



المقارنة بين طرائق تحليل الانحدار الـ QTL باستخدام المحاكاة

Comparison of QTL Rregression Analysis Methods Using Simulation

ا. دكتور عباس لفته كنهير⁽²⁾

حميدة خضر ابراهيم⁽¹⁾

(Abstract) : المستخلص

في هذا البحث تم دراسة النموذج الجيني الذي يتناول خواصاً جينية أو سمات (Markers) لمناطق محددة من الكروموسومات التي تكون خواصها طبيعية تتأثر بالبيئة لذلك فإن استعمال نماذج الانحدار الخطي المتعدد تكون مناسبة لقياس هذا التأثير ولذلك تم استخدام طرائق منها انحدار هيلي نوت Haley- Knott regression و انحدار هيلي نوت الموسع Extended Haley- Knott regression و انحدار الواسمات Regression Marker واستخدام خوارزمية الانحدار أو خوارزمية (EM) EM Regression OR EM algorithm و لغرض التطبيق تم محاكاة مجتمع (F2 intercross) بالاعتماد على نموذج مواقع الصفات الكمية التجميعي (additive QTL model) باستعمال الدالة (sim.cross) ضمن بيئة البرنامج R .

ولغرض المقارنة بين أداء طرائق تحليل مواقع الصفات الكمية (QTL) تم محاكاة مجتمع التضارب العكسي (backcross) بالاعتماد على نموذج مواقع الصفات الكمية التجميعي (additive QTL model) بوجود سلالتين فطريتين هما (A) و (B) باستعمال الدالة (sim.cross) ضمن بيئة البرنامج R وتطبيق هذه الطرائق .

Abstract:

The genetic model is concerned with genetic traits or (Markers) for particular chromosomal regions whose traits are typically influenced by the environment. Multiple linear regression models can therefore be used to measure this effect. Consequently, techniques were employed, such as Knott regression. Extended Haley-Knott regression and the Haley model the Knott regression utilizing the regression technique and the Haley-Extended and Regression Marker (EM) For the purposes of the application, the (sim.cross) function within the R environment was used to simulate the F2 intercross population based on the additive QTL model.

The (sim.cross) function has been used to simulate a backcross population based on the additive QTL model with the presence of two fungal strains (A) and (B) in

order to compare the effectiveness of the quantitative trait loci (QTL) methodologies.

1. المقدمة: (Introduction)

إنّ موضع السمة الكمية Quantitative Attribute position السمة الكمية (QA) هي سمة خاصة تُستخدم لمقارنة قيمتين ، أي تُستخدم لمقارنة قيمة معرفة من قبل المستخدم بالحد الأعلى والحد الأدنى. والنمط الفردي Haplotype مجموعة من الجينات المرتبطة التي تحتل مواقع متجاورة على الكروموسوم نفسه والتي عادة ما يتم توريثها سوياً كوحدة واحدة. وان الواسمات أو السمات Markers وهي مواد كيميائية تستخدم لكشف مواقع جينية معينة على شريط ال DNA، والنمط الجيني Genotype وهو التعبير الوراثي المسؤول عن اظهار صفة معينة بمعنى اخر الجين صيغة الجين المسؤول عن تلك الصفة المعينة وهو علم يمكن أن يكون امتداداً للنظرية المنديلية البسيطة التي تقول بأن التأثير المشترك لجين أو أكثر من الجينات والبيئات ويتم التعبير عن الأنماط الظاهرية (Phenotype) بالآتي:.

$$P = G + E$$

إذ إن قيمة النمط الظاهري (P) (Phenotype) للفرد هو الأثر المشترك لقيمة النمط الجيني (G) (Genotype) والتأثير البيئي (E) (Environment). وأن قيمة النمط الجيني هو الأثر الناتج عن جميع الآثار الوراثية المشتركة، بما في ذلك الجينات النووية والجينات الميتوكوندريا والتفاعلات بين الجينات ومن الممكن استخدام بعض الطرائق للكشف عن ال-QTL ومعرفة افضل طريقة لتشخيص الصفات الكمية على الكروموسومات ضمن التجربة .

2. مشكلة البحث: (The research problem)

غالبا ما تواجه الباحثين تغيرات أو طفرات وراثية مرغوب أو غير مرغوب بها تحدث في الأجيال المتعاقبة نتيجة لموروثات جينية وكذلك تكون متأثرة بالعوامل البيئية الأمر الذي يتطلب صياغة أنموذج إحصائي يمكن إن يبين هذه التأثيرات ويقاس اتجاه وعلاقة هذه التأثيرات ومعرفة افضل طريقة لتشخيص الصفات الكمية على الكروموسومات للموروثات الجينية ضمن التجربة .

3. هدف البحث: (Purpose of search)

يهدف البحث الى دراسة وتحليل تأثير أثر مواقع الصفات الكمية للموروثات الجينية عبر الأجيال المتعاقبة من خلال تقدير طرائق ال-QTL كطريقة انحدار هيلي نوت Haley- Knott regression وانحدار هيلي نوت الموسع Extended Haley- Knott regression وانحدار الواسمات Regression Marker واستخدام خوارزمية الانحدار أو خوارزمية Regression OR EM algorithm (EM) ولغرض ايجاد افضل تقدير ل-QTL .

