



## تحليل أنموذج قياس الصفات الكمية لتشخيص تأثير الموروثات الجينية باستعمال طريقة الامكان الأعظم

### Analysis of a Quantitative Trait Measurement Model to Diagnose the Effect of Genetic Inheritance using the Max Likelihood Method

ا. د. دكتور عباس لفته كنهير<sup>(2)</sup>

حميدة خضر ابراهيم<sup>(1)</sup>

#### المستخلص : (Abstract)

في هذه الدراسة تم دراسة وتحليل علاقة التأثير المتبادل للمورثات الجينية وإن موضع السمات الكمية (Quantitative Trait Loci Analysis QTL) هو موضع (قسم من الحمض النووي) يرتبط باختلاف سمة الكمية في النمط الظاهري لمجموعة من الكائنات الحية . و يعد تحليل مواقع الصفات الكمية (Quantitative Trait Loci Analysis (QTL) طريقة إحصائية للقياس الكمي للصفات التي تربط نوعين من المعلومات الصفات الظاهرية (Phenotypes) مع الصفات الجينية (Genotypes) التي تستخدم عادة لتحليل الموقع الجيني في مناطق محددة من الكروموسومات التي تتأثر بصورة طبيعية بجين واحد أو أكثر أو بطبيعة البيئة ( المحيط الخارجي) ويسمى هذا بالتحليل الإحصائي الجيني<sup>(1)</sup> (Genetic statistical Analysis) .

تم استخدام دالة الامكان الاعظم لتقدير جزء اعادة التركيب r في الكروموسوم التحليل الوراثي Genetic Analysis وفيما بعد تم استخدام اختبارات التباديل Permutation Tests لغرض تقييم فرضية الاختبار. تم عمل اختبار التباديل من خلال تنفيذ (5000) عينة تبادلية ومن ثم حساب الإحصاء (  $-2\log \lambda$  ) لكل عينة ومن ثم تحديد القيم الحرجة باستعمال الأشكال البيانية لتوزيعات تباديل العينات .

#### **Abstract:**

A locus (section of DNA) known as a quantitative trait locus (QTL) corresponds with quantitative trait variance in an organism group's phenotype. It is common practice to utilize quantitative trait loci analysis (QTL), a statistical technique for the quantitative measurement of traits certain chromosomal areas that are naturally influenced by a gene. Genetic statistical analysis is the study of one or more environmental factors, specifically the external environment .

Later on, permutation tests were employed to assess the test hypothesis. In order to perform a permutation test, 5000 exchange samples were used. After computing the statistic  $(-2\log \lambda)$  for each sample, the critical values were then calculated using the graphical representations of the sample permutation distributions.

## 1. المقدمة introduction

إن موضع السمات الكمية (Quantitative Trait Loci Analysis QTL) هو موضع (قسم من الحمض النووي) يرتبط باختلاف سمة كمية في النمط الظاهري لمجموعة من الكائنات الحية التي يتم تعيينها من خلال تحديد العلامات الجزيئية مثل تعدد أشكال النوكليوتيدات المفردة (SNP) (Amplified fragment length polymorphism) أو تعدد الأشكال لطول الجزء المضخم (AFLP) (Amplified Fragment Length Polymorphism) المرتبطة بسمة ملحوظة في الفرد . غالبًا ما تكون خطوة تحديد الجينات الفعلية خطوة مبكرة تسبب اختلاف السمات وتسلسلها.<sup>(1)</sup>

ويمكن تعريف الوراثة بصورة عامة : هو العلم الذي يدرس كل ما يتعلق بالمواد الحية التي تتشكل بين أجيال الكائنات الحية . بعبارة أخرى هو العلم الذي يهتم في أوجه التشابه والاختلاف في الصفات بين الأفراد الذين تربطهم صلة القرابة [9]

## 2. مشكلة البحث The research problem

إن دراسة مواقع الصفات الكمية تهدف لمعرفة تأثير الكروموسوم في نقل الصفات الوراثية وتحسينها، وكذلك معرفة تأثير الأمراض وانتقالها إلى الأجيال القادمة عبر الجينات وتستخدم طريقة الإمكان الأعظم (MLE) لتقدير معالم (QTL) للحصول على تقديرات معالم (QTL) للنموذج الجيني ومعرفة أفضل تقدير .

## 3. هدف البحث Purpose of search

يهدف البحث الى دراسة وتحليل تأثير أثر مواقع الصفات الكمية للموروثات الجينية عبر الأجيال المتعاقبة من خلال صياغة أنموذج إحصائي باستعمال طريقة الامكان الاعظم وتمثيل هذا التأثير والذي يتمثل بتأثير الصفات الظاهرية (Phenotypes) والتأثيرات الجينية (Genetic types) ومن ثم إخضاعها للقوانين الإحصائية بغية التحكم أو التنسيق بهذه التأثيرات.

## 4. فرضية البحث

الكشف عن الترابط بين الواسمات والصفات المظهرية من خلال معامل اعادة التركيب الجزيئي (r) فاذا كان  $r = 0.5$  فهذا يعني عدم الترابط بين الواسمات والصفات المظهرية الكمية اي ان متوسطي المجموعتين من الصفات المظهرية يكونان متساويين .

## 5. الجانب النظري

### 1.5 المقدمة

يعد تحليل مواقع الصفات الكمية (Quantitative Trait Loci Analysis (QTL) طريقة إحصائية للقياس الكمي للصفات التي تربط نوعين من المعلومات الصفات الظاهرية (Phenotypes) مع الصفات الجينية (Genetic types) والتي تستخدم عادة لتحليل الموقع الجيني في مناطق محددة من الكروموسومات التي تتأثر بصورة طبيعية بجين واحد أو

أكثر أو بطبيعة البيئة ( المحيط الخارجي) ويسمى هذا بالتحليل الإحصائي الجيني (Genetic statistical Analysis) (1)

وأن مصطلح Quantitative Trait Loci (QTL) يمكن توضيحه بالآتي :

- Quantitative التي يرمز لها بـ ( Q ) تشير إلى التحقيق التجريبي المنهجي للظواهر عن طريق الإحصاء باستخدام بيانات كمية وتقنيات حسابية على هذه البيانات .
- Trait يرمز لها بـ ( T) ويمكن تعريفها بأنها سمة الربط الجيني عندما تكون المورثات محكومة بجينين أو أكثر وتفاعله مع البيئة أو هي الصفة التي تمتلك توزيع مستمر من النوع الظاهري بالعديد من الجينات الداخلة في البيئة.
- Loci يرمز له بالرمز (L) يمثل منطقة من جين واحد أو عدد من الجينات أو هي موقع تأثير السمات للمنطقة البسيطة من الجين أو أي جزء من الجين المفرد إلى أكبر جزء من الكروموسوم والذي يسمى الموقع (2).

## 2.5 التحليل الوراثي Genetic Analysis

يمكن أن نستنتج وجود معدل ضمنى لصفة الكمية عن طريق اختبار t بسيط أو جدول تحليل التباين ANOVA على علامة معينة في الـ QTL المرتبطة بعلامة مع جزء إعادة التركيب r . إن قيم التوقع الشرطي لتأثير الجيني والمرتبطة بكل نمط جيني تحسب من الاحتمالات الشرطية لأنماط الجينية لـ QTL مع إعطاء علامة النمط الجيني من مختلف الأنماط الجينية لـ QTL. ويمكن ان نعرف علامة التأثير الجيني لـ Aa (1) و aa (0) التي يمكن أن نشق الاحتمالات الشرطية لاثنتين من انماط التأثير الجيني qq (0) , Qq (1) لـ QTL ويمكن أن نكون الجدول (1) التضارب العكسي لقيم التأثير الجيني وكالاتي (3) :

### الجدول (1)

التضارب العكسي لقيم التأثير الجيني لـ QTL

Marker Genotypic Value	QTL Genotypic Value	
	$\mu_1$	$\mu_0$
$m_1$	$(1 - r)$	$r$
$m_0$	$r$	$(1 - r)$

إن قيم التأثير الجيني لاثنتين من انماط تأثير التضارب العكسي لـ QTL يمكن أن تمثل بـ  $\mu_1 = \mu + \frac{1}{2}a$  و  $\mu_0 = \mu - \frac{1}{2}a$  على التوالي .

