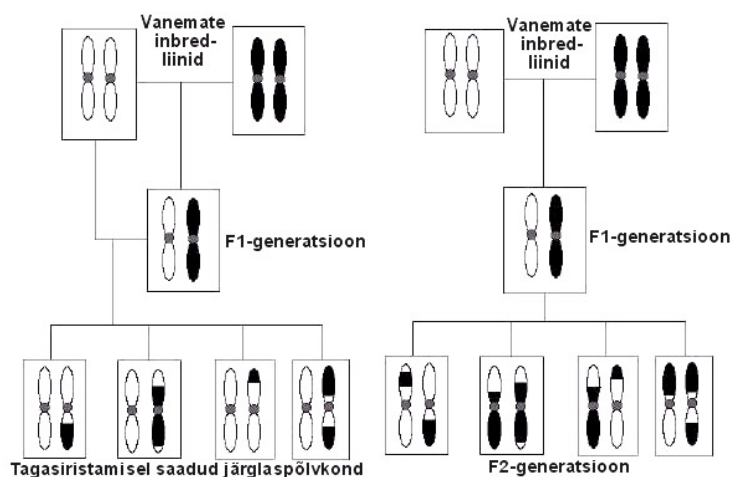


3.3 Põhigeeni tuvastamine eksperimentaalsel ristamisel

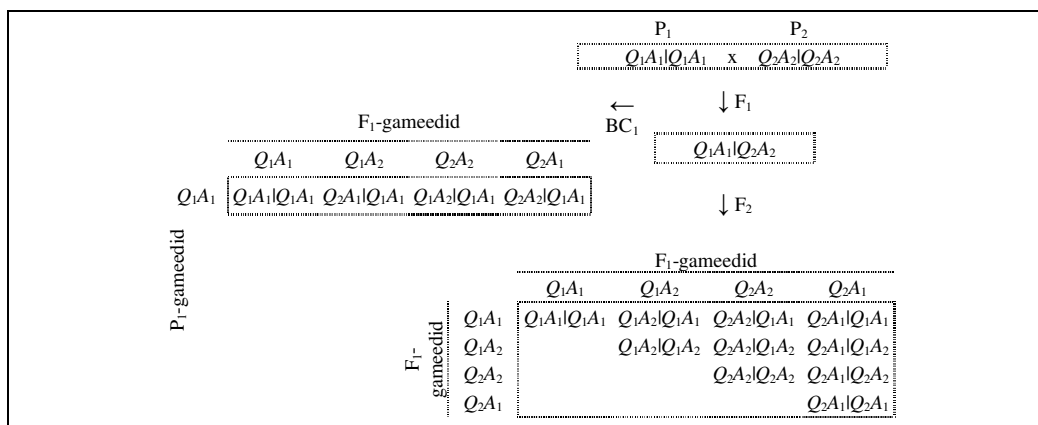
3.3.1 Marker-gruppide vahelise erinevuse testimine dispersioonanalüüsiga

Enamus põhigeeni leidmiseks korraldatud eksperimente võtavad aluseks kaks inbredliini (st, et vanemateks valitakse täielikult homosügootsed isendid). Selliste vanemate ristamisel saame geneetiliselt identsed heterosügootsed järglased (põlvkond F₁). Olgu oletatavaks uuritavat tunnust märkimisväärselt mõjutavaks põhigeeni lookuseks e QTL-ks dialleelne lookus *Q* (alleelid vastavalt *Q*₁ ja *Q*₂) ning vaadeldud indiviididel määratud markerlookuseks dialleelne lookus *A* (alleelid vastavalt *A*₁ ja *A*₂). Inbredliinidest pärit homosügootsed vanemad omavad siis vastavalt genotüüpe *Q*₁*A*₁*Q*₁*A*₁ ja *Q*₂*A*₂*Q*₂*A*₂ ning nende ristamisel saadud F₁-põlvkonna järglased genotüüpe *Q*₁*A*₁*Q*₂*A*₂.

Selgitamaks, kas lookused *A* ja *Q* on aheldunud, analüüsitakse andmeid fenotüübi ja markeri(te) kohta kas põlvkonnas F₂ või tagasiristamise teel saadud järglastel. Mõlemad need ristamisskeemid on kujutatud skemaatiliselt joonistel 16 ja 17.



