

Puntos de rotura identificados en los cromosomas J y U CODE ▾

Merce Merayo Pastor

2024-11-16

- Introducción
 - Cargar datos procedentes de Breakdancer
 - Cromosoma J
 - Cromosoma U

Introducción

Este informe analiza los puntos de rotura identificados en los cromosomas J y U, utilizando el software Breakdancer .

Cargar datos procedentes de Breakdancer

Cromosoma J

SHOW

```
# A tibble: 6 × 12
  Cromosoma Posicion_1 Orientacion_1 Cromosoma2 Posicion_2 Orientacion_2 Tipo_SV
<chr>      <dbl> <chr>      <chr>      <dbl> <chr>      <chr>
1 Drosophi... 26658 3+0-      Drosophil... 27108 1+5-      DEL
2 Drosophi... 28176 4+9-      Drosophil... 29159 48+43-    INV
3 Drosophi... 28176 4+9-      Drosophil... 29614 1+10-     INV
4 Drosophi... 86444 45+2-     Drosophil... 91558 0+44-     DEL
5 Drosophi... 95272 4+3-      Drosophil... 103385 0+14-     DEL
6 Drosophi... 154831 4+6-     Drosophil... 155325 0+18-     DEL
# i 5 more variables: Longitud <dbl>, Score <dbl>, Num_Reads <dbl>,
#   num_Reads_orig <chr>, Valor <dbl>
```

Se limpia el dataset para dejar solo los campos de interes, concretamente se elimina el nombre del cromosoma 2, ya que solo se trabaja con un cromosoma y el nombre del fichero.

SHOW

```
# A tibble: 6 × 9
  Cromosoma Posicion_1 Orientacion_1 Posicion_2 Orientacion_2 Tipo_SV Longitud
<chr>      <dbl> <chr>      <dbl> <chr>      <chr>      <dbl>
1 Drosophila... 26658 3+0-      27108 1+5-      DEL        430
2 Drosophila... 28176 4+9-      29159 48+43-    INV        617
3 Drosophila... 28176 4+9-      29614 1+10-     INV        1125
4 Drosophila... 86444 45+2-     91558 0+44-     DEL        5188
5 Drosophila... 95272 4+3-      103385 0+14-     DEL        8133
6 Drosophila... 154831 4+6-     155325 0+18-     DEL        424
# i 2 more variables: Score <dbl>, Num_Reads <dbl>
```

Se filtran las inversiones con un Score superior al 90%

SHOW

```
[1] "Se detectan 145 inversiones en el cromosoma J."
```

SHOW

```
[1] "De las inversiones detectadas en el cromosoma J 59 tienen un Score superior al 90%."
```

Las inversiones se espera que tengan más de 1 millón de pares de bases, y se visualiza si al menos alguna de ellas podría ser válida.

SHOW

[1] "Hay 3 inversiones a analizar."

SHOW

A tibble: 3 × 9

```
  Cromosoma  Posicion_1 Orientacion_1 Posicion_2 Orientacion_2 Tipo_SV Longitud
  <chr>      <dbl> <chr>          <dbl> <chr>          <chr>    <dbl>
1 Drosophila... 2122245 5+5-          3313109 6+6-          INV     1190573
2 Drosophila... 10732355 10+0-        12247089 12+2-          INV     1514427
3 Drosophila... 10738885 16+2-        12242422 39+13-         INV     1503162
```

i 2 more variables: Score <dbl>, Num_Reads <dbl>

SHOW

Tabla de inversiones a analizar en el cromosoma J.

| Cromosoma | Posición 1 | Orientación 1 | Posición 2 | Orientación 2 | Tipo SV | Longitud | Score | Número de Lecturas |
|-----------------------------------|------------|---------------|------------|---------------|---------|----------|-------|--------------------|
| Drosophila_subobscura_NC_048532.1 | 2122245 | 5+5- | 3313109 | 6+6- | INV | 1190573 | 99 | 5 |
| Drosophila_subobscura_NC_048532.1 | 10732355 | 10+0- | 12247089 | 12+2- | INV | 1514427 | 99 | 10 |
| Drosophila_subobscura_NC_048532.1 | 10738885 | 16+2- | 12242422 | 39+13- | INV | 1503162 | 99 | 13 |

Cromosoma U

Se sigue el mismo proceso en el cromosoma U que el realizado en el cromosoma J.