وان لكل نمط وراثي علامة يتم خلط نمطين جينيين مختلفين ويمكن ترجيحها بالاحتمالات الشرطية وهكذا فان التوقع الشرطي لقيم الاحتمالات الجينية المرتبطة بعلامتين جينيتين مختلفتين يمكن حسابها وكالاتي :

$$m_1 = (1 - r)\mu_1 + r\mu_0 = (1 - r)\left(\mu + \frac{1}{2}a\right) + r\left(\mu - \frac{1}{2}a\right) = \mu + \frac{1}{2}(1 - 2r)a$$

$$m_0 = r\mu_1 + (1 - r)\mu_0 = r\left(\mu + \frac{1}{2}a\right) + (1 - r)\left(\mu - \frac{1}{2}a\right) = \mu - \frac{1}{2}(1 - 2r)a$$

لذلك فان اختلاف متوسطي العلامتين يعطى بالشكل الاتي :

$$\mu_1 - \mu_0 = (1 - 2r)a \quad \dots \quad (1)$$

الجدول (2)

التضارب العكسي لقيم التأثير الجيني لـ QTL

Marker Genotypic Value	QTL Genotypic Value	
	$\mu_1$	$\mu_0$
$m_1$	$(1 - r)$	$r$
$m_0$	$r$	$(1 - r)$

وان فرضية الاختبار تكون كالآتي :

$$H_0: \mu_1 = \mu_0$$

$$H_1: \mu_1 \neq \mu_0$$

وإحصاء الاختبار تكون بالشكل :

$$t = \frac{m_1 - m_0}{\sqrt{s^2 \frac{1}{n_1} + \frac{1}{n_2}}} \quad \dots \quad (2)$$

إذا كانت  $a$  غير معنوية ومختلفة عن الصفر، بناءً على إحصاء الاختبار  $t$  بالنسبة إلى المعادلة (2) عندما تكون إحصاء الاختبار أصغر من القيمة المحسوبة ، وبهذا المعنى يمكن أن يوفر اختبار  $t$  معلومات أكثر أهمية بالنسبة إلى تأثير QTL. ولكن قيمة  $t$  تكون غير معنوية بسبب عدم وجود ارتباط بين العلامتين وأن قيمة  $r$  تفترض أنها تساوي 0.5 بالنسبة إلى QTL في المعادلة (1) لذلك يعطي اختبار  $t$  اختباراً مركباً فقط لتأثير QTL وارتباط علامات الـ QTL. (3)

بالنسبة للمجتمع F2 يمكن اشتقاق ثلاث متوسطات علامة بالمثل مشابهة للاحتتمالات الشرطية والتي يمكن التعبير عنها في الجدول (3) كما يلي :

### الجدول (3)

القيم الجينية لمتوسطات العلامات

Marker Genotypic Value	QTL Genotypic Value		
	$\mu_0$	$\mu_1$	$\mu_2$
$m_2$	$r^2$	$2r(1-r)$	$(1-r)^2$
$m_1$	$r(1-r)$	$(1-r)^2 + r^2$	$r(1-r)$
$m_0$	$(1-r)^2$	$2r(1-r)$	$r^2$

وتفترض القيم الجينية لمتوسطات العلامات والتي يعبر عنها بالاتي :

$$m_2 = (1-r)^2\mu_2 + 2r(1-r)\mu_1 + r^2\mu_0$$

$$= \mu + (1-r)a + 2r(1-r)d \quad \dots (3)$$

$$m_1 = 2r(1-r)\mu_2 + [(1-r)^2 + r^2]\mu_1 + r(1-r)\mu_0$$

$$= \mu + (1-2r+r^2)d \quad \dots (4)$$

$$m_0 = (1-r)^2\mu_2 + 2r(1-r)\mu_1 + r^2\mu_0$$

$$= \mu - (1-2r)a + 2r(1-r)d \quad \dots (5)$$

إن إحصاء اختبار التأثير المضاف a والمجال d يكونا كالاتي :

$$t_1 = \frac{m_2 - m_0}{\sqrt{s^2 \frac{1}{n_2} + \frac{1}{n_0}}} \quad \dots (6)$$

$$t_2 = \frac{m_1 - \frac{1}{2}(m_2 - m_0)}{\sqrt{s^2 \frac{1}{4n_2} + \frac{1}{n_1} + \frac{1}{4n_0}}} \quad \dots (7)$$

إن اختبار التأثير المضاف a في المعادلة (6) وتأثير المجال d في المعادلة (7) يكون مكافئاً لاختبار المعالم المركبة والتي تكون بالشكل الاتي :

$$\frac{1}{2}(m_2 - m_0) = (1-2r)a = a - 2ra$$

$$m_1 - \frac{1}{2}(m_2 + m_0) = (1-2r)^2d = d - 4(1-r)d + d$$

والمعادلتان مساويتان للصفر .

من التحليل أعلاه على الرغم من أنه يمكن استخدام اختبار t و ANOVA لاختبار معنوية الاختلافات بين العلامات فإنه لا يمكن تمييز الاوساط الجينية لـ QTL وجزء اعادة التركيب بين علامة واحدة والـ QTL الا إذا كان الفرق بين العلامات معنوي ، فإننا لا نزال لا نعرف ما إذا كان هذا الاختلاف يرجع إلى ارتباط ضيق (يعني r صغيرة) بين العلامة

والـ QTL يكون التأثير صغيراً أو يكون ارتباط واسع (يعني  $r$  كبيرة) بين العلامة والـ QTL يكون التأثير كبيراً. في الواقع يتم التقليل من آثار التجميع والمجال لـ QTLs من  $2r$  الى  $4r(1 - r)$  على التوالي ومن مقارنة بسيطة من اوساط العلامات . أيضا أن اختبار  $t$  و ANOVA لا يمكن تمييز آثار الـ QTLs الفردية على النمط الظاهري إذا كان هناك اثنان أو أكثر QTL على نفس الكروموسوم .<sup>(4) (5)</sup>

### 3.5 دالة الامكان Likelihood

إنّ بيانات النمط الجيني المشاهدة سيكون إما  $M_1M_1$  أو  $M_1M_2$ ، ولكن بالنظر إلى ذلك النمط الجيني المشاهدة ، سيكون النمط الوراثي QTL إما  $Q_1Q_1$  أو  $Q_1Q_2$  وفقاً للاحتمالات المعطاة في الجدول (4) وكالاتي :

#### الجدول (4)

#### احتمالات النمط الوراثي لـ QTL

	$M_1M_1$	$M_1M_2$
$Q_1Q_1$	$\frac{1-r}{2}$	$\frac{r}{2}$
$Q_1Q_2$	$\frac{r}{2}$	$\frac{1-r}{2}$