Joonis 16. Tagasiristamisega ja tavaline ristamisskeem (joonistel on kujutatud vaid üht homologsete kromosoomide paari ja 4 analüüsitava järglast – reaalsetes uuringutes on vaatluse all kogu genotüüp ja üle 100 järglase).



Joonis 17. Kahe, markerlookuse *A* ja QTL-i *Q* suhtes homosügootse, vanemliini ristamine – F₁-põlvkonnas on kõik indiviidid geneetiliselt identsed heterosügootid genotüübiga *Q*₁*A*₁*Q*₂*A*₂, F₂-põlvkonnas leidub 10 erinevat ja tagasiristamisel saadud BC₁-generatsioonis 4 erinevat genotüüpi.

Tagasiristamise (*backcrossing*) korral ristatakse F₁-indiviidid ühega vanemliinidest. Selle tagajärjel on järglaspõlvkonna üks kromosoomidest identne ristamisel kasutatud inbredliinist pärit vanema kromosoomiga, teine kromosoom on aga tänu ristsiidlele kahe inbredliini kombinatsioon.

Tavalise ristamisskeemi (*intercrossing*) korral ristatakse omavahel põlvkonna F_1 isendeid. Generatsiooni F_2 indiviidide genotüüp on kombinatsioon oma vanemate genotüüpidest, kusjuures iga üksiku genotüübi esinemise sagedus sõltub QTL-i ja markerlookuse vahelisest rekombinatsioonimäärast r .

Järgevalt tuleb teha konkreetseid teste vaid tagasiristamisega eksperimendi tarvis, sest selle puhul on võrreldavaid kombinatsioone vähem ja valemite kirjapanek lihtsam.

Olgu meil ristatud indiviidid inbredselt genotüübiga $Q_1A_1|Q_1A_1$ ja F_1 -generatsioonist genotüübiga $Q_1A_1|Q_2A_2$. Kõik järglaspõlvkonna isendid pärivad tõenäosusega 1 gameedi Q_1A_1 . Generatsioonist F_1 pärit vanemalt saadakse alleelid Q_1A_1 tõenäosusega $\frac{1}{2}(1-r)$, alleelid Q_1A_2 tõenäosusega $r/2$, alleelid Q_2A_1 tõenäosusega $r/2$ ja alleelid Q_2A_2 tõenäosusega $\frac{1}{2}(1-r)$. Tagasiristamisel saadud indiviidid võib jagada kahte gruppi: marker-genotüübiga $A_1|A_1$ ja marker-genotüübiga $A_1|A_2$. Neist esimesed omavad QTL-genotüüpi $Q_1|Q_1$ tõenäosusega $1-r$ ja QTL-genotüüpi $Q_1|Q_2$ tõenäosusega r , teised analoogselt genotüüpi $Q_1|Q_1$ tõenäosusega r ja genotüüpi $Q_1|Q_2$ tõenäosusega $1-r$. Marker-genotüübiga $A_1|A_1$ indiviidide oodatav fenotüübiväärtus avaldub siis seosena

$$\mu_{A_1|A_1} = (1-r)\mu_{Q_1|Q_1} + r\mu_{Q_1|Q_2}$$

kus $\mu_{Q_1|Q_1}$ märgib uuritava tunnuse keskvaartust indiviididel QTL-genotüübiga $Q_1|Q_1$ ja $\mu_{Q_1|Q_2}$ indiviididel QTL-genotüübiga $Q_1|Q_2$. Samaselt on marker-genotüübiga $A_1|A_2$ indiviidide oodatav fenotüübiväärtus esitatav seosena

$$\mu_{A_1|A_2} = r\mu_{Q_1|Q_1} + (1-r)\mu_{Q_1|Q_2}.$$

Nende oodatavate fenotüübiväärtuste erinevus on

$$\mu_{A_1|A_1} - \mu_{A_1|A_2} = \underbrace{(\mu_{Q_1|Q_1} - \mu_{Q_1|Q_2})}_{\delta}(1-2r) = \delta(1-2r). \quad (3.14)$$