4. الجانب النظري:

1.4 المقدمة:

يعد تحليل مواقع الصفات الكمية (QTL) Quantitative Trait Loci Analysis 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 (7) (Genetic statistical Analysis)

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

- 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 توزيعاً مستمراً 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 (8).
- إذ أنه يمكن تعريف (النموذج الجيني) بأنه يتناول خواصاً جينية أو سمات (Markers) لمناطق محددة من الكروموسومات التي تكون خواصها طبيعية تتأثر بالبيئة وأن أزواج هذه الكروموسومات متماثلة identical مثل الوالدين والعلاقة بين الصفات البيئية والصفات الطبيعية كذلك فإن استعمال نماذج الانحدار الخطي المتعدد تكون مناسبة لقياس هذا التأثير (7).
- وهذا يجعل تحليل الانحدار طريقة مرغوبة لاستخدامه في طرائق إعادة أخذ العينات وأن طريقة إعادة العينات تحدد الاختبارات الإحصائية لاختبار العينات وسوف يتم عرض رسم الخرائط لـ QTL للعلامات المتعددة أكثر من 2 وطريقة حساب أكثر من نقطة نوعية واحدة .

2.4 طريقة الانحدار Regression method

يمكن لتحليل التباين أن يكون بنوعين وكالاتي :

2.4.1 تحليل التباين ANOVA باستخدام الأنماط الجينية ذات السمة المفردة

Analysis ANOVA using single marker genotypes

يمثل النمط الوراثي للسمة (النمط الفردي لسمة) فئة تأثير الثابت بالاتي :

$$Y = M + MG_1 + e \quad \dots \quad (1)$$

إذ إن :

Y : النمط الظاهري للمشاهدة .

M : الاليل او الواسمة Marker .

MG₁ : الواسمة للجين الاول .

وأن تحليل إحصاءة F يوفر طريقة سهلة وبسيطة للكشف عن السمات المرتبطة مع الـ QTL⁽⁸⁾⁽²⁾ .

2.4.2 تحليل التباين ANOVA باستخدام الأنماط الجينية ذات السمات المتعددة

Analysis ANOVA using multiple marker genotypes

يتمثل كل نمط وراثي محدد للسمات (أو النمط الفرداني للسمات فئة التأثيرات ثابتة) بالاتي:

$$Y = M + MG_1 + MG_2 + \dots + MG_n \quad \dots \quad (2)$$

هذا هو نموذج الانحدار المتعدد والسمات يمكن إن تحذف من النموذج إذا لم تكن كبيرة . تشير مجموعة السمات العامة في التحليل النهائي إلى وجود تأثير الـ QTL كبير (أو كثير اعتمادا على مدى تباعد السمات) .

وبهذا المعنى فإنه يمكن مقارنته مع الانحدار على النمط الجيني لسمة الفردية . وإن طريقة السمة المتعددة هي أكثر قوة في تحليل الواسم الفردي عندما تكون الواسمات منتشرة بشكل جيد على الجينوم فإنها تكون أكثر قدرة على تمييز موقع الـ QTL ويوصى عادة بعد الكشف عن مثل هذا الموقع في التحليل الثاني لرسم خرائط الفترة⁽⁴⁾ .

2.4.3 الانحدار على احتمال مواقع الصفات الكمية في الـ QTL مشروطا بعلامات الأنماط الفردية

Regression on QTL probability conditional on Marker haplotypes

تعطى سمة النمط الوراثي أو النمط الفرداني لعلامة الموروثة من الأب ويمكننا حساب احتمال وراثته الاليل Q أو q . لذلك ومن الطبيعي يمكن ان يتراجع النمط الظاهري على الاحتمال Q-prob والانموذج هو :

$$Y = M + \alpha x + e \quad \dots \quad (28.2)$$

إذ إن :

Y : النمط الظاهري للمشاهدة .

X : هو الاحتمال المعطى للأنماط الأبوية Q وهو احتمال الوراثة الجينية لسمة المشاهدة ، وموضع السمة (أو العلامة) الـ QTL (marker/QTL positions P(Q|mg1, mg2, r₁, r₁₂)) .

α : معامل الـ x .

ويمكن تحديد المواقع المتبقية لكل الـ QTL كما يمكن تقدير موقع القرب QTL في الحالات التي يكون فيها معيار SEE اقل ما يمكن⁽³⁾ .

2.2.4 انحدار هيلي نوت Haley- Knott regression

اقترح العالم Haley (1992)⁽²⁾ إعادة تحديد طفيف لنموذج الانحدار السابق ولكن المبدأ مماثل ، بدلاً من التعامل مع الأنماط الفردانية الواسمة ، فإنه قدم نموذجاً أكثر عمومية حيث اعتمد الأنماط الجينية QTL بدلا من الأنماط الجينية الواسمة (ذات السمة) . إن احتمال عمل النموذج جيني معين يعتمد على الأنماط الجينية المميزة والتصميم يكون كالاتي :

$$Y = Mx_1 + \beta x_2 + e \quad \dots \quad (3)$$

إذ ان :

Y : المشاهدات الظاهرية .

x_2, x_1 : هو الاحتمال الشرطي لـ **QTL** المشروط بالأنماط الجينية المصاحبة وان :

$$x_1 = p(QQ/M_i) - p(qq/M_j) \quad \dots \quad (4)$$

$$x_2 = p(Qq/M_i) \quad \dots \quad (5)$$

α, β : يمثل معامل الانحدار الفرق بين الأنماط الجينية لـ **QTL** المتماثلة الزيجات وتأثير الهيمنة لـ **QTL** على التوالي .

