          2927288… 10    10    9     10    10    -/-   11    11
## # ℹ 145 more rows
## # ℹ 967 more variables: TA057 <chr>, TA137 <chr>, TA153 <chr>, TA205 <chr>,
## #   TA217 <chr>, TA242 <chr>, TA285 <chr>, TA293 <chr>, TAMex0091.1 <chr>,
## #   TAMex0153.1 <chr>, TAMex0238.1 <chr>, TAMex0396.1 <chr>, TAMex0446.1 <chr>,
## #   TAMex0455.2 <chr>, TAMex0520.1 <chr>, TAMex0608.5 <chr>, TAMex0624.1 <chr>,
## #   TAMex0643.1 <chr>, TAMex0775.2 <chr>, TAMex0930.1 <chr>, TAMex1083.1 <chr>,
## #   TAMex1111.1 <chr>, TAMex1195.2 <chr>, TAMex1222.2 <chr>, …
##
## [[2]]
## # A tibble: 127 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
##  1 PZD00022… 2          2331291… 6     6     6     6     6     -/-   8     8
##  2 PZD00022… 2          2331290… 11    8     11    11    11    11    8     8
##  3 zap1.2    2          2331285… 2     4     4     2     -/-   4     9     4
##  4 PZD00022… 2          2331285… 11    6     8     8     6     6     6     8
##  5 PZA00672… 2          2289910… -/-   9     -/-   9     9     -/-   9     9
##  6 PZA00672… 2          2289910… 4     9     9     2     9     9     4     9
##  7 PZA00487… 2          2221352… 9     -/-   9     4     2     9     9     2
##  8 PZA00487… 2          2221349… 4     2     2     4     2     2     2     4
##  9 PZA00487… 2          2221349… 6     6     6     6     6     6     6     7
## 10 PZA00163… 2          2202647… 6     -/-   -/-   3     3     6     6     6
## # ℹ 117 more rows
## # ℹ 967 more variables: TA057 <chr>, TA137 <chr>, TA153 <chr>, TA205 <chr>,
## #   TA217 <chr>, TA242 <chr>, TA285 <chr>, TA293 <chr>, TAMex0091.1 <chr>,
## #   TAMex0153.1 <chr>, TAMex0238.1 <chr>, TAMex0396.1 <chr>, TAMex0446.1 <chr>,
## #   TAMex0455.2 <chr>, TAMex0520.1 <chr>, TAMex0608.5 <chr>, TAMex0624.1 <chr>,
## #   TAMex0643.1 <chr>, TAMex0775.2 <chr>, TAMex0930.1 <chr>, TAMex1083.1 <chr>,
## #   TAMex1111.1 <chr>, TAMex1195.2 <chr>, TAMex1222.2 <chr>, …
##
## [[3]]
## # A tibble: 107 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
##  1 PZA00088… 3          2286142… 7     6     7     6     7     7     -/-   7
##  2 PZA00420… 3          2268722… 8     6     6     6     8     6     8     6
##  3 PZA02789… 3          2221146… 6     6     2     6     6     6     6     6
##  4 PZA02789… 3          2221145… 10    9     9     9     10    9     9     9
```