Kui QTL ja marker ei ole aheldunud, siis $r = \frac{1}{2}$ ja $\mu_{A_1|A_1} - \mu_{A_1|A_2} = 0$. Sellest lähtuvalt on loomulik marker-gruppide võrdlemisel kontrollida hüpoteeside paari $H_0: \mu_{A_1|A_1} - \mu_{A_1|A_2} = 0$, $H_1: \mu_{A_1|A_1} - \mu_{A_1|A_2} \neq 0$. Ligikaudu sobib toodud hüpoteeside kontrollimiseks tavaline t-test (ligikaudu seepärast, et uuritava tunnuse jaotuseks võrreldavais gruppides on segujaotus). Üldjuhul kasutatakse marker-gruppide võrdlemiseks siiski dispersioonanalüüsi (siit ka kirjeldatud meetodi populaarsem nimetus – **ANOVA-** e **dispersioonanalüüsi meetod**). Samas ei ole saadud tulemused ühesed, sest mitmed δ ja r väärtused annavad tulemuseks sama vahe $\mu_{A_1|A_1} - \mu_{A_1|A_2}$.

Leidmaks selliselt konstrueeritud testi võimsust, eeldame, et QTL-genotüübiga $Q_1|Q_1$ indiviidide fenotüübijaotus on normaaljaotus parameetritega $\mu_{Q_1|Q_1}$ ja σ^2 ning genotüübiga $Q_1|Q_2$ indiviidide fenotüübijaotus normaaljaotus parameetritega $\mu_{Q_1|Q_2}$ ja σ^2 . Tähistame marker-genotüübiga $A_1|A_1$ indiviidi j fenotüübiväärtust y_{1j} ja marker-genotüübiga $A_1|A_2$ indiviidi j fenotüübiväärtust y_{2j} . Esimese marker-grupi sisest varieeruvust näitav dispersioon on siis esitatav QTL-genotüüpidega $Q_1|Q_1$ ja $Q_1|Q_2$ isendite gruppide sisese ja vahelise varieeruvuse summana:

$$\text{Var}(y_{1j}) = E[\text{Var}(y_{1j} | \text{QTL-genotüüp})] + \text{Var}[E(y_{1j} | \text{QTL-genotüüp})]. \quad (3.15)$$

Selle summa esimene liidetav (QTL-gruppide sisene varieeruvus) on

$$E[\text{Var}(y_{1j} | \text{QTL-genotüüp})] = \sigma^2(1-r) + \sigma^2r = \sigma^2$$

ja teine liidetav (QTL-gruppide vaheline varieeruvus)

$$\begin{aligned} & \text{Var}[E(y_{1j} | \text{QTL-genotüüp})] \\ &= E[E(y_{1j} | \text{QTL-genotüüp})^2] - \{E[E(y_{1j} | \text{QTL-genotüüp})]\}^2 \\ &= E[(1-r)\mu_{Q_1|Q_1}^2 + r\mu_{Q_1|Q_2}^2] - [(1-r)\mu_{Q_1|Q_1} + r\mu_{Q_1|Q_2}]^2 \\ &= (1-r)r\mu_{Q_1|Q_1}^2 + (1-r)r\mu_{Q_1|Q_2}^2 - (1-r)2r\mu_{Q_1|Q_1}\mu_{Q_1|Q_2} \\ &= (\mu_{Q_1|Q_1}^2 - 2\mu_{Q_1|Q_1}\mu_{Q_1|Q_2} + \mu_{Q_1|Q_2}^2)(1-r)r = (\mu_{Q_1|Q_1} - \mu_{Q_1|Q_2})^2(1-r)r = \delta^2(1-r)r. \end{aligned}$$