اقترح Haley استخدام الاختبار التالي المشار إليه على أنه اختبار تقريبي لنسبة الاحتمال أو (نسبة احتمال تقريبي):

$$LR = n \ln \left(\frac{SSE_{reduced}}{SSE_{full}} \right) = -n \cdot \ln(1 - r^2) \quad \dots \quad (6)$$

التي هي نسبة مجموع مربعات الأخطاء في النموذج الكامل SSE_{full} لـ **QTL** والنموذج المختزل $SSE_{reduced}$ ، إما الحد r^2 هو مربع r المعروف .

2.2.4 انحدار هيلي نوت الموسع Extended Haley- Knott regression

في طريقة انحدار هيلي نوت الموسع يفترض ان P_{ij} و μ_{ij} وكالاتي⁽²⁾:

$$y_i | M_i \sim N(\sum_j P_{ij} \mu_j, \sum_j P_{ij} (\mu_j - \sum_k P_{ik} \mu_k)^2 + \sigma^2)$$

إذ انه :

$$\begin{aligned} var(y_i | M_i) &= var [E(y_i | g_i) | M_i] + E[var(y_i | g_i) | M_i] \\ &= var(\mu_{gi} | M_i) + E(\sigma^2 | M_i) = \sum_j P_{ij} (\mu_j - \sum_k P_{ik} \mu_k)^2 \dots (7) \end{aligned}$$

إذ أنه :

$$m_i = \sum_j P_{ij} \mu_j$$

$$v_i(\mu, \sigma^2) = \sum_j P_{ij} [\mu_j - m_i(\mu)]^2 + \sigma^2 = \sum_j P_{ij} \mu_j^2 + (\sum_j P_{ij} \mu_j)^2 + \sigma^2 \dots (8)$$

إن طريقة هيلي نوت الموسع تفترض :

$$y_i | M_i \sim N(m_i(\mu), v_i(\sigma^2))$$

3.2.4 انحدار الواسمات Regression Marker

إنّ انحدار السمات او الواسمات يكون على انواع منها: انحدار السمات في التضارب الداخلي Intercross يعني التضارب الطبيعي بين الذكر والانثى في F_2 ، وانحدار السمات للتضارب العكسي Backcross وهو تهجين احد الافراد مع أحد الوالدين أو فرد مشابه وراثياً لوالده ، لتحقيق ذرية وراثية أقرب إلى هوية الوالد ويتم استخدامه في البستنة ، وتربية الحيوانات. وان الانحدار الخطي البسيط وفي هذه الانواع ينقسم حسب نوعية السمات انحدار سمة واحدة او انحدار متعدد السمات.⁽¹⁾

1.3.2.4 انحدار الواسمات التضارب العكسي Backcross

نفترض التضارب العكسي من P_1 يكون $F_1 = P_1 \times P_2$ ويمكن فرض حجم العينة كدراسة $500 - 100 \leq n$ لمجموعة الأفراد $300 - 75 \leq m$ وأن السمات ليست بالضرورة أن تكون مرتبة كخريطة ربط والهدف منه تحديد السمات المرتبطة بـ QTL للنظر في كل سمة على حدة ، وأن تقسيم الأفراد إلى مجموعتين يكون حسب النمط الجيني للسمات . وان تقدير قيم التركيب الجيني و الانحراف المعياري بفرض النمط الظاهري G_{AA} ، G_{Aa} يعني AA ، Aa ، وتقدير متوسطات العينة داخل المجموعة والانحراف المعياري المشترك (SD) من σ والمتوسط المرجح لـ SDs داخل المجموعة فان صيغة المتوسط والانحراف تكون كالآتي⁽⁵⁾⁽¹⁾ :

$$\hat{G}_{AA} = \sum_i (Y_i 1(X_{ij} = AA)) / n_{AA} \quad \dots \quad (9)$$

$$s_{AA}^2 = \sum_i (Y_i - \hat{G}_{AA})^2 1(X_{ij} = Aa) / n_{AA} \quad \dots \quad (10)$$

$$SD = \hat{\sigma}_{pool} = \sqrt{\frac{(n_{AA}-1)s_{AA}^2 + (n_{Aa}-1)s_{Aa}^2}{n_{AA} + n_{Aa} - 2}} \quad \dots \quad (11)$$

ولاختبار الفرضية الصفرية التي تفترض $G_{AA} = G_{Aa}$ وباستخدام إحصاء الاختبار t وكالاتي:

$$t = \frac{\hat{G}_{AA} - \hat{G}_{Aa}}{\hat{\sigma}_{pool} \sqrt{\frac{1}{n_{AA}} + \frac{1}{n_{Aa}}}} \quad \dots \quad (12)$$

وعليه يكون القرار حسب القيمة الجدولية وان الصيغ المستخدمة للانحدار السمات الداخلي فقط باختلاف مواقع الجينات .