```
##  5 PZA00234… 3          2217306… 6     6     2     6     6     6     6     3
##  6 PZA00219… 3          2193090… -/-   6     -/-   6     11    6     8     11
##  7 PZA02824… 3          2173296… -/-   4     2     -/-   4     9     4     2
##  8 PZA02824… 3          2173296… 6     6     11    6     6     6     6     6
##  9 sh2.9     3          2148411… 2     2     2     2     2     2     2     2
## 10 sh2.5     3          2148410… 11    11    11    8     11    11    11    11
## # ℹ 97 more rows
## # ℹ 967 more variables: TA057 <chr>, TA137 <chr>, TA153 <chr>, TA205 <chr>,
## #   TA217 <chr>, TA242 <chr>, TA285 <chr>, TA293 <chr>, TAMex0091.1 <chr>,
## #   TAMex0153.1 <chr>, TAMex0238.1 <chr>, TAMex0396.1 <chr>, TAMex0446.1 <chr>,
## #   TAMex0455.2 <chr>, TAMex0520.1 <chr>, TAMex0608.5 <chr>, TAMex0624.1 <chr>,
## #   TAMex0643.1 <chr>, TAMex0775.2 <chr>, TAMex0930.1 <chr>, TAMex1083.1 <chr>,
## #   TAMex1111.1 <chr>, TAMex1195.2 <chr>, TAMex1222.2 <chr>, …
##
## [[4]]
## # A tibble: 91 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
##  1 PZA00051… 4          2450516… 9     2     2     2     4     2     2     2
##  2 PZA00682… 4          2449928… 9     9     6     6     -/-   9     9     6
##  3 PZA03017… 4          2432228… 6     8     11    6     6     6     6     11
##  4 PZA03017… 4          2432228… 6     6     6     11    11    11    11    6
##  5 PZA00005… 4          2430034… 8     11    11    -/-   8     8     6     6
##  6 PZA00005… 4          2430033… 10    9     9     10    10    10    11    11
##  7 PZA00529… 4          2407691… 4     4     9     9     9     4     9     4
##  8 PZA00636… 4          2268846… 4     4     2     4     9     9     2     9
##  9 PZA02865… 4          2208578… 4     2     4     2     4     2     2     2
## 10 PZA03081… 4          2153934… 7     7     6     7     9     9     6     7
## # ℹ 81 more rows
## # ℹ 967 more variables: TA057 <chr>, TA137 <chr>, TA153 <chr>, TA205 <chr>,
## #   TA217 <chr>, TA242 <chr>, TA285 <chr>, TA293 <chr>, TAMex0091.1 <chr>,
## #   TAMex0153.1 <chr>, TAMex0238.1 <chr>, TAMex0396.1 <chr>, TAMex0446.1 <chr>,
## #   TAMex0455.2 <chr>, TAMex0520.1 <chr>, TAMex0608.5 <chr>, TAMex0624.1 <chr>,
## #   TAMex0643.1 <chr>, TAMex0775.2 <chr>, TAMex0930.1 <chr>, TAMex1083.1 <chr>,
## #   TAMex1111.1 <chr>, TAMex1195.2 <chr>, TAMex1222.2 <chr>, …
##
## [[5]]
## # A tibble: 122 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
##  1 PZA00408… 5          2144428… 5     5     5     2     2     2     2     2
##  2 PZA00069… 5          2144249… 6     6     6     8     8     -/-   6     6
##  3 PZA00433… 5          2132715… 9     4     2     4     -/-   4     9     9
##  4 PZA03335… 5          2108908… 9     9     10    9     10    10    9     11
##  5 PZA03335… 5          2108907… 11    11    11    11    11    8     8     -/-
##  6 PZA00545… 5          2066650… 9     9     10    9     10    9     9     10
##  7 PZA00545… 5          2066649… 6     6     8     6     8     6     6     8
##  8 PZA00545… 5          2066647… 6     6     6     8     6     8     6     6
##  9 PZA00395… 5          2028472… -/-   -/-   6     8     6     6     6     6
## 10 PZA00395… 5          2028471… 2     2     6     3     6     6     6     6
## # ℹ 112 more rows
## # ℹ 967 more variables: TA057 <chr>, TA137 <chr>, TA153 <chr>, TA205 <chr>,
```