Seega on dispersioon (3.15) avaldatav kujul

$$\text{Var}(y_{1j}) = \sigma^2 + \delta^2(1-r)r. \quad (3.16)$$

Analoogselt on esitatav ka marker-grupi $A_1|A_2$ sisene dispersioon. Marker-gruppide keskmiste vahelise erinevuse jaotus on siis ligikaudu normaaljaotus:

$$\bar{y}_1 - \bar{y}_2 \sim N(\delta(1-2r), 2[\sigma^2 + \delta^2(1-r)r]/n),$$

kus n on indiviidide arv igas grupis. Juhul, kui n on suur, on t-jaotus lähendatav normaaljaotusega ja testi võimsuse leidmiseks võime rakendada normaaljaotusele vastavat teststatistikut

$$Z = \frac{(\bar{y}_1 - \bar{y}_2)}{\sqrt{2[\sigma^2 + \delta^2(1-r)r]/n}}.$$

Testi võimsus on leitav kui tõenäosus $P(Z > z_{\alpha/2})$, kus z_{α} on normaaljaotuse α -täiendkvantiil, st $P(Z > z_{\alpha}) = \alpha$. Sisuka hüpoteesi kehtimise korral on statistiku keskväärtus esitatav seosena

$$E(Z) = \frac{\delta(1-2r)}{\sqrt{2[\sigma^2 + \delta^2(1-r)r]/n}}. \quad (3.17)$$

Testi võimsus on siis esitatav tõenäosusena

$$P[Z - E(Z) > z_{\alpha/2} - E(Z)].$$

Kuna vahe $Z - E(Z)$ on standardse normaaljaotusega, siis peab võimsuse $1 - \beta$ saavutamiseks kehtima võrdus $z_{\alpha/2} - E(Z) = z_{1-\beta} = -z_{\beta}$. Asendades viimases võrduses $E(Z)$ suhtega (3.17) ja lahendades saadud võrrandi n suhtes, saame võimsuse $1 - \beta$ jaoks vajalikud gruppide suurused:

$$n = (z_{\alpha/2} + z_{\beta})^2 \frac{2[\sigma^2 + \delta^2(1-r)r]}{\delta^2(1-2r)^2}. \quad (3.18)$$

Näide. Olgu meil eksperiment läbi viidud sarnaselt joonisel 17 kujutatud skeemile. Geenide aditiivne efekt põlvkonnas F_2 on kirjeldatav aditiivdispersiooni σ_a^2 abil. Tähistagu p QTL-i osa generatsiooni F_2 aditiivgeneetilises varieeruvuses. Märkus: vanempõlvkonnad (P_1 ja P_2) on täielikult homosügootsed, põlvkonnas F_2 on aga juba 50% isendeid heterosügootid – seega on QTL-st põhjustatud varieeruvus põlvkonnas F_2 poole suurem võrreldes vanempõlvkondadega. QTL-st tingitud osa aditiivgeneetilisest varieeruvusest avaldub siis suhtena

$$p = \frac{1}{2} \delta^2 / \sigma_a^2,$$

kus teatavasti $\delta = \mu_{Q_1Q_1} - \mu_{Q_1Q_2}$. Siis

$$\sigma_a^2 = \frac{1}{2} \delta^2 / p.$$

Uuritava tunnuse päritavus h^2 generatsioonis F_2 on

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_e^2},$$

kus $\sigma_e^2 = \sigma_a^2(1-h^2)/h^2$ on keskkonna mõjudest tingitud varieeruvus. Aditiivgeneetiline varieeruvus tagasiristamisel saadud isendite vahel on poole väiksem võrreldes F_2 -generatsiooniga. Seega, kuna geenide aditiivsest efektist tingitud varieeruvus on QTL-i olemasolul vastavalt väiksem: $\sigma_a^2 - p \times \sigma_a^2 = \sigma_a^2 - \frac{1}{2} \delta^2$, on koguvareeruvus põlvkonnas BC_1 esitatav summana

$$\sigma^2 = \sigma_e^2 + \frac{1}{2}(\sigma_a^2 - \frac{1}{2} \delta^2).$$

Asendades σ_e^2 avaldisega $\sigma_e^2 = \sigma_a^2(1-h^2)/h^2$, saame viimatisele dispersioonile kuju

$$\sigma^2 = \frac{1}{2} \delta^2 \frac{(1-h^2)}{ph^2} + \frac{1}{4} \delta^2 \left(\frac{1}{p} - 1 \right) = \delta^2 \left[\frac{1}{2} \frac{(1-h^2)}{ph^2} + \frac{1}{4} \left(\frac{1}{p} - 1 \right) \right].$$