4.2.4 خوارزمية الانحدار أو خوارزمية (EM) Regression OR EM algorithm

بتجاهل تأثيرات المتغيرات المشتركة ، يمكن التعبير عن القيمة المظهيرية y_i ، للفرد i في اختبار التضارب العكسي من حيث تأثير QTL وقيمة الخطأ بالآتي⁽¹⁾ :

$$y_i = \sum_{j_1 \leq j_2 = 1}^4 \xi_{j_1 j_2 | i} \mu_{j_1 j_2} + e_i \quad \dots \quad (13)$$

إذ انه :

$\xi_{j_1 j_2 | i}$: هو مؤشر المتغير فاذا كان (1) يعبر عن الفرد i الذي لديه نمط جيني لـ QTL بـ $j_1 j_2$ واذا كان $(j_1 \leq j_2 = Q_1 Q_2 Q_3 Q_4)$ فيما عدا ذلك يعبر عنه 0 .

e_i : الخطأ العشوائي الذي يتوزع توزيعاً طبيعياً بمتوسط 0 وتباين σ^2 .

$\mu_{j_1 j_2}$: هو متوسط القيمة الوراثية للنمط الجيني $j_1 j_2$ والذي يمكن تقسيم قيمه إلى تأثيرات وراثية مضافة والسائدة لأنواع مختلفة بافتراض ان هناك 10 أنماط وراثية مختلفة لأشراج QTL في رباعي الصيغة الصبغية متعددة التكافؤ ، معبراً عنه كالاتي :

$$\mu_{11} = \mu + a_1 \quad \text{for } Q_1 Q_1$$

$$\mu_{22} = \mu + a_2 \text{ for } Q_2 Q_2$$

$$\mu_{33} = \mu + a_3 \text{ for } Q_3 Q_3$$

$$\mu_{44} = \mu + a_1 - a_2 - a_3 \text{ for } Q_4 Q_4$$

$$\mu_{12} = \mu + a_1 + a_2 + d_{12} \text{ for } Q_1 Q_2$$

$$\mu_{13} = \mu + a_1 + a_3 + d_{13} \text{ for } Q_1 Q_3 \quad \dots(14)$$

$$\mu_{14} = \mu - a_2 - a_3 + d_{14} \text{ for } Q_1 Q_4$$

$$\mu_{23} = \mu + a_2 + a_3 + d_{23} \text{ for } Q_2 Q_3$$

$$\mu_{24} = \mu - a_1 - a_3 + d_{24} \text{ for } Q_2 Q_4$$

$$\mu_{34} = \mu - a_1 - a_2 + d_{34} \text{ for } Q_3 Q_4$$

حيث μ هو المتوسط العام وان a_1, a_2, a_3 هي التأثيرات الوراثية المضافة للأليلات Q_1, Q_2, Q_3, Q_4 بالنسبة للأليل $d_{34}, d_{24}, d_{23}, d_{14}, d_{13}, d_{12}$ هي التأثيرات الوراثية السائدة بسبب التفاعلات بين الأليلات المختلفة Q_1, Q_2 و Q_1, Q_3 و Q_1, Q_4 و Q_2, Q_3 و Q_2, Q_4 و Q_3, Q_4 على التوالي.

ومن منظومة المعادلات (14) ، يمكننا ايجاد المتوسط العام والتأثيرات المضافة والسائدة وكالاتي:

$$\mu = \frac{1}{4}(\mu_{11} + \mu_{22} + \mu_{33} + \mu_{44}) \quad \dots (15)$$

$$a_1 = \frac{1}{4}(3\mu_{11} + \mu_{22} + \mu_{33} + \mu_{44}) \quad \dots (16)$$

$$a_2 = \frac{1}{4}(3\mu_{11} + \mu_{22} + \mu_{33} + \mu_{44}) \quad \dots (17)$$

$$a_3 = \frac{1}{4}(3\mu_{11} + \mu_{22} + \mu_{33} + \mu_{44}) \quad \dots (18)$$

$$d_{12} = \mu_{12} - \frac{1}{4}(3\mu_{33} + 3\mu_{44} - \mu_{11} - \mu_{22}) \quad \dots (19)$$

$$d_{13} = \mu_{13} - \frac{1}{4}(3\mu_{22} + 3\mu_{44} - \mu_{11} - \mu_{33}) \quad \dots (20)$$

$$d_{14} = \mu_{14} - \frac{1}{4}(3\mu_{22} + 3\mu_{33} - \mu_{11} - \mu_{44}) \quad \dots (21)$$

$$d_{23} = \mu_{23} - \frac{1}{4}(3\mu_{11} + 3\mu_{44} - \mu_{22} - \mu_{33}) \quad \dots (22)$$

$$d_{24} = \mu_{24} - \frac{1}{4}(3\mu_{11} + 3\mu_{33} - \mu_{22} - \mu_{44}) \quad \dots (23)$$

$$d_{34} = \mu_{34} - \frac{1}{4}(3\mu_{11} + 3\mu_{22} - \mu_{33} - \mu_{44}) \quad \dots (24)$$