```
## #   TA217 <chr>, TA242 <chr>, TA285 <chr>, TA293 <chr>, TAMex0091.1 <chr>,
## #   TAMex0153.1 <chr>, TAMex0238.1 <chr>, TAMex0396.1 <chr>, TAMex0446.1 <chr>,
## #   TAMex0455.2 <chr>, TAMex0520.1 <chr>, TAMex0608.5 <chr>, TAMex0624.1 <chr>,
## #   TAMex0643.1 <chr>, TAMex0775.2 <chr>, TAMex0930.1 <chr>, TAMex1083.1 <chr>,
## #   TAMex1111.1 <chr>, TAMex1195.2 <chr>, TAMex1222.2 <chr>, …
##
## [[6]]
## # A tibble: 76 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
##  1 PZA03089… 6          1657249… 9     9     6     9     9     9     7     9
##  2 PZA00198… 6          1644703… 6     6     6     -/-   6     6     -/-   6
##  3 PZB01222… 6          1644194… 6     6     -/-   7     7     6     6     9
##  4 PZB01222… 6          1644193… 6     6     6     8     8     6     6     6
##  5 PZA00266… 6          1615701… 6     6     6     8     6     6     -/-   6
##  6 PZA03348… 6          1598081… 3     2     2     2     3     2     6     3
##  7 PZA00042… 6          1571293… 5     2     11    2     2     2     5     2
##  8 PZA00042… 6          1571291… 8     11    6     11    11    11    11    11
##  9 PZA02958… 6          1519934… 8     8     11    8     11    8     11    6
## 10 PZA00241… 6          1504706… 11    -/-   11    8     11    11    11    11
## # ℹ 66 more rows
## # ℹ 967 more variables: TA057 <chr>, TA137 <chr>, TA153 <chr>, TA205 <chr>,
## #   TA217 <chr>, TA242 <chr>, TA285 <chr>, TA293 <chr>, TAMex0091.1 <chr>,
## #   TAMex0153.1 <chr>, TAMex0238.1 <chr>, TAMex0396.1 <chr>, TAMex0446.1 <chr>,
## #   TAMex0455.2 <chr>, TAMex0520.1 <chr>, TAMex0608.5 <chr>, TAMex0624.1 <chr>,
## #   TAMex0643.1 <chr>, TAMex0775.2 <chr>, TAMex0930.1 <chr>, TAMex1083.1 <chr>,
## #   TAMex1111.1 <chr>, TAMex1195.2 <chr>, TAMex1222.2 <chr>, …
##
## [[7]]
## # A tibble: 97 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
##  1 PZA02850… 7          1690679… 2     2     2     2     -/-   2     2     3
##  2 PZA02850… 7          1690677… 6     6     6     6     6     7     9     6
##  3 PZA00424… 7          1690175… 11    8     11    8     11    11    11    8
##  4 PZA00043… 7          1672182… 11    8     11    8     11    11    11    11
##  5 PZA00423… 7          1670607… 6     6     -/-   -/-   -/-   6     6     -/-
##  6 PZA00423… 7          1670606… 6     -/-   6     8     6     8     6     6
##  7 PZA03295… 7          1670148… 6     6     6     6     6     6     6     6
##  8 PZB00603… 7          1659948… 9     7     9     7     6     6     -/-   9
##  9 PZB00603… 7          1659948… 7     7     6     -/-   -/-   6     7     6
## 10 PZB00603… 7          1659948… 2     2     2     2     2     2     2     2
## # ℹ 87 more rows
## # ℹ 967 more variables: TA057 <chr>, TA137 <chr>, TA153 <chr>, TA205 <chr>,
## #   TA217 <chr>, TA242 <chr>, TA285 <chr>, TA293 <chr>, TAMex0091.1 <chr>,
## #   TAMex0153.1 <chr>, TAMex0238.1 <chr>, TAMex0396.1 <chr>, TAMex0446.1 <chr>,
## #   TAMex0455.2 <chr>, TAMex0520.1 <chr>, TAMex0608.5 <chr>, TAMex0624.1 <chr>,
## #   TAMex0643.1 <chr>, TAMex0775.2 <chr>, TAMex0930.1 <chr>, TAMex1083.1 <chr>,
## #   TAMex1111.1 <chr>, TAMex1195.2 <chr>, TAMex1222.2 <chr>, …
##
## [[8]]
## # A tibble: 62 × 978
```

```
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
##  1 PZA00298… 8          1669800… 2     2     2     2     2     2     4     2
##  2 PZA00298… 8          1669799… 11    11    6     11    11    11    11    11
##  3 PZA00641… 8          1661827… 2     2     2     2     2     -/-   -/-   9
##  4 PZA00641… 8          1661827… 9     9     9     9     9     9     9     6
##  5 PZA00516… 8          1635723… 4     4     2     9     -/-   4     4     2
##  6 PZA00460… 8          1635654… 6     6     6     6     6     6     8     6
##  7 PZA00460… 8          1635654… 8     6     8     8     11    8     6     -/-
##  8 PZA00460… 8          1635653… -/-   8     6     8     8     8     11    6
##  9 PZA00706… 8          1598979… 8     6     6     8     6     8     11    6
## 10 PZB00160… 8          1574730… 11    8     11    11    8     11    8     8
## # ℹ 52 more rows
## # ℹ 967 more variables: TA057 <chr>, TA137 <chr>, TA153 <chr>, TA205 <chr>,
## #   TA217 <chr>, TA242 <chr>, TA285 <chr>, TA293 <chr>, TAMex0091.1 <chr>,
## #   TAMex0153.1 <chr>, TAMex0238.1 <chr>, TAMex0396.1 <chr>, TAMex0446.1 <chr>,
## #   TAMex0455.2 <chr>, TAMex0520.1 <chr>, TAMex0608.5 <chr>, TAMex0624.1 <chr>,
## #   TAMex0643.1 <chr>, TAMex0775.2 <chr>, TAMex0930.1 <chr>, TAMex1083.1 <chr>,
## #   TAMex1111.1 <chr>, TAMex1195.2 <chr>, TAMex1222.2 <chr>, …
##
## [[9]]
## # A tibble: 60 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
##  1 PZA00184… 9          1512899… 9     6     9     7     6     7     6     9
##  2 PZA00184… 9          1512898… 9     9     9     9     9     9     9     9
##  3 PZA00123… 9          1493987… 6     8     6     6     6     6     6     8
##  4 PZA00579… 9          1473849… 9     9     9     9     9     9     7     9
##  5 PZA02997… 9          1464706… 6     6     6     11    6     6     6     6
##  6 PZA02997… 9          1464705… 2     2     2     2     2     2     2     2
##  7 PZA00323… 9          1425000… 6     8     6     6     6     6     6     6
##  8 PZA00323… 9          1424998… 11    11    11    11    11    11    11    11
##  9 PZB00221… 9          1422714… 11    11    11    -/-   -/-   11    11    11
## 10 PZB00221… 9          1422710… 2     2     2     2     2     2     4     2
## # ℹ 50 more rows
## # ℹ 967 more variables: TA057 <chr>, TA137 <chr>, TA153 <chr>, TA205 <chr>,
## #   TA217 <chr>, TA242 <chr>, TA285 <chr>, TA293 <chr>, TAMex0091.1 <chr>,
## #   TAMex0153.1 <chr>, TAMex0238.1 <chr>, TAMex0396.1 <chr>, TAMex0446.1 <chr>,
## #   TAMex0455.2 <chr>, TAMex0520.1 <chr>, TAMex0608.5 <chr>, TAMex0624.1 <chr>,
## #   TAMex0643.1 <chr>, TAMex0775.2 <chr>, TAMex0930.1 <chr>, TAMex1083.1 <chr>,
## #   TAMex1111.1 <chr>, TAMex1195.2 <chr>, TAMex1222.2 <chr>, …
##
## [[10]]
## # A tibble: 53 × 978
##    SNP_ID    Chromosome Position S0881 S1057 S1087 S1689 S1697 S1698 S1703 S1704
##    <chr>     <chr>      <chr>    <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
##  1 PZD00075… 10         1465994… 6     6     -/-   6     -/-   6     6     6
##  2 PZD00075… 10         1465991… 7     7     9     6     6     7     6     9
##  3 PZA02969… 10         1436761… 6     6     6     6     6     8     6     6
##  4 PZA00727… 10         1428547… 4     2     2     2     4     9     2     2
##  5 PZA00727… 10         1428546… 6     6     8     8     8     6     6     6
##  6 PZB00055… 10         1404931… 9     7     9     7     7     7     6     6
```