إذا أن  $r$  هو جزء إعادة التركيب بين العلامة و QTL وافترض أن سمة النمط الظاهري ( $y$ ) تتبع التوزيع الطبيعي وأنه يمكن ربط تأثير الأنماط الظاهرية للسمة بالنسبة للأوساط المظهرية الخاصة بالأنماط الجينية للعلامة ، ويمكن اعطاء نموذج مختلط وكالاتي :

$$M_1M_1 \text{ المشاهدة } \rightarrow y \sim (1 - r)N(\mu_1, \sigma^2) + r N(\mu_2, \sigma^2)$$

$$M_1M_2 \text{ المشاهدة } \rightarrow y \sim r N(\mu_1, \sigma^2) + (1 - r)N(\mu_2, \sigma^2) \dots (8)$$

إذا أنه :

$\mu_1, \mu_2$  : تمثل الاوساط الحسابية لتأثير الظاهري للأنماط الجينية ( أو القيم الجينية ) لكمية الـ QTL لتأثير الجيني لـ  $Q_1Q_1$  و  $Q_1Q_2$  على التوالي . وبالنظر لنموذج (8) سيكون توزيع الوسط والتباين كالاتي :

$$\mu_{M_1M_1} = (1 - r)\mu_1 + r\mu_2$$

$$\mu_{M_1M_2} = r\mu_1 + (1 - r)\mu_2$$

$$\sigma^2_{M_1M_1} = \sigma^2_{M_1M_2} = \sigma^2 + r(1 - r)(\mu_1 - \mu_2)^2$$

ويمكن ملاحظة ان الارتباط بين العلامات والـ QTL إذا كان  $r = \frac{1}{2}$  وان  $\mu_{M_1M_1}$  تساوي  $\mu_{M_1M_2}$  فأن فرضية الاختبار يفرض عدم وجود ارتباط وكالاتي :

$$H_0 : r = \frac{1}{2} \text{ or } H_0 : \mu_{M_1M_1} - \mu_{M_1M_2} = (1 - 2r)(\mu_1 - \mu_2) = 0$$

من المهم ملاحظة أنه في ظل ظروف  $H_0$  لا يمكننا معرفة ما إذا كانت  $r = \frac{1}{2}$  أو  $\mu_1 = \mu_2$  في كلتا الحالتين ومع ذلك لا يمكن اكتشافه من الناحية العملية ايجاد التباين في النمط الظاهري. إذا افترضنا أن  $y_1, \dots, y_{n_1}$  من مجموعة العلامات  $M_1M_1$  وأن  $y_{n_1+1}, \dots, y_n$  من مجموعة العلامات  $M_1M_2$  ، فإن دالة الامكان معتمدة بناءً على النموذج . (8)

$$L(\mu_1, \mu_2, \sigma^2, r|y) = \prod_{i=1}^{n_1} (1-r) f(y_i|\mu_1, \sigma^2) + r f(y_i|\mu_2, \sigma^2) \times \prod_{i=1}^{n_1+1} r f(y_i|\mu_1, \sigma^2) + (1-r) f(y_i|\mu_2, \sigma^2) \dots (9)$$

لاختبار فرضية العدم التي تفرض عدم وجود ارتباط لـ QTL يعني  $H_0 = 0$ ، ويمكننا استخدام احتمال نسبة الامكان وكالاتي :

$$\lambda = \frac{\max_{\mu_1=\mu_{12}, \sigma^2, r} L(\mu_1, \mu_2, \sigma^2, r|y)}{\max_{\mu_1, \mu_2, \sigma^2, r} L(\mu_1, \mu_{12}, \sigma^2, r|y)} \dots (10)$$

سترفض فرضية العدم  $H_0$  إذا كانت إحصائية الاختبار  $\lambda$  صغيرة جدًا ، بدلاً من ذلك يمكننا تحويل الاختبار إلى  $-2\log \lambda$  ، والذي سيرفض إذا كان كبيرًا وغالبًا ما يقترب الى توزيع  $\chi^2$  ومع ذلك فإن نموذج المختلط يأخذ افتراضات توزيع  $\chi^2$  ولذلك يتم إجراء اختبار بديل لـ  $-2\log \lambda$  أو تغيير نتيجة درجة LOD<sup>(6)</sup>.

### 3.5.1 اختبار نسبة الامكان Likelihood Ratio Test

لحساب إحصائية الاختبار  $\lambda$  علينا تعظيم كل من البسط والمقام في المعادلة (10) ، إن تعظيم البسط يكون سهلاً لأنه تحت الفرضية العدم  $\mu = \mu_1 = \mu_{12} = H_0$  يصبح الاحتمال في المعادلة (9) كالاتي<sup>(7)(8)</sup>:

$$L(\mu_1, \mu_{12}, \sigma^2, r|y) = L(\mu, \sigma^2, r|y) - \prod_{i=1}^n f(y_i|\mu, \sigma^2) \dots (11)$$

إذا ان تقدير معالم MLE سيكون  $\hat{\mu} = \bar{y}$  وان  $\sigma^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2$

لتعظيم مقام المعادلة (10) ، نحتاج إلى تعظيم الاحتمال (9). وللقيام بذلك نشق بالنسبة للصفر ثم نحل.

$$\begin{aligned} \frac{\partial}{\partial \mu_1} \log L(\mu_1, \mu_2, \sigma^2, r|y) &= \frac{\partial}{\partial \mu_1} \sum_{i=1}^{n_1} \log((1-r)f(y_i|\mu_1, \sigma^2) + rf(y_i|\mu_2, \sigma^2)) + \\ &\frac{\partial}{\partial \mu_1} \sum_{i=n_1+1}^n \log(rf(y_i|\mu_1, \sigma^2) + (1-r)f(y_i|\mu_2, \sigma^2)) \\ &= \sum_{i=1}^{n_1} \frac{(1-r)\frac{\partial}{\partial \mu_1} f(y_i|\mu_1, \sigma^2)}{(1-r)f(y_i|\mu_1, \sigma^2) + rf(y_i|\mu_2, \sigma^2)} + \sum_{i=n_1+1}^n \frac{r\frac{\partial}{\partial \mu_1} f(y_i|\mu_1, \sigma^2)}{rf(y_i|\mu_1, \sigma^2) + (1-r)f(y_i|\mu_2, \sigma^2)} \dots (12) \end{aligned}$$

ويمكن الاشتقاق باستخدام المعادلة (13) وكالاتي :

$$\sigma_g^2 = \frac{MS_1 - MS_2}{k} \dots (13)$$

$$\begin{aligned} p_1(y_i) &= \frac{(1-r)f(y_i|\mu_1, \sigma^2)}{(1-r)f(y_i|\mu_1, \sigma^2) + rf(y_i|\mu_2, \sigma^2)} \\ p_2(y_i) &= \frac{(1-r)f(y_i|\mu_1, \sigma^2)}{rf(y_i|\mu_1, \sigma^2) + (1-r)f(y_i|\mu_2, \sigma^2)} \dots (14) \end{aligned}$$

ويمكن ان نحصل على ما يأتي :

$$\frac{\partial}{\partial \mu_1} \log L(\mu_1, \mu_2, \sigma^2, r|y) = \sum_{i=1}^{n_1} p_1(y_i)(y_i - \mu_1) + \sum_{i=n_1+1}^n p_2(y_i)(y_i - \mu_1) \dots (15)$$

نساوي المعادلة (15) للصفر ونشتق بطريقة Yield لإيجاد تقدير  $\mu$  وكالاتي :

$$\hat{\mu} = \frac{\sum_{i=1}^{n_1} p_1(y_i)(y_i) + \sum_{i=n_1+1}^n p_2(y_i)(y_i)}{\sum_{i=1}^{n_1} p_1(y_i) + \sum_{i=n_1+1}^n p_2(y_i)} \quad \dots(16)$$