بافتراض Ω للإشارة إلى المتجه المجهول $\mu_{34}, \mu_{24}, \mu_{23}, \mu_{14}, \mu_{13}, \mu_{12}, \mu_{44}, \mu_{33}, \mu_{22}, \mu_{11}$ ولرسم خرائط QTL ، يمكن ملاحظة الأنماط الجينية للعلامة. لنفترض أن n هي مشاهدات النمط الجيني للعلامة l_1 ($l_1 \leq l_2$) ، $l_2 = M_1, M_2, M_3, M_4$) يتم بناء نموذج الاحتمال لبيانات النمط الظاهري (y) وبيانات العلامة (M) ، ضمن إطار نموذج مختلط وكالاتي :

$$L(\Omega | y, M) = \prod_{i=1}^n \prod_{l_1=M_1}^{M_4} \prod_{l_2=M_1}^{M_4} \sum_{j_1=Q_1}^{Q_4} \sum_{j_2=Q_1}^{Q_4} \pi_{j_1 j_2 | l_1 l_2} f_{j_1 j_2}(y_i) \quad l_1 \leq l_2, j_1 \leq j_2 \quad \dots$$

(25)

حيث $\pi_{j_1 j_2 | l_1 l_2}$ هو الاحتمال الشرطي للنمط الجيني $j_1 j_2$ المعطى لـ QTL وان النمط الجيني للعلامة $l_1 l_2$ ، ويفترض أن $f_{j_1 j_2}(y_i)$ يتبع التوزيع الطبيعي بمتوسط $\mu_{j_1 j_2}$ والتباين σ^2 .

إن الاحتمال الشرطي السابق $\pi_{j_1 j_2 | l_1 l_2}$ يتم حساب $j_1 j_2 | l_1 l_2$ على أنه النمط الوراثي للعلامة المشتركة مقسومًا على تكرار النمط الجيني للعلامة $l_1 l_2$.

يمكن الحصول على تقديرات المعلمات المجهولة التي تزيد من الاحتمالية (25) من خلال تنفيذ خوارزمية EM. نحسب الاحتمال السابق للنمط الجيني QTL بالنظر إلى النمط الوراثي للعلامة المحددة للفرد i وكالاتي :

$$\theta_{j_1 j_2 | l_1 l_2 i} = \frac{\pi_{j_1 j_2 | l_1 l_2} f_{j_1 j_2}(y_i)}{\sum_{j_1=Q_1}^{Q_4} \sum_{j_2=Q_1}^{Q_4} \pi_{j_1 j_2 | l_1 l_2} f_{j_1 j_2}(y_i)} \quad \dots (26)$$

ونحسب تكرارات الاحتمالات الجينية ويمكن ملاحظتها بناءً على الاحتمالات المحسوبة مسبقا للجينات مما أدى إلى تقديرات تكرارات الاختزال المزدوج وكالاتي :

$$\hat{\alpha} = \hat{g}_1 + \hat{g}_2 + \hat{g}_3 + \hat{g}_4$$

$$= \frac{1}{n}(n_{11} + n_{22} + n_{33} + n_{44}) \quad \dots (27)$$

$$\hat{\beta} = \hat{g}_1 + \hat{g}_2 + \hat{g}_5 + \hat{g}_6$$

$$\hat{r} = \frac{1}{2}(\hat{g}_3 + \hat{g}_5 + 2(\hat{g}_2 + \hat{g}_4 + \hat{g}_6) + 2\emptyset\hat{g}_7 + (1 + \psi)\hat{g}_8)$$

إذ انه يمكن التعويض \emptyset و ψ بالمعادلات الآتية :

$$\emptyset = \frac{r^2}{10r^2 - 18r + 9}$$

$$\psi = \frac{r}{3 - 2r} \quad \dots (28)$$

يتم تقدير القيمة الوراثية للنمط الجيني $j_1 j_2$ لـ QTL والتباين المتبقي وكالاتي :

$$\hat{\mu}_{j_1 j_2} = \frac{\sum_{i=1}^n \sum_{l_1=M_1}^{M_4} \sum_{l_2=M_1}^{M_4} \theta_{j_1 j_2 | l_1 l_2 i} y_i}{\sum_{l_1=M_1}^{M_4} \sum_{l_2=M_1}^{M_4} \theta_{j_1 j_2 | l_1 l_2 i}} \quad , l_1 \leq l_2; j_1 \leq j_2 \quad \dots (29)$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \sum_{l_1=M_1}^{M_4} \sum_{l_2=M_1}^{M_4} \sum_{l_1=Q_1}^{Q_4} \sum_{l_2=Q_1}^{Q_4} \theta_{j_1 j_2 | l_1 l_2 i} (y_i - \hat{\mu}_{j_1 j_2})^2 \quad l_1 \leq l_2; j_1 \leq j_2$$

... (30)

يتم التكرار بين الخطوة (28) و (26) وأن معادلات (28) تعتبر خطوة (3) بالتجربة ، وفي الخطوة اللاحقة يتم التكرار بين الخطوة (26) و (30) ، حتى يتم الحصول على تقديرات مستقرة وأن هذه التقديرات هي تقديرات الاحتمالية العظمى (MLEs) للمعلمات .

5. الجانب العملي

1.5 وصف تجربة المحاكاة

لغرض المقارنة بين أداء عدد من طرائق تحليل مواقع الصفات الكمية (QTL) (mr,em,hk,ehk) تم محاكاة مجتمع التضارب العكسي (backcross) بالاعتماد على نموذج مواقع الصفات الكمية التجميعي (additive QTL model) بوجود سلالتين فطريتين هما (A) و (B) باستعمال الدالة (sim.cross) ضمن بيئة البرنامج R .