```
##  7 PZB01111… 10         1340348… 8     11    11    11    8     8     11    8
##  8 PZB01111… 10         1340346… 9     -/-   9     9     9     9     9     9
##  9 PZB01111… 10         1340344… 6     6     6     6     6     6     6     -/-
## 10 PZA00342… 10         1301707… 9     9     4     4     9     4     9     9
## # ℹ 43 more rows
## # ℹ 967 more variables: TA057 <chr>, TA137 <chr>, TA153 <chr>, TA205 <chr>,
## #   TA217 <chr>, TA242 <chr>, TA285 <chr>, TA293 <chr>, TAMex0091.1 <chr>,
## #   TAMex0153.1 <chr>, TAMex0238.1 <chr>, TAMex0396.1 <chr>, TAMex0446.1 <chr>,
## #   TAMex0455.2 <chr>, TAMex0520.1 <chr>, TAMex0608.5 <chr>, TAMex0624.1 <chr>,
## #   TAMex0643.1 <chr>, TAMex0775.2 <chr>, TAMex0930.1 <chr>, TAMex1083.1 <chr>,
## #   TAMex1111.1 <chr>, TAMex1195.2 <chr>, TAMex1222.2 <chr>, …
```

Generate long file

```
genotypes_long <- raw_genotypes %>%
  pivot_longer(cols = -c(Sample_ID, JG_OTU, Group), names_to = "SNP_ID", values_to = "Genotype")