نساوي المعادلة (15) للصف ونشتق بطريقة Yield لإيجاد تقدير  $\sigma^2$  وكالاتي :

$$\hat{\sigma}^2 = \frac{1}{n} (\sum_{i=1}^{n_1} [p_1(y_i)(y_i - \hat{\mu}_1)^2 (1 - p_1(y_i))(y_i - \hat{\mu}_2)^2] + \sum_{i=n_1+1}^n p_2(y_i)(y_i - \hat{\mu}_1)^2 (1 - p_1(y_i))(y_i - \hat{\mu}_2)^2) \quad \dots(17)$$

إنّ المعادلتين (16) و (17) لا تحل الاحتمالية لجميع معاملاتهما لأن  $p_1(y_i)$  و  $p_2(y_i)$  تعتمدان على المعلمات ويعتمد أيضاً على  $r$ . نملك جزءاً من مخطط التكرار للعثور على المعلمات. بعد تقدير  $\mu_1, \mu_2, \sigma^2$  نحن نستخدم القيم الحالية لإيجاد  $r, p_1, p_2$  وحتى نحصل على التقارب يجب افتراض الاتي :

- 1- يعتبر  $r$  ثابت .
- 2- نستخدم المعادلتين (16) و (17) لإيجاد تقدير  $\mu_1, \mu_2, \sigma^2$  .
- 3- نستخدم التقدير الحالي لـ  $\mu_1, \mu_2, \sigma^2$  بأعظم قيمة

$$\log L(\mu_1, \mu_2, \sigma^2, r|y) = \frac{\partial}{\partial \mu_1} \sum_{i=1}^{n_1} \log((1-r)f(y_i|\mu_1, \sigma^2) + rf(y_i|\mu_2, \sigma^2)) + \frac{\partial}{\partial \mu_1} \sum_{i=n_1+1}^n \log(rf(y_i|\mu_1, \sigma^2) + (1-r)f(y_i|\mu_2, \sigma^2))$$

- 4- نعمل على التكرار بين الخطوة 2 والخطوة 3 حتى نحصل على التقارب. (8)

## 6. الجانب العملي

### 1.6 تحليل الواسمات للنمط الظاهري Marker Analysis of Phenotype

سيتم في هذا الجزء اجراء تجارب المحاكاة لغرض توليد مجموعتين من البيانات (صفات النمط الظاهري) تتبع كل منهما التوزيع الطبيعي حيث تكون المجموعة الاولى ناتجة من اعادة التركيب الجزيئي بين واسمتين ( $M_1M_1$ ) بينما تمثل المجموعة الثانية مجموعة البيانات الناتجة من اعادة التركيب الجزيئي بين الواسمتين ( $M_1M_2$ ) اذ سيتم توليد مجموعتي البيانات وفقاً للنموذج المختلط في العلاقة (19):

$$M_1M_1 \text{ المشاهدة } \rightarrow y \sim (1-r)N(\mu_1, \sigma^2) + rN(\mu_2, \sigma^2)$$

$$M_1M_2 \text{ المشاهدة } \rightarrow y \sim rN(\mu_1, \sigma^2) + (1-r)N(\mu_2, \sigma^2) \quad \dots(19)$$

اذ ان :

$\mu_1, \mu_2$  : تمثل الاوساط الحسابية للصفة المظهرية الناتجة من ازواج الواسمات المشار اليها على التوالي .

ومن الجدير بالملاحظة هنا سيتم الكشف عن الترابط بين الواسمات والصفات المظهرية من خلال معامل اعادة التركيب الجزيئي ( $r$ ) فاذا كان  $r = 0.5$  فهذا يعني عدم الترابط بين الواسمات والصفات المظهرية الكمية اي ان متوسطي المجموعتين من الصفات المظهرية يكونان متساويين .

فرضية الاختبار لعد وجود الترابط هي :

$$H_0: r = 0.5 \text{ or } H_0: \mu_{M_1M_1} - \mu_{M_1M_2} = (1-2r)(\mu_1 - \mu_2) = 0$$

الجدول (5) يبين نتائج تجارب المحاكاة المنفذة وفقاً لقيم افتراضية مختلفة وحسب ما تم تعيينه في الجدول .



بالنسبة للتجربة الأولى الخاصة بأحجام العينات ( $n_1 = n_2 = 100$ ) كانت القيمة التقديرية لمتوسط المجموعة الأولى (94.188) ومتوسط المجموعة الثانية (100.3213) وهي قريبة من القيم المفترضة لهذه المتوسطات في حين كانت قيمة التباين المقدر (3.799) وهي أيضاً تمثل قيمة تقديرية جيدة للتباين ، اما بالنسبة لقيمة (r) المقدره فقد بلغت (0.0378) وهي أيضاً تمثل قيمة تقديرية قريبة جداً من قيمة (r) المفترضة .

أما في حالة كون احجام العينات ( $n_1 = n_2 = 100$ ) فقد كانت القيمة التقديرية لمتوسط المجموعة الأولى (94.7938) والقيمة التقديرية لمتوسط المجموعة الثانية (99.7775) وكانت قيمة التباين المقدر (4.1508) في حين كانت قيمة (r) المقدره بلغت (0.362) .

أما بالنسبة للتجربة الثانية فقد تم توليد مجموعتي البيانات بالقيم الافتراضية نفسها ما عدا قيمة (r) فقد افترضت انها مساوي (0.25) .

القيم التقديرية لمتوسطي المجموعتين والتباين ولمختلف حجوم العينات اظهرت تقارباً مع القيم الافتراضية في حين ان القيم الافتراضية (r) سجلت قيماً (0.1727) للمجموعة الأولى من حجوم العينات و (0.1993) للمجموعة الثانية من حجوم العينات .

التجربة الثالثة تم توليد مجموعتين من البيانات بالقيم الافتراضية نفسها مع تغيير قيمة (r) المفترضة لتصبح (0.45) ، القيم التقديرية لمتوسطي المجموعتين والتباين اظهرت تقارباً مع القيم المفترضة لكل منهم في حين ان القيم التقديرية لـ (r) سجلت قيمة (0.2947) و (0.3270) للمجموعتين من احجام العينات على التوالي .

خلاصة التجارب الثلاثة الأولى بثبات احجام العينات والقيم الافتراضية للمتوسطات والتباين مع اختلاف القيم الافتراضية لـ (r) تبين لنا أن القيم التقديرية لمعامل اعادة التركيب الجزيئي (r) تتناقص كلما زادت القيم الافتراضية لها وبما أن (r) يعد مؤشراً على وجود الترابط بين الصفات المظهرية و الواسمات لذلك عند تغير في قيم (r) التقديرية يكون هنالك مؤشراً واضحاً على وجود ترابط معنوي بين الصفات المظهرية و الواسمات ، اي ان عملية العبور من الواسمة ( $M_1$ ) الى ( $M_2$ ) توفر دليلاً واضحاً على الكشف عن الصفات المظهرية (QTL).

ولغرض تقييم فرضية الاختبار أعلاه تم عمل اختبار التباديل من خلال تنفيذ (5000) عينة تبادلية ومن ثم حساب الإحصاء ( $-2\log \lambda$ ) لكل عينة ومن ثم تحديد القيم الحرجة باستعمال الأشكال البيانية لتوزيعات تباديل العينات .

نتائج الاختبار كما مبينة في العمودين التاسع والعاشر من الجدول (5) (cutoff) لـ 95% من توزيع تباديل العينات في ظل الفرضية البديلة ونظريتها (cutoff) لـ 95% في ظل فرضية العدم على التوالي .