التجربة تتضمن (120) مفردة بفرض صفة ظاهرية واحدة (Phenotype) بوجود (20) كروموسوم (chromosomes) وعدد الواسمات الكلية (187) واسمة موزعة على كل كروموسوم (chromosomes) إذ أن عدد الواسمات الموجودة على الكروموسوم الاول (14) واسمة (Marker) بطول (127CM) وبمتوسط مسافة (9.8CM) وبحد اعلى للمسافة بين كل علامتين (9.8CM) . اما الكروموسوم الثاني فإنه يحتوي على (12) واسمة بطول (114CM) وبمتوسط مسافة (10.4 CM) وبحد اعلى للمسافة بين كل علامتين (10.4 CM) وهكذا الحال بالنسبة الى الكروموسوم التاسع عشر فإنه يحتوي على (7) علامات بطول (55.7CM) وبمتوسط مسافة (9.3CM) وبحد اعلى للمسافة بين كل علامتين (9.3CM) . الجدول (1) يمثل ملخصاً للخارطة الجينية المستخدمة في عملية التوليد .

الجدول (1)

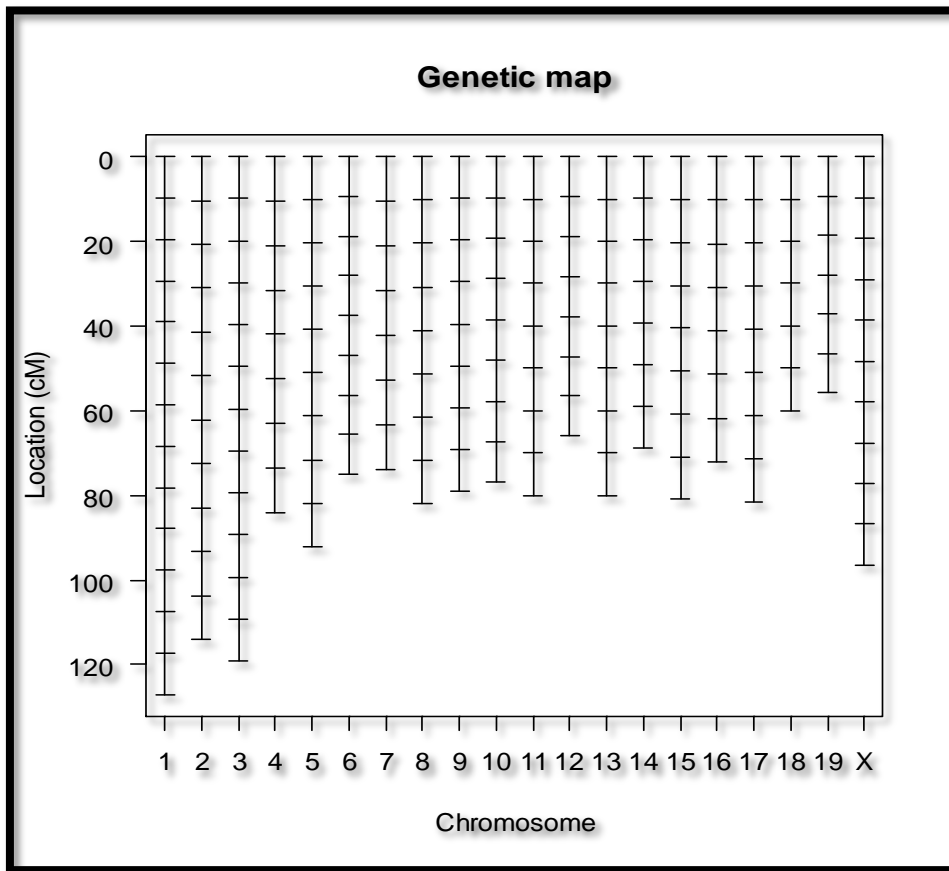
الخارطة الجينية المستخدمة في عملية التوليد لتقدير الـ(QTL) في التضارب العكسي Backcross

Autosomes	n.mar	length ave	spacing max
1	14	127	9.8
2	12	114	10.4
3	13	119.2	9.9
4	9	84	10.5
5	10	92	10.2
6	9	75	9.4
7	8	74	10.6
8	9	82	10.2
9	9	79	9.9
10	9	77	9.6
11	9	80	10
12	8	66	9.4
13	9	80	10
14	8	69	9.9
15	9	81	10.1
16	8	72	10.3
17	9	81.6	10.2