joint_snpALL_long <- genotypes_long %>%
  left_join(snp_position, by = "SNP_ID") %>%
  filter(Group %in% c("ZMMIL", "ZMMLR", "ZMMMR","ZMPBA", "ZMPIL", "ZMPJA"))
```

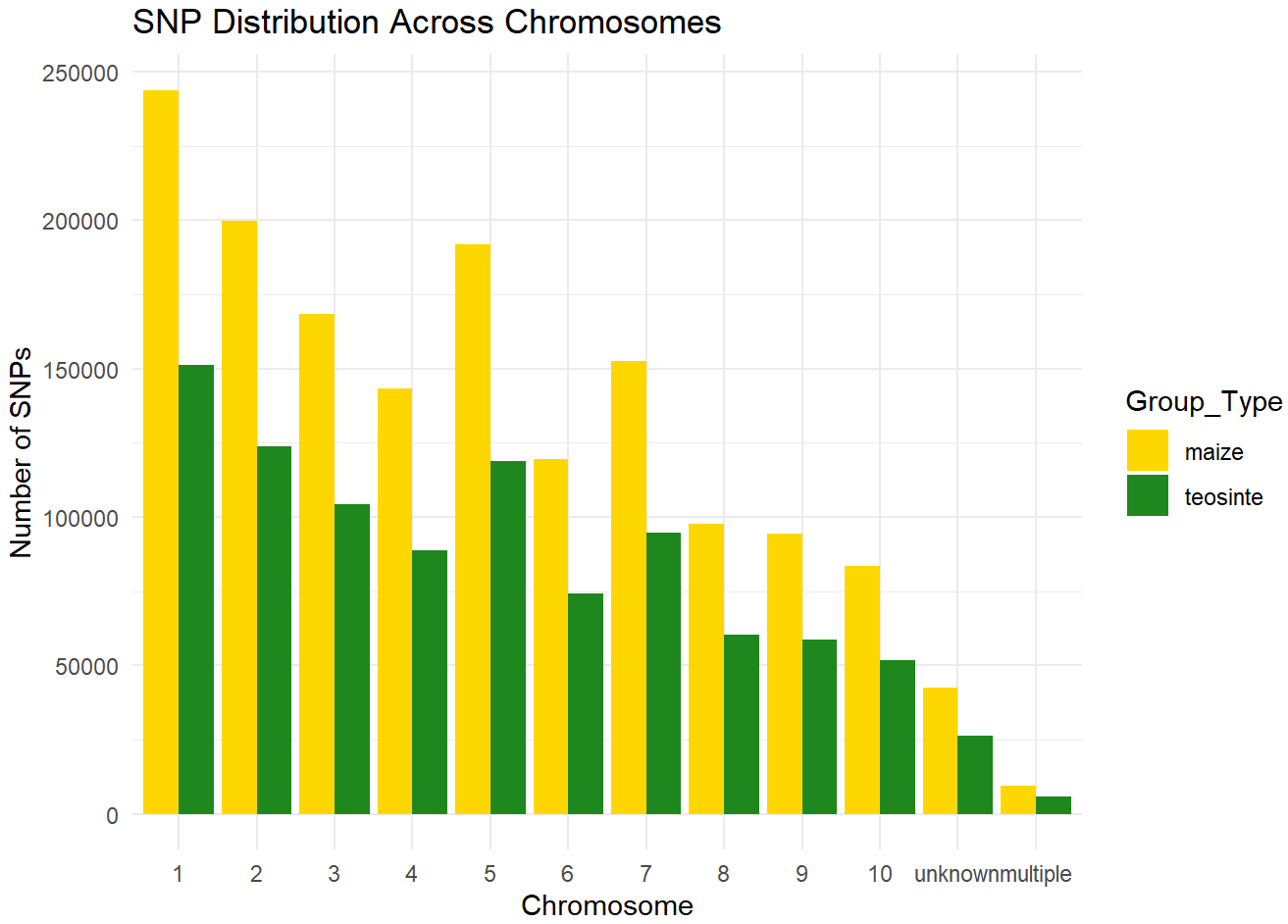
Part II

SNPs per chromosome

```
joint_snpALL_long <- joint_snpALL_long %>%
  mutate(Chromosome_numeric = suppressWarnings(as.numeric(Chromosome)),
         Chromosome = if_else(is.na(Chromosome_numeric), Chromosome, as.character(Chromosome_num
eric)))


joint_snpALL_long <- joint_snpALL_long %>%
  mutate(Chromosome = factor(Chromosome, levels = c(sort(as.numeric(unique(Chromosome[!Chromosom
e %in% c("unknown", "multiple")]))), "unknown", "multiple")))