SHOW

A tibble: 6 × 12

```
  Cromosoma Posicion_1 Orientacion_1 Cromosoma2 Posicion_2 Orientacion_2 Tipo_SV
  <chr>      <dbl> <chr>          <chr>      <dbl> <chr>          <chr>
1 Drosophi...    1 6186+5107- Drosophil... 1119 6186+5107- ITX
2 Drosophi... 76625 46+0- Drosophil... 77349 0+46- DEL
3 Drosophi... 254869 17+0- Drosophil... 257080 0+16- DEL
4 Drosophi... 278564 14+19- Drosophil... 317583 17+0- INV
5 Drosophi... 278564 14+19- Drosophil... 317931 0+28- INV
6 Drosophi... 310775 0+4- Drosophil... 310836 0+4- INV
```

i 5 more variables: Longitud <dbl>, Score <dbl>, Num_Reads <dbl>,

num_Reads_orig <chr>, Valor <dbl>

Tras dar formato a las columnas

SHOW

A tibble: 6 × 9

```
  Cromosoma  Posicion_1 Orientacion_1 Posicion_2 Orientacion_2 Tipo_SV Longitud
  <chr>      <dbl> <chr>          <dbl> <chr>          <chr>    <dbl>
1 Drosophila...    1 6186+5107-    1119 6186+5107- ITX     -250
2 Drosophila... 76625 46+0-    77349 0+46- DEL       794
3 Drosophila... 254869 17+0-    257080 0+16- DEL     2197
4 Drosophila... 278564 14+19-    317583 17+0- INV     38642
5 Drosophila... 278564 14+19-    317931 0+28- INV     38976
6 Drosophila... 310775 0+4-    310836 0+4- INV     -205
```

i 2 more variables: Score <dbl>, Num_Reads <dbl>

Se filtran las inversiones con un Score superior al 90%

SHOW

[1] "Se detectan 205 inversiones en el cromosoma U."

SHOW

[1] "De las inversiones detectadas en el cromosoma J 97 tienen un Score superior al 90%."

Las inversiones se espera que tengan más de 2 millones de pares de bases, ya que la inversión U_8 se espera que sea bastante grande, y se visualiza si al menos alguna de ellas podría ser válida.

SHOW

```
[1] "Hay 8 inversiones a analizar."
```

SHOW

```
# A tibble: 8 × 9
  Cromosoma Posicion_1 Orientacion_1 Posicion_2 Orientacion_2 Tipo_SV Longitud
  <chr>      <dbl> <chr>          <dbl> <chr>          <dbl>
1 Drosophila... 7772887 49+0-          14722269 48+53-        INV      6948660
2 Drosophila... 606914 12+0-          19484944 12+18-        INV      18877738
3 Drosophila... 607359 0+18-          19484944 12+18-        INV      18877320
4 Drosophila... 16322606 17+26-        23730942 44+0-        INV      7407920
5 Drosophila... 16322606 17+26-        23732120 0+53-        INV      7409125
6 Drosophila... 16189452 8+3-          24359964 7+7-        INV      8170264
7 Drosophila... 16189746 0+5-          24359964 7+7-        INV      8169914
8 Drosophila... 727789 57+17-        24594486 35+18-        INV      22161830
# i 2 more variables: Score <dbl>, Num_Reads <dbl>
```

SHOW

Tabla de inversiones a analizar en el cromosoma U.

| Cromosoma | Posición 1 | Orientación 1 | Posición 2 | Orientación 2 | Tipo SV | Longitud | Score | Número de Lecturas |
|-----------------------------------|------------|---------------|------------|---------------|---------|----------|-------|--------------------|
| Drosophila_subobscura_NC_048534.1 | 7772887 | 49+0- | 14722269 | 48+53- | INV | 6948660 | 99 | 48 |
| Drosophila_subobscura_NC_048534.1 | 606914 | 12+0- | 19484944 | 12+18- | INV | 18877738 | 99 | 12 |
| Drosophila_subobscura_NC_048534.1 | 607359 | 0+18- | 19484944 | 12+18- | INV | 18877320 | 99 | 18 |
| Drosophila_subobscura_NC_048534.1 | 16322606 | 17+26- | 23730942 | 44+0- | INV | 7407920 | 99 | 17 |
| Drosophila_subobscura_NC_048534.1 | 16322606 | 17+26- | 23732120 | 0+53- | INV | 7409125 | 99 | 26 |
| Drosophila_subobscura_NC_048534.1 | 16189452 | 8+3- | 24359964 | 7+7- | INV | 8170264 | 99 | 7 |
| Drosophila_subobscura_NC_048534.1 | 16189746 | 0+5- | 24359964 | 7+7- | INV | 8169914 | 99 | 5 |
| Drosophila_subobscura_NC_048534.1 | 727789 | 57+17- | 24594486 | 35+18- | INV | 22161830 | 99 | 14 |