18	7	60	10
19	7	55.7	9.3
X	11	96.5	9.7
overall	187	1665	10

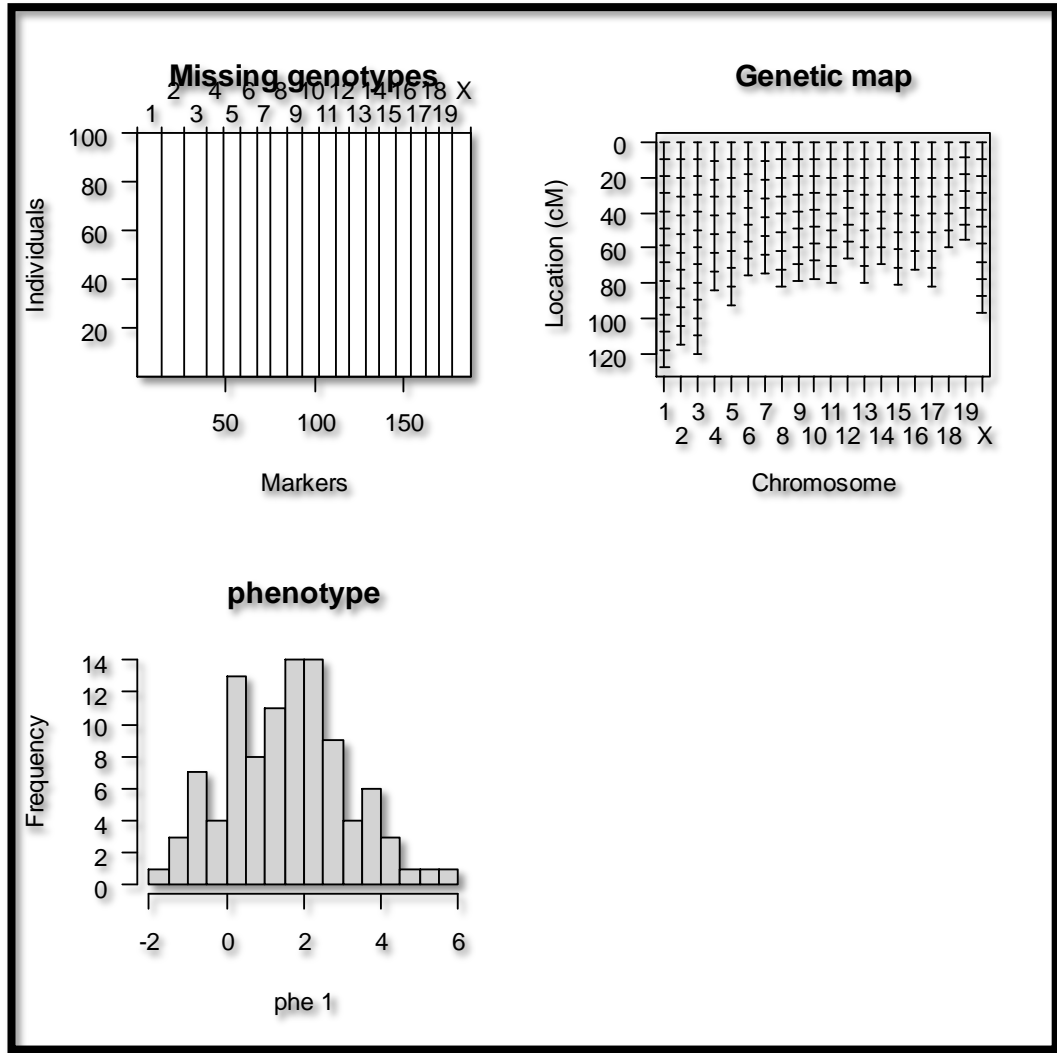
الشكل (1)

الخارطة الجينية المستخدمة في عملية التوليد التضارب العكسي Backcross



الشكل (2)

الخارطة الجينية المستخدمة في عملية التوليد التضارب العكسي Backcross



2.5 المقارنة بين طرائق تحليل مواقع الصفات الكمية

الجدول (2) يحتوي على أربعة أعمدة يمثل العمود الاول طرائق تحليل مواقع الصفات الكمية وهي (mr,em,hk,ehk)، أما العمود الثاني فيمثل الكروموسوم، في حين يمثل العمود الثاني موقع الواسمة المؤثرة (cM)، اما العمود الأخير فيمثل إحصاءة الاختبار (Lod score). إن النتائج المعروضة في الجدول (1) تمثل خلاصة تحليل الارتباطات المعنوية بين الواسمات ومواقع الصفات الكمية (QTL) لـ (20) كروموسوم بوجود (187) واسمة كلية.

1.2.5 انحدار الواسمات (Marker regression)

بالنسبة لانحدار الواسمات (Marker regression) فقد اظهر هذا الاسلوب أن هنالك (QTL) على الكروموسوم الاول ضمن الواسمة (DIM6) في الموقع (48.8) بقيمة lod مقدارها (4.32) وهي اكبر من قيمة العتبة الحرجة (3) والتي تؤشر على وجود (QTL) وكذلك وجود (QTL) على الكروموسوم الخامس ضمن الواسمة (D5M6) ايضاً الموقع (51.1) بقيمة lod مقدارها (7.91) وهي اكبر من قيمة العتبة الحرجة (3) والتي تؤشر على وجود (QTL) علماً بأن قيمة lod المتعلقة بالكروموسوم الخامس توفر دليلاً اكبر على وجود (QTL) وذلك لكون قيمتها اكبر من قيمة lod مع نظيرتها في الكروموسوم الاول .

2.2.5 انحدار أو خوارزمية EM Regression OR EM algorithm

بالنسبة (em) فقد اظهر هذا الاسلوب أن هنالك (QTL) على الكروموسوم الاول ضمن الواسمة (C1.loc47) في الموقع (47) بقيمة lod مقدارها (4.59) وهي اكبر من قيمة العتبة الحرجة (3) والتي تؤثر على وجود (QTL) وكذلك بوجود (QTL) على الكروموسوم الخامس ضمن الواسمة (D5M6) ايضاً الموقع (51.1) بقيمة lod مقدارها (7.91) وهي اكبر من قيمة العتبة الحرجة (3) والتي تؤثر على وجود (QTL) علماً بأن قيمة lod المتعلقة بالكروموسوم الخامس توفر دليلاً اكبر على وجود (QTL) وذلك لكون قيمتها اكبر من قيمة lod مع نظيرتها في الكروموسوم الاول .