joint_snpALL_long <- joint_snpALL_long %>%
  mutate(Position = suppressWarnings(as.numeric(Position)))
```

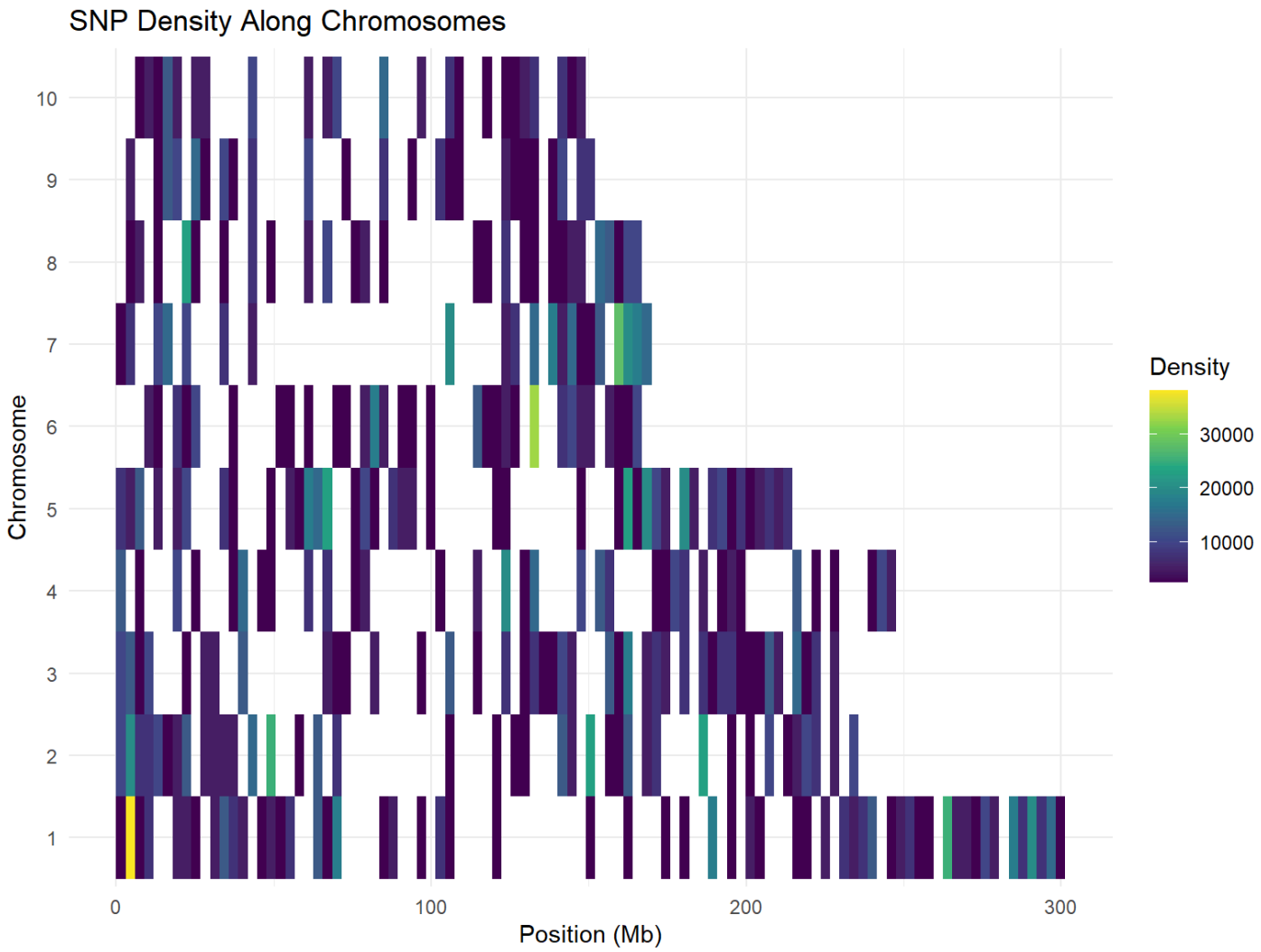
Distribution of SNP on Chromosome

```
joint_snpALL_long <- joint_snpALL_long %>%
  mutate(Group_Type = case_when(
    Group %in% c("ZMMIL", "ZMMLR", "ZMMMR") ~ "maize",
    Group %in% c("ZMPBA", "ZMPIL", "ZMPJA") ~ "teosinte"
  ))

ggplot(joint_snpALL_long, aes(x = Chromosome, fill = Group_Type)) +
  geom_bar(position = "dodge") +
  labs(title = "SNP Distribution Across Chromosomes",
       x = "Chromosome", y = "Number of SNPs") +
  scale_fill_manual(values = c("maize" = "#FFD700", "teosinte" = "#228B22")) +
  theme_minimal()
```



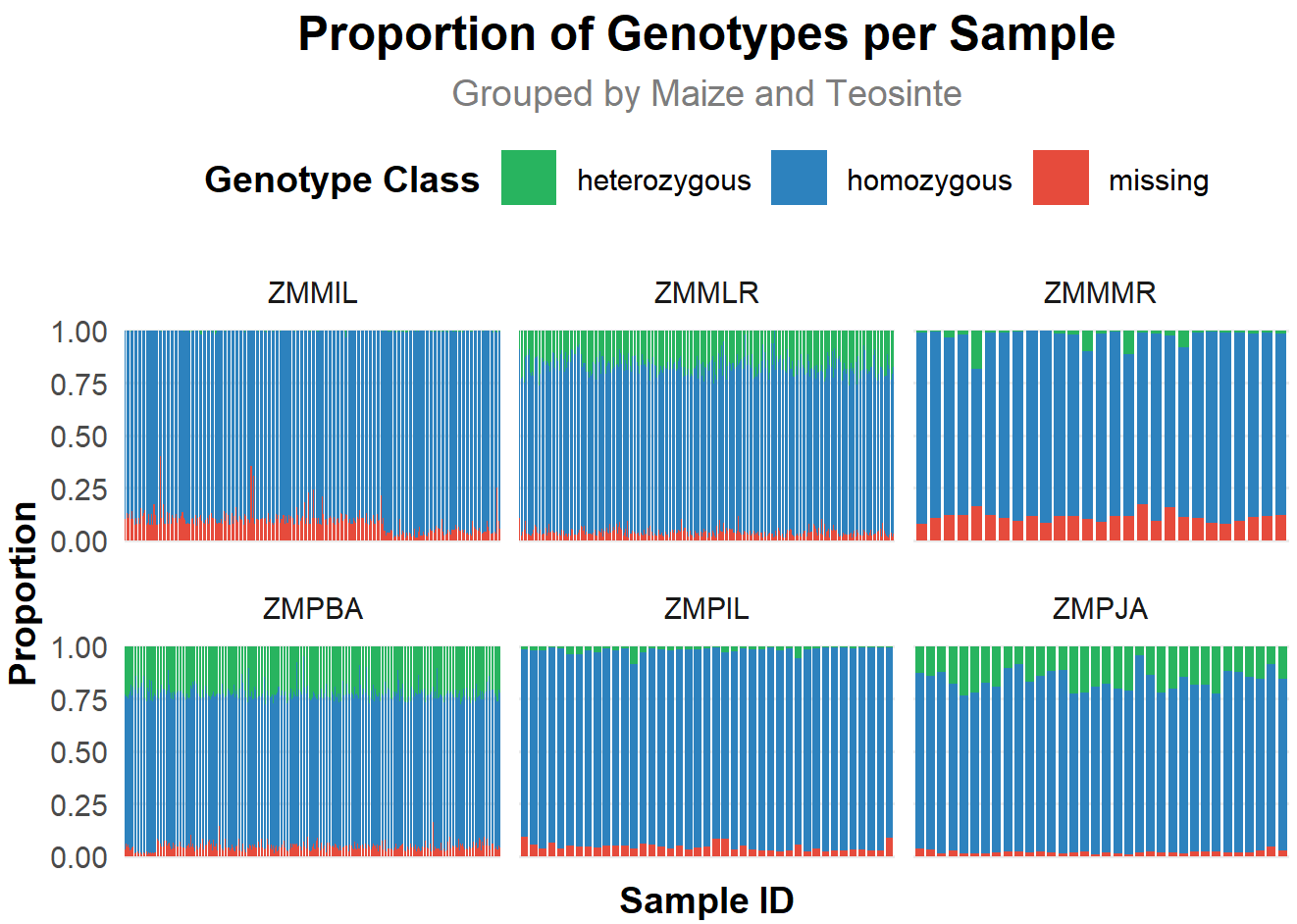
Distribution of SNP Across Chromosome