قاعدة القرار رفض الفرضية الصفرية في حال كون إحصاء اختبار التباديل في ظل الفرضية البديلة اصغر من نظريتها في ظل فرضية العدم .

( 5 ) الجدول

القيم التقديرية لطريقة MLE لـ  $\mu_1, \mu_2, \sigma^2$ ، عندما تكون

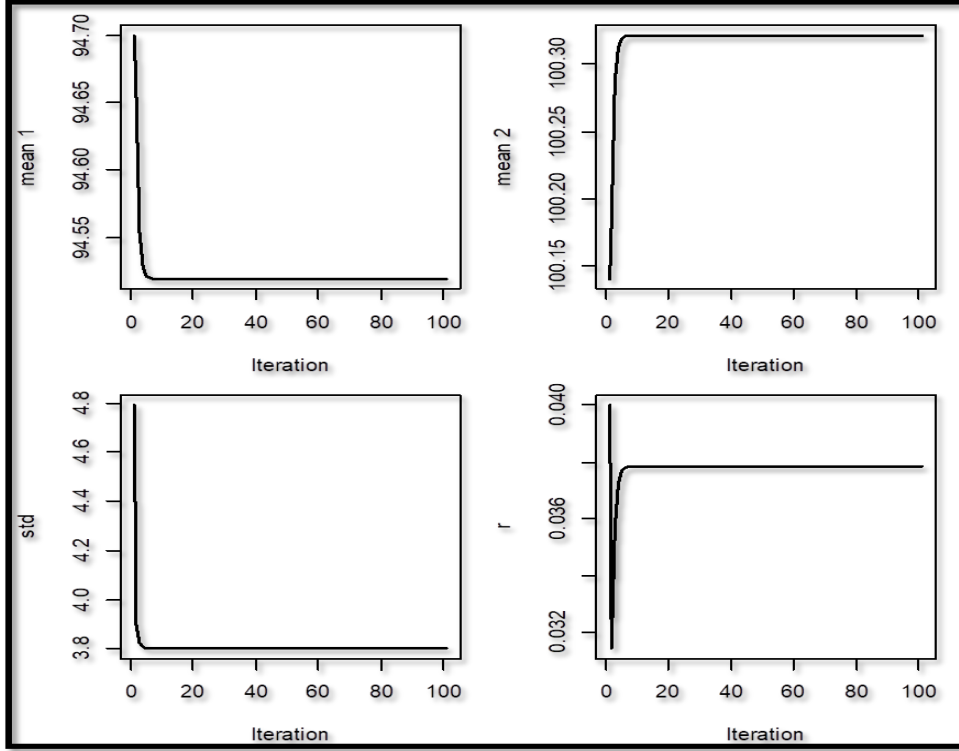
$$\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4$$

$\mu_1, \mu_2$	$\sigma^2$	r	SAMPLE SIZES	MLE				Cutoff	$-2\log \lambda$ under $H_0$	Decision
				$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}^2$	$\hat{r}$			
$\mu_1 = 95$ $\mu_2 = 100$	4	0.0 4	$n_1 = 100 \quad n_2 = 100$	94.5188	100.321 3	3.7990	0.037 8	5.573 2	10	significant
			$n_1 = 200 \quad n_2 = 200$	94.7938	99.7775	4.1508	0.036 2	3.979 4	9	significant
		0.2 5	$n_1 = 100 \quad n_2 = 100$	93.5908	100.701 7	3.1583	0.172 7	8,7438	8	Non- significant
			$n_1 = 200 \quad n_2 = 200$	93.9204	100.869 7	3.4586	0.199 3	3.889 4	5	significant
		0.4	$n_1 = 100 \quad n_2 = 100$	94.0986	100.786	3.8984	0.294	3.854	8	significant

		5			3		7	2		
			$n_1 = 200$ $n_2 = 200$	94.4837	100.493 3	3.8333	0.327 0	9.963 8	3	Non-significant

الشكل (1)

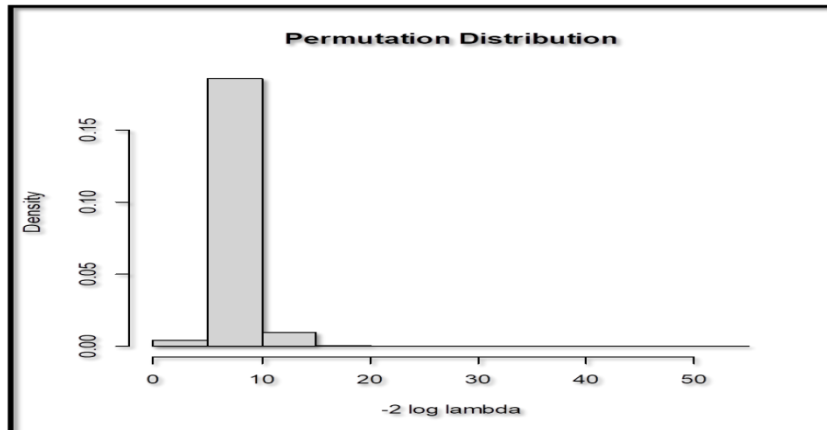
الرسم للقيم التقديرية لطريقة MLE لـ  $\mu_1, \mu_2, \sigma^2, r$  عندما تكون  $\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.04, n_1 = 100, n_2 = 100$



الشكل (2)

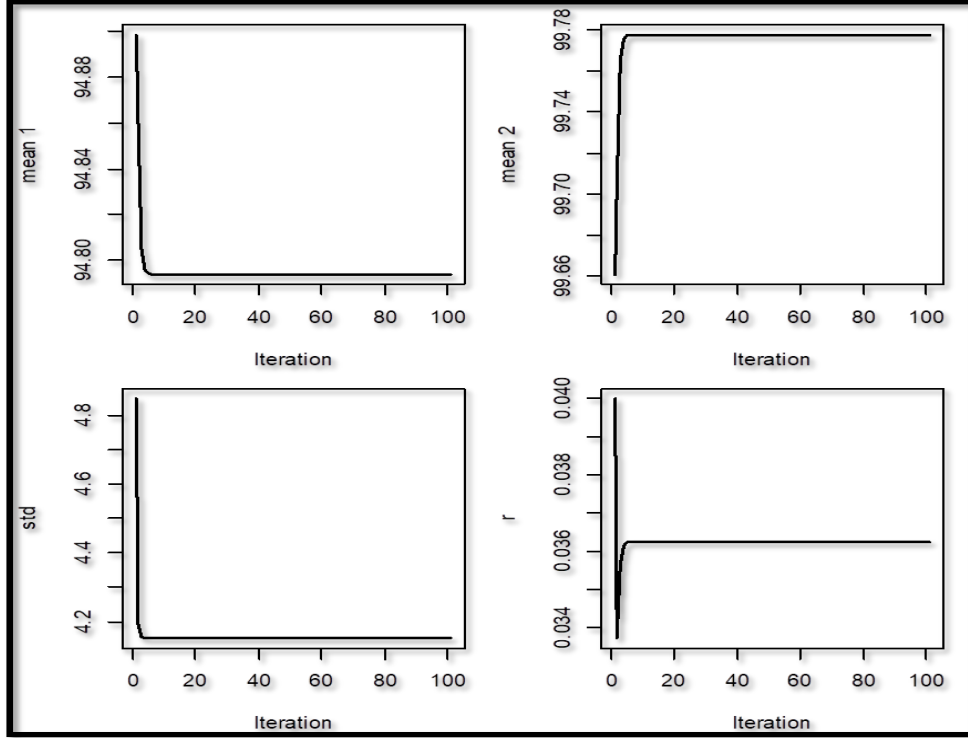
توزيع التباديل لإحصاءة نسبة الامكان  $(-2\log \lambda)$  تحت فرضية العدم بالاعتماد على (5000) من تباديل البيانات المولدة بالقيم الافتراضية