Võrduse (3.18) saame nüüd oma näite jaoks esitada seosena

$$n = 2 \frac{(z_{\alpha/2} + z_{\beta})^2}{(1-2r)^2} \left[\frac{1}{2} \frac{(1-h^2)}{ph^2} + \frac{1}{4} \left(\frac{1}{p} - 1 \right) + r(1-r) \right].$$

Tagasiristamisega eksperimentis tunnusele päritavusega $h^2 = 0,25$ põhigeeni tuvastamise võimsuse 0,9 saavutamiseks vajalikud valimi mahud on toodud järgnevas tabelis.

r	P		
	0,04	0,08	0,16
0	1828	909	449
0,05	2260	1125	557
0,1	2863	1426	708
0,2	5097	2543	1266
0,4	45959	22974	11482

3.2.2 Suurima tõepära meetod

Nagu oli näha valemist (3.14), pole marker-gruppide fenotüübierinevuste alusel QTL-efekti võimalik üheselt hinnata. Üks võimalus probleemi lahendamiseks on kirjutada välja andmetele vastav tõepärafunktsioon ja maksimeerida see paralleelselt nii oletatavale QTL-genotüübile vastavate jaotusparameetrite kui ka rekombinatsioonimäära suhtes.

Vaatleme lihtsuse mõttes jällegi tagasiristamisega skeemi ning otsime dialleelset QTL-i. Eeldame, et QTL-genotüübiga $Q_1|Q_1$ indiviidide fenotüübijaotus on normaaljaotus parameetritega $\mu_{Q_1|Q_1}$ ja σ^2 ning QTL-genotüübiga $Q_1|Q_2$ indiviidide fenotüübijaotus normaaljaotus parameetritega $\mu_{Q_1|Q_2}$ ja σ^2 . Siis on nende indiviidide, kes on markerlookuse suhtes homosügootid $A_1|A_1$, fenotüübijaotus esitatav kahe normaaljaotuse seguna:

$$f_1(y; \mu_{Q_1}, \mu_{Q_2}, \sigma, r) = (1-r) \times \phi((y - \mu_{Q_1})/\sigma) + r \times \phi((y - \mu_{Q_2})/\sigma),$$

kus r märgib markerlookuse ja QTL-i vahelist rekombinatsioonimäära ja ϕ tähistab standardse normaaljaotuse tihedusfunktsiooni. Analoogselt on markerlookuse suhtes heterosügootsete $A_1|A_2$ indiviidide fenotüübijaotus esitatav seguna:

$$f_2(y; \mu_{Q_1}, \mu_{Q_2}, \sigma, r) = (1-r) \times \phi((y - \mu_{Q_1})/\sigma) + r \times \phi((y - \mu_{Q_2})/\sigma).$$

Tähistades indiviidi i fenotüübi y_i ja tuues sisse muutuja x_i , mis võrdub 1-ga, kui i indiviid on markerlookuse suhtes heterosügootne, ja 0-ga markerlookuse suhtes homosügootse indiviidi korral, on andmetele vastav tõepärafunktsioon avaldatav korrutisena

$$L(\theta, r; y, x) = \prod_i [f_1(y_i; \theta, r)]^{1-x_i} [f_2(y_i; \theta, r)]^{x_i},$$

kus $\theta = (\mu_{Q_1}, \mu_{Q_2}, \sigma)$.

Saadud tõepärafunktsioon maksimeeritakse esmalt θ suhtes, andes selleks rekombinatsioonimäärale r mingi esialgse väärtuse; seejärel, kasutades leitud suurima tõepära hinnangut $\hat{\theta}$, avaldatakse tõepärafunktsiooni maksimum parameetri r suhtes ja lahendatakse, saades tulemuseks suurima tõepära hinnangu \hat{r} ; leitud rekombinatsioonimäära hinnangut kasutatakse omakorda täpsustamaks parameetri θ hinnangut, jne kuni protseduur koondub.