```
ggplot(joint_snpALL_long %>% filter(!is.na(Position)),
       aes(x = Position / 1e6, y = Chromosome)) +
  geom_bin2d(bins = 100) +
  scale_fill_viridis_c() +
  labs(title = "SNP Density Along Chromosomes",
       x = "Position (Mb)", y = "Chromosome", fill = "Density") +
  theme_minimal()
```


SNP Density Along Chromosomes
Chromosome
10
9
8
7
6
5
4
3
2
1
0
100
200
300
Position (Mb)
Density
30000
20000
10000

Missing data and amount of heterozygosity

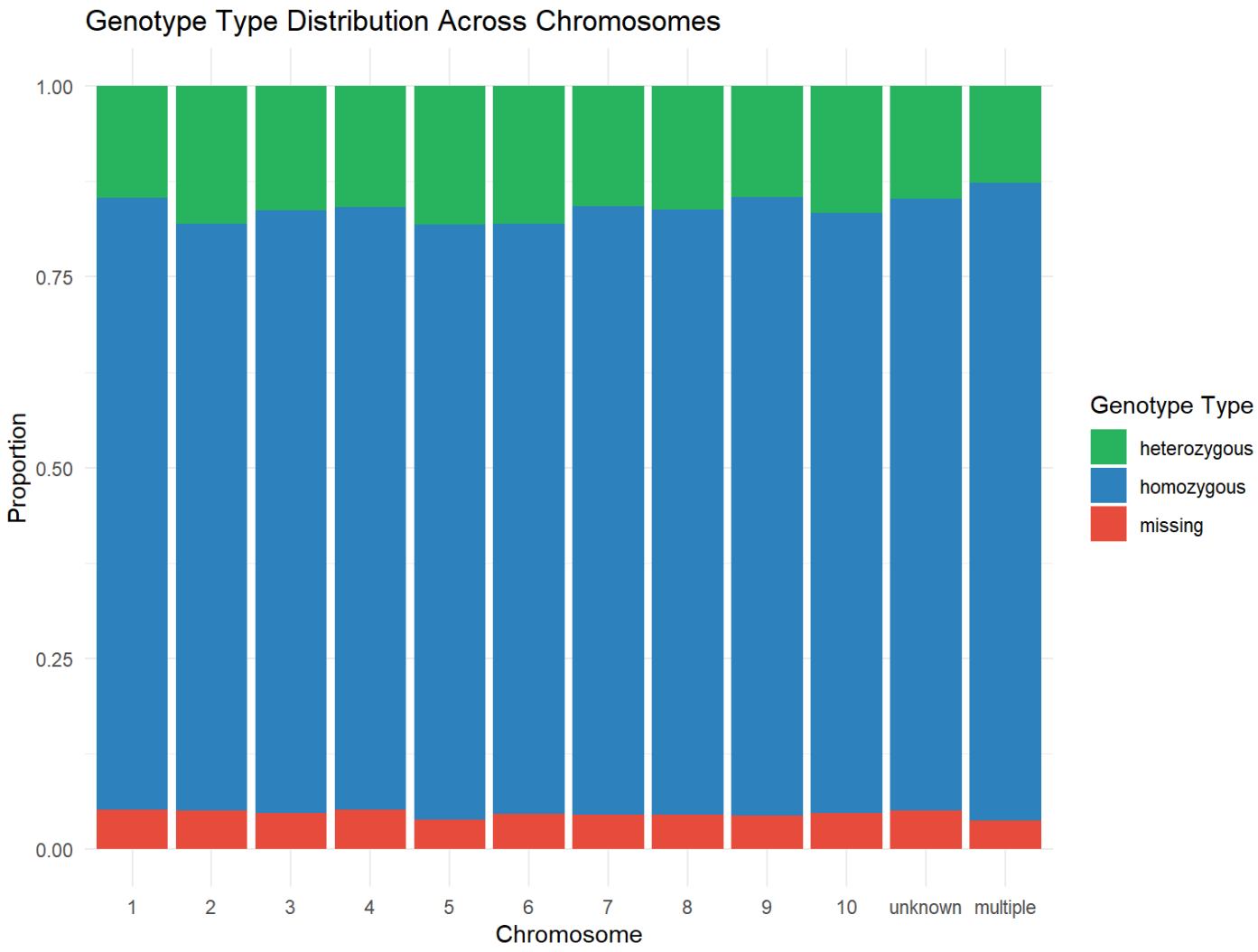
```
genotype_summary <- joint_snpALL_long %>%
  mutate(Genotype_Class = case_when(
    Genotype == "?/?" ~ "missing",
    substr(Genotype, 1, 1) == substr(Genotype, 3, 3) ~ "homozygous",
    TRUE ~ "heterozygous"
  )) %>%
  count(Sample_ID, Group, Genotype_Class) %>%
  group_by(Sample_ID, Group) %>%
  mutate(Proportion = n / sum(n))

