(1)

(a)

	A	C	A
A	4	2	9
T	2	2	6
A	9	6	6

(b)

(i) x = A G T A C A

y = A C C A T A

(2) Gaps in Y arise due to moving horizontally in the scoring matrix (provided Y is aligned vertically to the left of the matrix).

Therefore, having the rule,

$$H_{x,y} = \max \begin{cases} H_{x-1,y-1} + m & \text{if } x_i = y_i \\ H_{x-1,y-1} - s & \text{if } x_i \neq y_i \\ H_{x,y-1} - d \end{cases}$$

Using this rule makes sure there are no horizontal movements and hence, the global alignment of Y will have no gaps.

⑤

(a) Here, we want the count of $y=1$ since that implies the individual has phenotype y. $y=0$ implies that individuals with a particular SNP do not have phenotype y.

Therefore, $SNP_{x_1} = \frac{8+4+4}{100}$

$SNP_{x_2} = \frac{5+10+1}{100}$

$SNP_{x_3} = \frac{14+2+0}{100}$

$\therefore SNP_{x_1} = 0.16$

$SNP_{x_2} = 0.16$

$SNP_{x_3} = 0.16$

$\boxed{\therefore \text{Answer} = 0.16}$ $\left(\because \frac{3 \times 0.16}{3} = 0.16\right)$

(b) $y = g(\beta_1 x_i^0 + \beta_2 x_i^1 + \beta_3 x_i^2)$ where $g(.)$ is sigmoid function.

The decision boundary is the straight line

$$L \equiv \beta_1 x^0 + \beta_2 x^1 + \beta_3 x^2 .$$

$\therefore$ For SNP_{x_1} : $\beta_1 > 0; \beta_2 > 0; \beta_3 > 0$

SNP_{x_2} ; $\beta_1 > 0 ; \beta_2 > 0 ; \beta_3 > 0$

SNP_{x_3} : $\beta_1 > 0 ; \beta_2 > 0 ; \beta_3 = 0$