$\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.04, n_1 = 100, n_2 = 100$



الشكل (3)

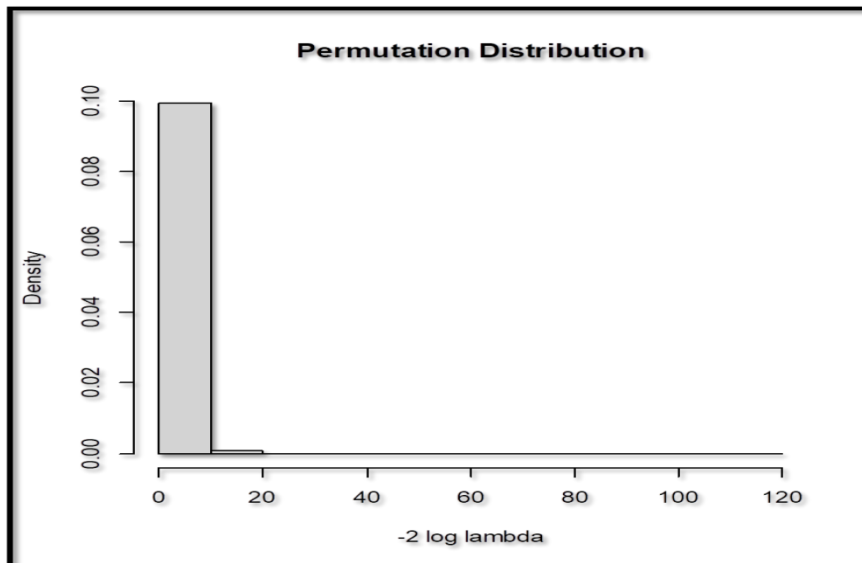
الرسم للقيم التقديرية لطريقة MLE لـ  $\mu_1, \mu_2, \sigma^2, r$  عندما تكون  $\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.04, n_1 = 200, n_2 = 200$



الشكل (4)

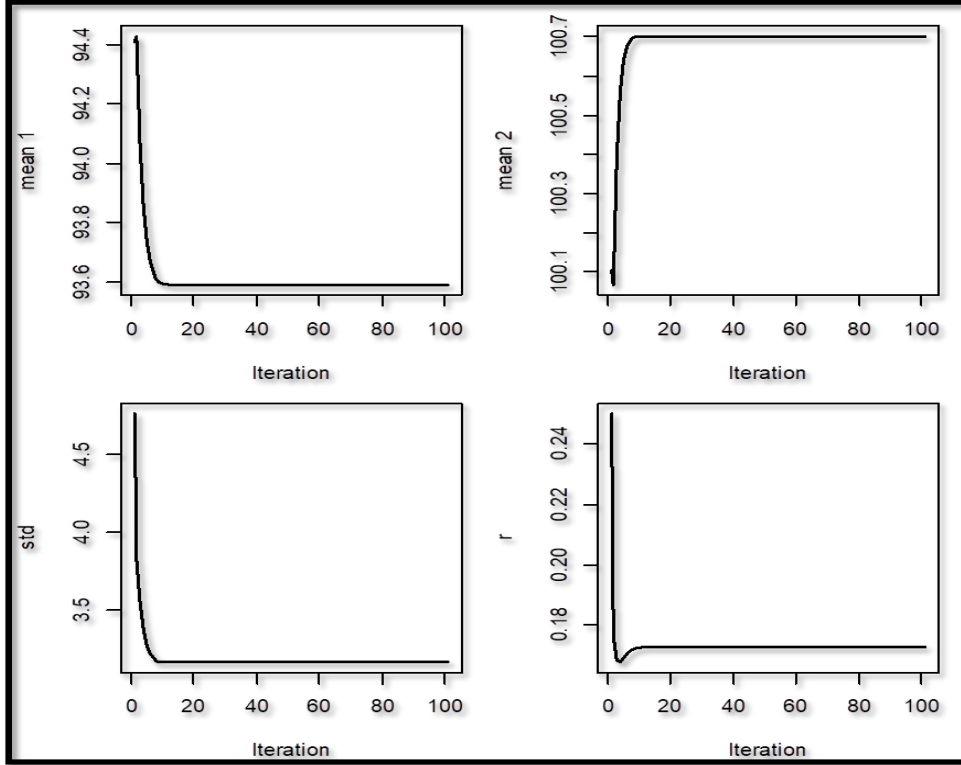
التوزيع للقيم التقديرية بطريقة MLE

$\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.04, n_1 = 200, n_2 = 200$



الشكل (5)

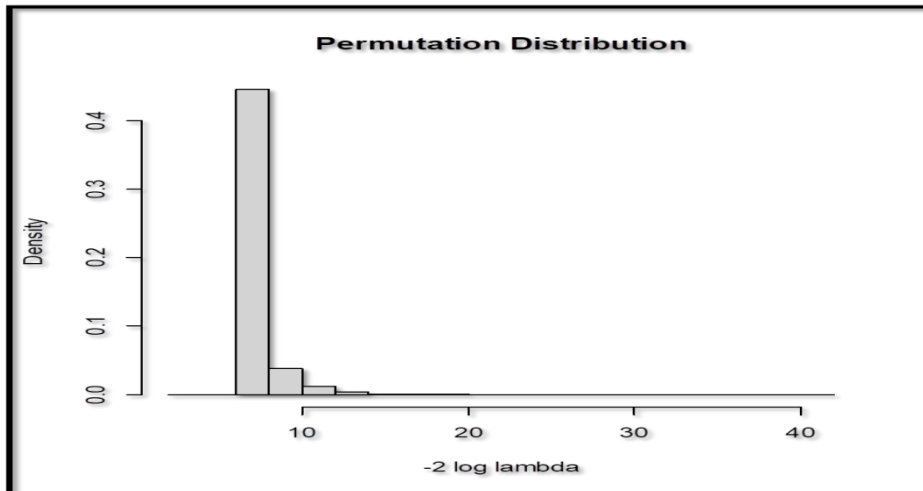
الرسم للقيم التقديرية لطريقة MLE لـ  $\mu_1, \mu_2, \sigma^2, r$  عندما تكون  $\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.25, n_1 = 200, n_2 = 200$



الشكل (6)