Hüpoteesi markerlookuse ja QTL-i ahelduse kohta kontrollitakse tõepärasuhte testi abil, võrreldes ülaltoodud ühe oletatava põhigeeni ja markeri ahelduse eeldusel konstrueeritud mudelit nullhüpoteesile vastava mudeliga, kus kõigi indiviidide fenotüübiväärtused eeldatakse jaotuvat vastavalt normaaljaotusele parameetritega μ ja σ ning tõepärafunktsioon avaldub korrutisena

$$L_0(\theta_0; y) = \prod_i \phi((y_i - \mu)/\sigma),$$

kus $\theta_0 = (\mu, \sigma)$. Seega on *LOD-score* statistik, mille täisarvulised väärtused 1, 2, 3 jne vastavad "ümmargustele" tõepärasuhtetele 10, 100, 1000 jne, kujul

$$LOD(r) = \log_{10} \left[\frac{\max L(\theta, r; y, x)}{\max L_0(\theta_0; y)} \right]. \quad (3.19)$$

3.2.3 Regressioonanalüüs

Oletatava põhigeeni lokaliseerimiseks eksperimentaalsel ristamisel vahest kõige enam kasutatav meetod seisneb lühidalt järgnevas: fikseeritakse oletatava põhigeeni asukoht markerite vahel ning püütakse seejärel lähendada fenotüübil mõõdetud tunnuse väärtusi oletatavatesse QTL-i gruppidesse kuulumise tõenäosustega, hinnates selleks regressioonanalüüsi abil erinevatele QTL-genotüüpidele vastavaid oodatavaid fenotüübiväärtuseid. Tehes sama läbi kõigi markerite vaheliste piirkondade kohta, saame tulemuseks kõige tõenäolisema põhigeeni asukoha – intervalli, kus regressioonanalüüsi jääkide ruutude summa on minimaalne.

Võtame vaatluse alla kaks dialleelset markerlookust A ja B ning oletatava dialleelse QTL-i Q nende vahel (selliste, oletatavat QTL-i sisaldava kromosoomiintervalli otstes paiknevate, markerite tarvis on ingliskeelses kirjanduses kasutusel spetsiaalne nimetus – *flanking markers* – **piiravad markerid**).

Individid inbredliinidest P_1 ja P_2 on vastavalt genotüüpidega $A_1Q_1B_1|A_1Q_1B_1$ ja $A_2Q_2B_2|A_2Q_2B_2$. Tähistame rekombinatsioonimäära lookuste A ja Q vahel r_A , lookuste Q ja B vahel r_B ning lookuste A ja B vahel r_{AB} , seejuures eeldatakse viimane teadaolevaks, st et toetutakse juba teadaolevale geneetilisele kaardile. Et rekombinatsioon intervallis A -st B -ni on fikseeritav vaid juhul, kui see leidis aset kas A -st Q -ni või Q -st B -ni, aga mitte mõlemal juhul korraga, siis kehtib järgmine rekombinatsioonimäärade vaheline seos: $r_{AB} = r_A + r_B - 2r_A r_B$.

Tähistame QTL-genotüüpidele vastavad oodatavad fenotüübiväärtused $E(y|Q_1Q_1) = \mu_{Q_1Q_1}$ ja $E(y|Q_1Q_2) = \mu_{Q_1Q_2}$ ning eeldame ka uuritava tunnuse võrdset varieeruvust mõlema genotüübi korral.

Generatsioonis F_1 on kõik indiviidid genotüübiga $A_1Q_1B_1|A_2Q_2B_2$, nende ristamisel vanemliiniga P_1 saadavas generatsioonis BC_1 on võimalik leida neli erinevat marker- ja kaks QTL-genotüüpi. Iga järglane omab ühes kromosoomis P_1 -vanemalt saadud alleelikomplekti $A_1Q_1B_1$, teises kromosoomis paiknev F_1 -vanemalt päritud alleelikomplekt sõltub juhusest ja lookustevahelistest võimalikest ristsiretest. Oletades, et järglane pärib teisena markeralleelid A_2 ja B_1 , saab ta tõenäosusega $r_A(1-r_B)$ ka alleeli Q_1 – selleks peab toimuma rekombinatsioon markerlookuse A ja QTL-i Q vahel (tõenäosus r_A) ja ei tohi toimuda rekombinatsiooni vahemikus Q -st B -ni (tõenäosus $1-r_B$). Tingliku tõenäosusena esitatuna on seos kujul

$$P\left(\frac{A_1Q_1B_1}{A_2Q_1B_1} \mid \frac{A_1Q_1B_1}{A_2B_1}\right) = \frac{r_A(1-r_B)}{r_A(1-r_B) + (1-r_A)r_B} = \frac{r_A(1-r_B)}{r_{AB}}$$