ggplot(genotype_summary, aes(x = reorder(Sample_ID, -Proportion, FUN = sum),
                             y = Proportion, fill = Genotype_Class)) +
  geom_col(position = "stack", width = 0.8) +
  scale_fill_manual(values = c("homozygous" = "#2E86C1",
                               "heterozygous" = "#28B463",
                               "missing" = "#E74C3C")) +
  labs(title = "Proportion of Genotypes per Sample",
       subtitle = "Grouped by Maize and Teosinte",
       x = "Sample ID", y = "Proportion", fill = "Genotype Type") +
  facet_wrap(~ Group, scales = "free_x") +
  theme_minimal(base_size = 14) +
  theme(
    plot.title = element_text(face = "bold", size = 18, hjust = 0.5),
    plot.subtitle = element_text(size = 14, hjust = 0.5, color = "gray50"),
    axis.title = element_text(size = 14, face = "bold"),
    axis.text.x = element_blank(),
    axis.ticks.x = element_blank(),
    panel.grid.major.x = element_blank(),
    panel.grid.minor = element_blank(),
    legend.position = "top",
    legend.title = element_text(face = "bold"),
    legend.key.size = unit(0.8, "cm")
  ) +
  guides(fill = guide_legend(title = "Genotype Class"))
```



My own visualization

Genotype Distribution Across Chromosomes

```
chromosome_genotype <- joint_snpALL_long %>%
  mutate(Genotype_Class = case_when(
    Genotype == "?/?" ~ "missing",
    substr(Genotype, 1, 1) == substr(Genotype, 3, 3) ~ "homozygous",
    TRUE ~ "heterozygous"
  )) %>%
  group_by(Chromosome, Genotype_Class) %>%
  summarise(Count = n(), .groups = "drop")

ggplot(chromosome_genotype, aes(x = factor(Chromosome, levels = sort(unique(Chromosome))),
                                y = Count, fill = Genotype_Class)) +
  geom_bar(stat = "identity", position = "fill") +
  scale_fill_manual(values = c("homozygous" = "#2E86C1", "heterozygous" = "#28B463", "missing" =
"#E74C3C")) +
  labs(title = "Genotype Type Distribution Across Chromosomes",
       x = "Chromosome", y = "Proportion", fill = "Genotype Type") +
  theme_minimal()
```


Genotype Type Distribution Across Chromosomes
1.00
0.75
0.50
0.25
0.00
Proportion
1
2
3
4
5
6
7
8
9
10
unknown
multiple
Chromosome
Genotype Type
heterozygous
homozygous
missing