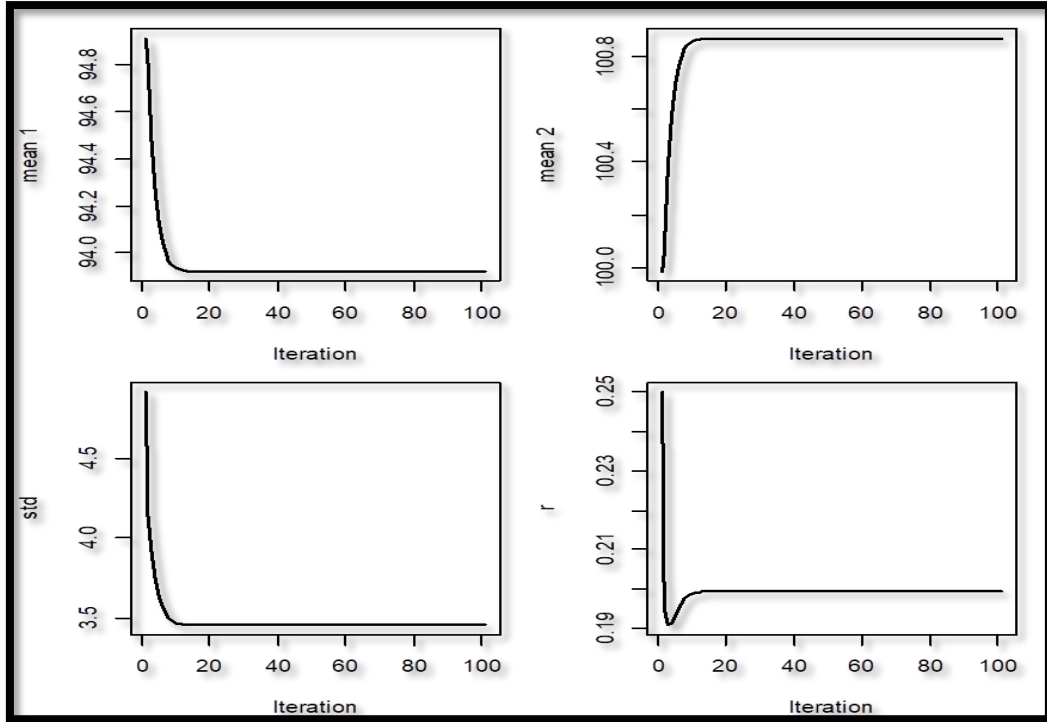
التوزيع للقيم التقديرية بطريقة MLE

$\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.25, n_1 = 100, n_2 = 100$



الشكل (7)

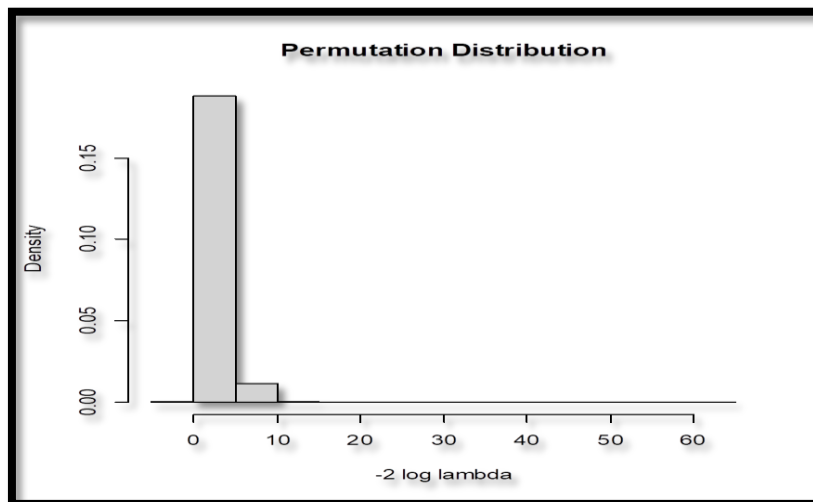
الرسم للقيم التقديرية لطريقة MLE لـ  $\mu_1, \mu_2, \sigma^2, r$  عندما تكون  $\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.25, n_1 = 200, n_2 = 200$



الشكل (8)

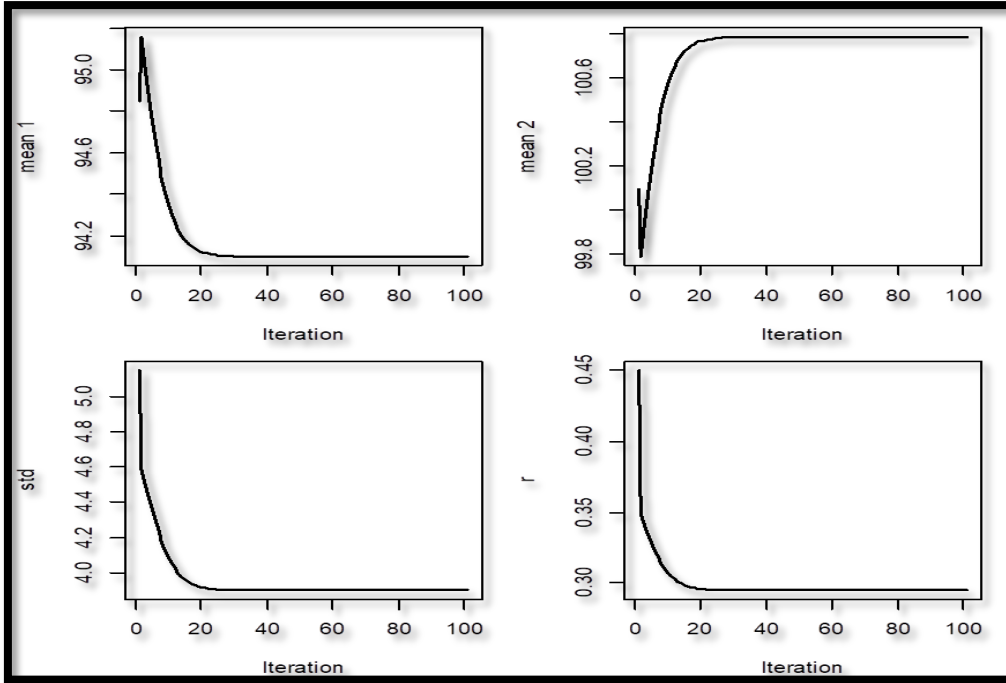
التوزيع للقيم التقديرية بطريقة MLE

$\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.25, n_1 = 200, n_2 = 200$



الشكل (9)

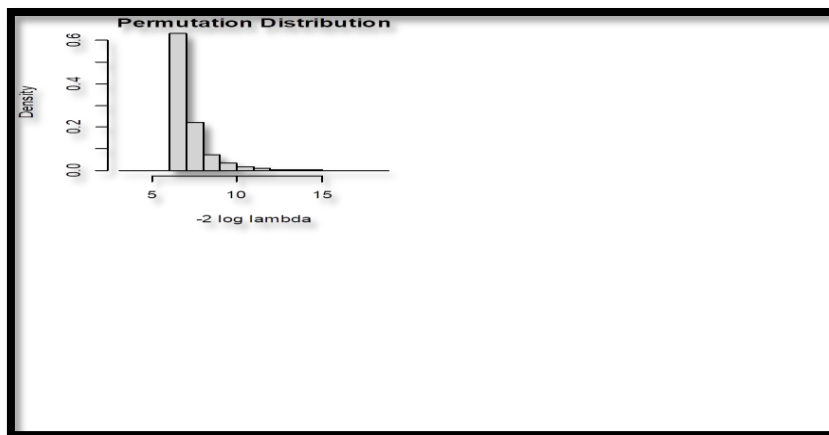
الرسم للقيم التقديرية لطريقة MLE لـ  $\mu_1, \mu_2, \sigma^2, r$  عندما تكون  $\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.45, n_1 = 100, n_2 = 100$



الشكل (10)