Analoogselt on leitav ka tõenäosus, et järglane pärib lisaks generatsiooni P_1 -vanemalt saadud QTL-alleelile Q_1 generatsiooni F_1 -vanemalt koos markeralleelidega A_2 ja B_1 QTL-alleeli Q_2 :

$$P\left(\frac{A_1Q_1B_1}{A_2Q_2B_1} \mid \frac{A_1Q_1B_1}{A_2B_1}\right) = \frac{(1-r_A)r_B}{r_A(1-r_B) + (1-r_A)r_B} = \frac{(1-r_A)r_B}{r_{AB}}$$

Samamoodi on tinglikud tõenäosused leitavad kõigi nelja võimaliku markergenotüübi ja kahe QTL-genotüübi korral (seda muidugi eeldades interferentsuse puudumist – vastasel juhul ei saaks markerite A ja B vahelist rekombinatsioonimäära esitada markerlookuste ja QTL-i vaheliste rekombinatsioonide toimumise ja/või mittetoimumise tõenäosuste korrutisena) – vt tabel 2.

Baseerudes leitud tõenäosustel püütakse uuritava tunnuse fenotüübiväärtusi lähendada mudeliga

$$y = \mathbf{X}\boldsymbol{\beta} + e,$$

kus y on fenotüübiväärtuste $n \times 1$ -vektor, \mathbf{X} on tabelis 2 toodud tõenäosuste $n \times 2$ -maatriks, $\boldsymbol{\beta} = (\mu_{Q_1Q_1}, \mu_{Q_1Q_2})^T$ on tundmatute QTL-genotüüpide ooteväärtuste 2×1 -vektor, e on juhuslike vigade $n \times 1$ -vektor.

Marker-genotüüp	QTL-genotüüp	
	$Q_1 Q_1$	$Q_1 Q_2$
$A_1B_1 A_1B_1$	$\frac{(1-r_A)(1-r_B)}{(1-r_{AB})}$	$\frac{r_A r_B}{(1-r_{AB})}$
$A_1B_1 A_1B_2$	$\frac{(1-r_A)r_B}{r_{AB}}$	$\frac{r_A(1-r_B)}{r_{AB}}$
$A_1B_1 A_2B_1$	$\frac{r_A(1-r_B)}{r_{AB}}$	$\frac{(1-r_A)r_B}{r_{AB}}$
$A_1B_1 A_2B_2$	$\frac{r_A r_B}{(1-r_{AB})}$	$\frac{(1-r_A)(1-r_B)}{(1-r_{AB})}$

Tabel 2. QTL-genotüüpide tinglikud tõenäosused markergenotüüpide suhtes tagasiristamisega saadud generatsioonis.

Fikseeritud markerlookuste A ja B ning nendevahelise oletatava QTL-i tarvis leitakse parameetervectori $\boldsymbol{\beta}$ vähimruutude hinnang

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$$

ning jääkide ruutude summa $SSE = \mathbf{y}^T\mathbf{y} - \hat{\boldsymbol{\beta}}^T\mathbf{X}^T\mathbf{y}$.

Teststatistikuna kasutatakse seosega (3.19) defineeritud tõepärasuhte statistikule lähedast statistikut, mille tähistuski on enamasti sama:

$$LOD(r) = \frac{n}{2} \log_{10} \left[\frac{SSE_0}{SSE} \right], \tag{3.20}$$

kus n on analüüsitud indiviidide arv, SSE_0 tähistab jääkide ruutude summat ilma QTL-efektita mudelist ja SSE sama regressiooniseose täpsust iseloomustavat suurust põhigeeni eeldusel konstrueeritud mudelist.