|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 |
| \*1 | T | A | G | C | T | G |
| \*2 |  | G |  |  |  |  |
| \*3 |  |  | C |  |  |  |
| \*4 |  |  |  | T |  |  |
| \*5 |  |  |  |  | A |  |
| \*6 |  |  |  |  |  | C |
| \*7 | A |  | C |  |  |  |
| \*8 |  | G |  |  |  | C |
| \*9 |  |  | C |  | A |  |
| \*10 |  | G |  | T |  |  |
| \*11 | A |  |  | T |  |  |
| \*12 |  |  |  |  | A | C |

\*1 TAGCTG

\*2 TGGCTG

\*3 TACCTG

\*4 TAGTTG

\*5 TAGCAG

\*6 TAGCTC

\*7 AACCTG

\*8 TGGCTC

\*9 TACCAG

\*10 TGGTTG

\*11 AAGTTG

\*12 TAGCAC

|  |  |
| --- | --- |
| Position | Allele |
| 1 | \*7, \*11 |
| 2 | \*2, \*8, \*10 |
| 3 | \*3, \*7, \*9 |
| 4 | \*4, \*10, \*11 |
| 5 | \*5, \*9, \*12 |
| 6 | \*6, \*8, \*12 |

|  |  |
| --- | --- |
| Allele | Position |
| \*2 | 2 |
| \*3 | 3 |
| \*4 | 4 |
| \*5 | 5 |
| \*6 | 6 |
| \*7 | 1,3 |
| \*8 | 2,6 |
| \*9 | 3,5 |
| \*10 | 2,4 |
| \*11 | 1,4 |
| \*12 | 5,6 |

วิธีหา diplotypes ที่เป็นไปได้ทั้งหมด กรณีมี missing positions (case phased genotypes)

ตัวอย่าง

คำตอบของ diplotype คือ \*2/\*3 , missing positions 2,4,5

T**G**GCTG \*2 (เส้นที่ 1)

TA**C**CTG \*3 (เส้นที่ 2)

หาคำตอบ star allele ทั้งหมดที่เป็นไปได้สำหรับแต่ละเส้น

เช่น คำตอบสำหรับเส้นที่ 1 = \*2 (ซึ่งเกี่ยวข้องกับ position 2)

1. หา all combinations ของ missing positions

2

4

5

2,4

2,5

4,5

2,4,5

2) หา position ที่เกี่ยวข้องกับคำตอบ

2

3) เติม position ที่เกี่ยวข้องกับคำตอบ เข้าไปในแต่ละ combinations ของข้อ 1)

position ที่เกี่ยวข้องกับคำตอบ = position 2

2,2

2,4

2,5

2,2,4

2,2,5

2,4,5

2,2,4,5

4) จากข้อ 1) เลือกมาเฉพาะ combination ที่ประกอบด้วย position ที่เกี่ยวข้องกับคำตอบ แล้วตัด position ที่เกี่ยวข้องกับคำตอบออกไป

combination ที่ประกอบด้วย position ที่เกี่ยวข้องกับคำตอบ = 2; 2,4; 2,5; 2,4,5

ตัด position 2 ออก

4

5

4,5

5) เอา combinations ทั้งหมดของข้อ 2),3),4) มารวมกัน

2

2,2

2,4

2,5

2,2,4

2,2,5

2,4,5

2,2,4,5

4

5

4,5

6) ในแต่ละ combination ข้อ 5) เอา positions ที่ซ้ำกันใน combination ตัวเองออกไป

2

2

2,4

2,5

2,4

2,5

2,4,5

2,4,5

4

5

4,5

7) เอา combination ที่ซ้ำกันในข้อ 6) ออกไป

2

~~2~~

2,4

2,5

~~2,4~~

~~2,5~~

2,4,5

~~2,4,5~~

4

5

4,5

เหลือ

2

2,4

2,5

2,4,5

4

5

4,5

8) หา star allele ที่เกี่ยวข้องกับ positions ในข้อ 7)

2 🡪 \*2

2,4 🡪 \*10

2,5 🡪 ไม่มี

2,4,5 🡪 ไม่มี

4 🡪 \*4

5 🡪 \*5

4,5 🡪 ไม่มี

All possible alleles ที่เป็นคำตอบของ missing ในเส้นที่ 1 = \*2,\*4,\*5,\*10

9) ถ้า positions ทั้งหมดที่เกี่ยวกับคำตอบเป็น missing ให้เพิ่ม \*1 ด้วย

positions ทั้งหมดที่เกี่ยวกับคำตอบ = 2 ซึ่งอยู่ใน list missing positions 2,4,5

ดังนั้น All possible alleles ที่เป็นคำตอบของ missing ในเส้นที่ 1 = \*1,\*2,\*4,\*5,\*10

10) ทำซ้ำข้อ 2)-9) สำหรับเส้นที่ 2

All possible alleles ที่เป็นคำตอบของ missing ในเส้นที่ 2 = \*3,\*9

11) จากข้อ 9) และ 10) หา diplotypes ทั้งหมดที่เป็น combinations ของ alleles ในแต่ละเส้น

T-G--G \*1,\*2,\*4,\*5,\*10

T-C--G \*3,\*9

All possible diplotypes = \*1/\*3, \*1/\*9, \*2/\*3, \*2/\*9, \*3/\*4, \*4/\*9, \*3/\*5, \*5/\*9, \*3/\*10, \*9/\*10

ตัวอย่าง case สำหรับคำตอบของ sample มี diplotype \*2/\*3 (เอาไว้เช็ค script ที่เขียน)

TGGCTG \*2 (เส้นที่ 1)

TACCTG \*3 (เส้นที่ 2)

Missing 1 position

1. missing position 2

เส้นที่ 1 \*1,\*2

เส้นที่ 2 \*3

1. missing position 3

เส้นที่ 1 \*2

เส้นที่ 2 \*1,\*3

1. missing position 4

เส้นที่ 1 \*2,\*10

เส้นที่ 2 \*3

1. missing position 5

เส้นที่ 1 \*2

เส้นที่ 2 \*3,\*9

1. missing position 6

เส้นที่ 1 \*2,\*8

เส้นที่ 2 \*3

Missing 2 positions

1. missing position 2,3

เส้นที่ 1 \*1,\*2,\*3

เส้นที่ 2 \*1,\*2,\*3

1. missing position 2,4

เส้นที่ 1 \*1,\*2,\*4,\*10

เส้นที่ 2 \*3

1. missing position 3,5

เส้นที่ 1 \*2

เส้นที่ 2 \*1,\*3,\*5,\*9

1. missing position 1,4

เส้นที่ 1 \*2,\*10

เส้นที่ 2 \*3,\*7

Missing 3 positions

1. missing position 2,3,4

เส้นที่ 1 \*1,\*2,\*3,\*4,\*10

เส้นที่ 2 \*1,\*2,\*3,\*4,\*10

1. missing position 2,4,5

เส้นที่ 1 \*1,\*2,\*4,\*5,\*10

เส้นที่ 2 \*3,\*9

1. missing position 3,4,5

เส้นที่ 1 \*2,\*10

เส้นที่ 2 \*1,\*3,\*4,\*5,\*9

1. missing position 4,5,6

เส้นที่ 1 \*2,\*8,\*10

เส้นที่ 2 \*3,\*9

Missing 4 positions

1. missing position 2,3,4,5

เส้นที่ 1 \*1,\*2,\*3,\*4,\*5,\*9,\*10

เส้นที่ 2 \*1,\*2,\*3,\*4,\*5,\*9,\*10