التوزيع للقيم التقديرية بطريقة MLE

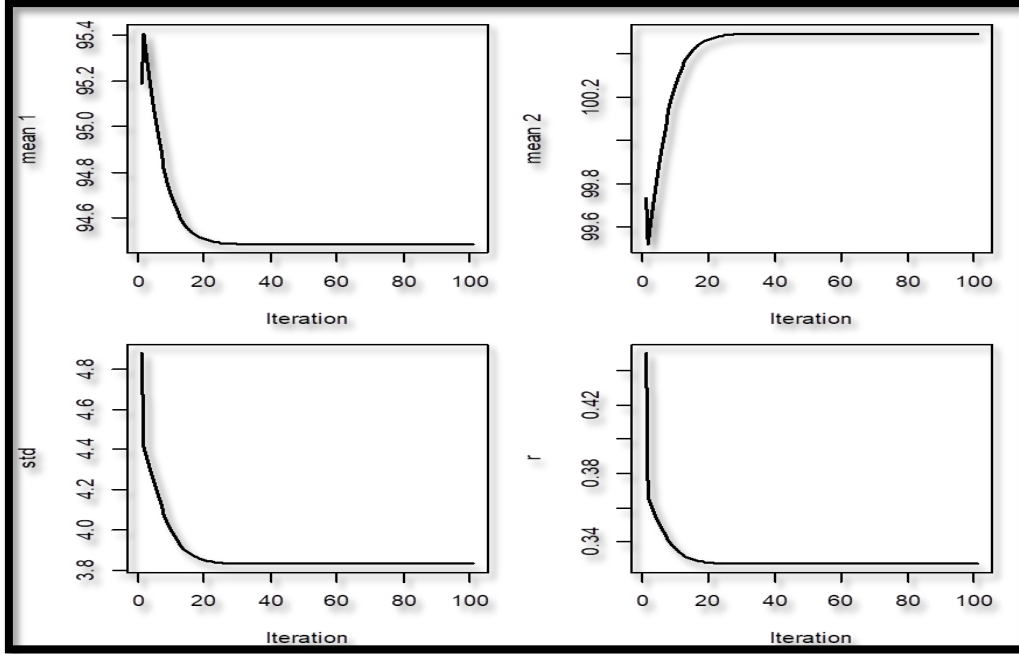
$\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.45, n_1 = 100, n_2 = 100$





الشكل (11)

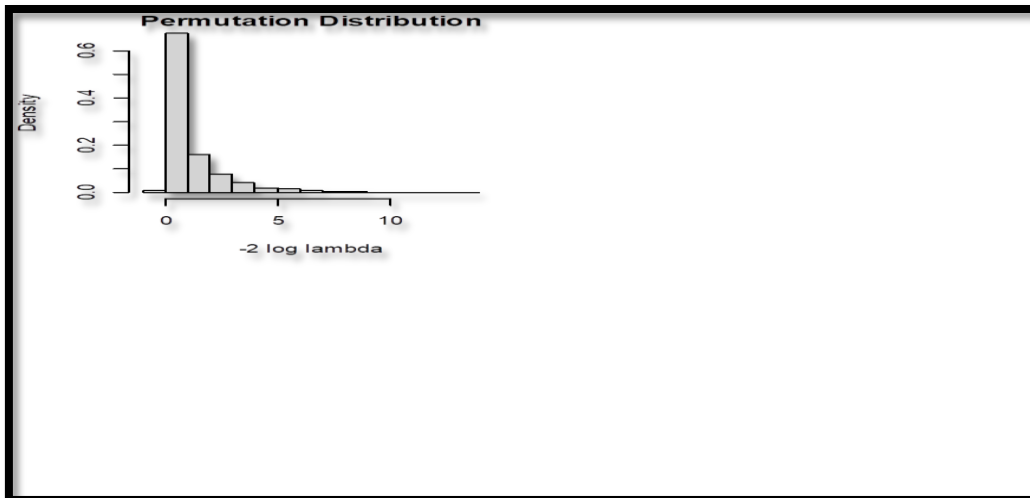
الرسم للقيم التقديرية لطريقة MLE لـ  $\mu_1, \mu_2, \sigma^2, r$  عندما تكون  $\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.45, n_1 = 200, n_2 = 200$



الشكل (12)

التوزيع للقيم التقديرية بطريقة MLE

$\mu_1 = 95, \mu_2 = 100, \sigma^2 = 4, r = 0.45, n_1 = 200, n_2 = 200$



## 7. الاستنتاجات :

1. بينت الدراسة باستخدام طريقة الامكان الاعظم لتقدير قيم (r) أن القيم التقديرية لمعامل اعادة التركيب الجزئي (r) تتناقص كلما زادت القيم الافتراضية له وبما أن (r) يعد مؤشراً على وجود الترابط بين الصفات المظهرية و الواسمات لذلك عند حصول تغير في قيم (r) التقديرية المتناقصة يكون هنالك مؤشراً واضحاً على وجود ترابط معنوي بين الصفات المظهرية و الواسمات ، اي ان عملية العبور من الواسمة (M<sub>1</sub>) الى (M<sub>2</sub>) توفر دليلاً واضحاً على الكشف عن الصفات المظهرية (QTL) في حالة كون حجم العينة (n<sub>1</sub> = n<sub>2</sub> = 100).
2. إنّ عملية العبور من الواسمة (M<sub>1</sub>) الى (M<sub>2</sub>) في حالة كون حجم العينة (n<sub>1</sub> = n<sub>2</sub> = 200) توفر دليلاً واضحاً على الكشف عن الصفات المظهرية (QTL) ومن الجدير بالملاحظة ايضاً هنالك انخفاض كبير في قيم (r) التقديرية في المجموعة الثانية من التجارب مقارنة مع مثيلاتها في المجموعة الاولى وهو ما يؤشر على وجود دليل أقوى للفرق بين متوسطي مجموعتي البيانات تحت تأثير الواسمات .
3. دل اختبار التباديل من خلال تنفيذ (5000) عينة تبادلية ومن ثم حساب الإحصاء ( -2log λ ) لكل عينة ومن ثم تحديد القيم الحرجة باستعمال الأشكال البيانية لتوزيعات تباديل العينات رفض الفرضية الصفرية في حال كون إحصاء اختبار التباديل في ظل الفرضية البديلة اصغر من نظيرتها في ظل فرضية العدم .

## 8. التوصيات

1. نوصي بتوسعة الدراسة باستعمال اكبر عدد من المتغيرات وكذلك التطبيق على متغيرات لعينات وراثية وتنفيذ التجارب على الامراض التي تصيب الانسان والعوامل الوراثية وكشفها بصورة مبكرة .
2. إنّ هناك طرائق ممكن استخدامها في تحليل الخرائط الجينية QTL مثل طريقة بيز وطريقة الامكان الاعظم الجزائية واقترح استعمالها من قبل الباحثين اللاحقين .

## المصادر :

1. Lander, E.S. , and D . Botstein , (1989) , " Proposed the use of 1 and 2 LOD support Intervals " , Genetics ,121:185-199 .
2. Whittaker , J.C. , Thompson ,R ., & Visscher , p.(1996), "On the mapping of QTL by regression of phenotype on marker Type ".Heredity 77:23-32 .
3. Walsh, B . & Lynch, M . (2001), " Genetics and Analysis of Quantitative Traits " , the American journal human genetics (Am J Hum Genet) . 2001,68(2):548-549 .
4. Vijay Rani Rajpal , Rama Rao & Raina , 2016 , " Quantitative Trait Loci Mapping in Plants: Concepts and Approaches " , Sustainable Development and Biodiversity book series (SDEB, volume 11) .

5. Weller&Bickhar,2018, "Determination of quantitative trait nucleotides by concordance analysis between quantitative trait loci and marker genotypes of US Holsteins", American Dairy Science Association .
6. Miles, C; Wayne, M (2008). "Quantitative trait locus (QTL) analysis". Nature Education.
7. Complex Trait Consortium (2003). " The nature and identification of quantitative trait loci :a community's view ". Nature Reviews Genetics.4 (11): 911–916. doi:10.1038/nrg1206. PMC 2063446. PMID 14634638.
- 8.Watanabe, Satoshi; Hideshima, Rumiko; Xia, Zhengjun; et al. (2009). "Map-Based Cloning of the Gene Associated With the Soybean Maturity Locus E3". Genetics. 182 (4): 1251–1262. doi:10.1534/genetics.108.098772.PMC 2728863.PMID 1947424