3.5.2 انحدار هيلي نوت Haley- Knott regression

بالنسبة (hk) فقد اظهر هذا الاسلوب أن هنالك (QTL) على الكروموسوم الاول ضمن الواسمة (D1M6) في الموقع (48.8) بقيمة lod مقدارها (4.32) وهي اكبر من قيمة العتبة الحرجة (3) والتي تؤثر على وجود (QTL) وكذلك بوجود (QTL) على الكروموسوم الخامس ضمن الواسمة (D5M6) ايضاً الموقع (51.1) بقيمة lod مقدارها (7.91) وهي اكبر من قيمة العتبة الحرجة (3) والتي تؤثر على وجود (QTL) علماً بأن قيمة lod المتعلقة بالكروموسوم الخامس توفر دليلاً اكبر على وجود (QTL) وذلك لكون قيمتها اكبر من قيمة lod مع نظيرتها في الكروموسوم الاول .

4.5.2 انحدار هيلي نوت الموسع Extent Haley- Knott regression

بالنسبة (ehk) فقد اظهر هذا الاسلوب أن هنالك (QTL) على الكروموسوم الاول ضمن الواسمة (C1.loc89) في الموقع (92.3) بقيمة lod مقدارها (3.74) وهي اكبر من قيمة العتبة الحرجة (3) والتي تؤثر على وجود (QTL) وكذلك بوجود (QTL) على الكروموسوم الرابع ضمن الواسمة (D4Mit164) ايضاً الموقع (37.3) بقيمة lod مقدارها (8.6) وهي اكبر من قيمة العتبة الحرجة (3) والتي تؤثر على وجود (QTL) علماً بأن قيمة lod المتعلقة بالكروموسوم الرابع توفر دليلاً اكبر على وجود (QTL) وذلك لكون قيمتها اكبر من قيمة lod مع نظيرتها في الكروموسوم الاول . خلاصة القول أن طرائق تحليل (QTL) قد تساهل في الكشف عن اثنين من الـ (QTL) ومع ذلك يمكن القول بأفضلية طريق انحدار (ehk) لكونها حققت اعلى قيمة lod .

الجدول (2)

المقارنة بين طرائق (mr,em,hk,ehk) لتقدير الـ (QTL) في التضارب العكسي Backcross

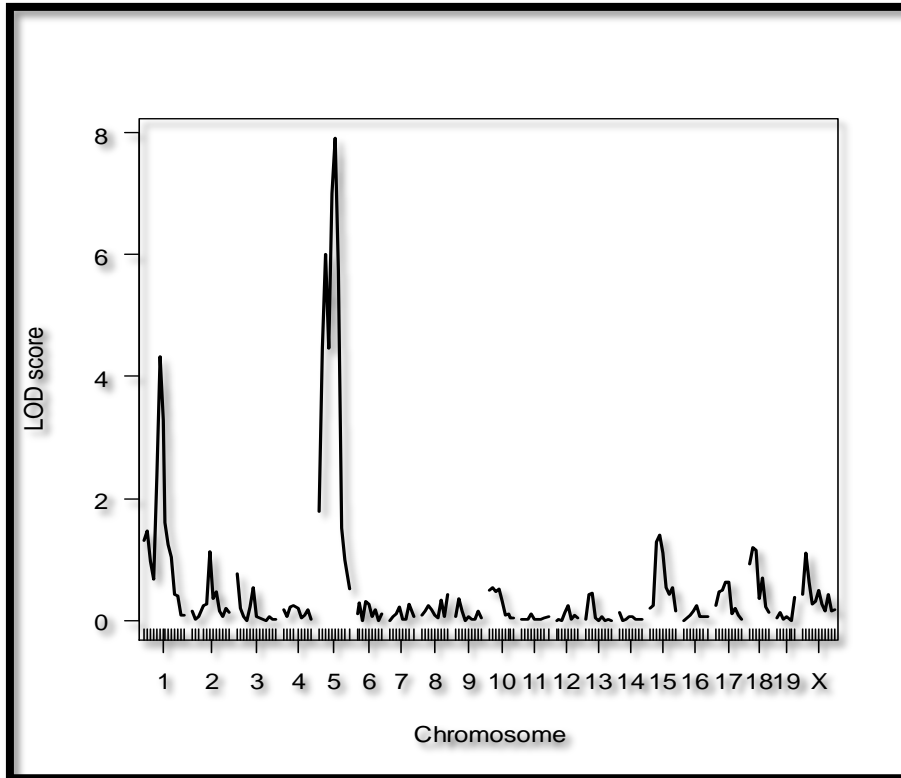
Method	Marker	chr	pos	lod
Mr	D1M6	1	48.8	4.32
	D5M6	5	51.1	7.91
EM	c1.loc47	1	47.0	4.59
	D5M6	5	51.1	7.91
hk	D1M6	1	48.8	4.32

	D5M6	5	51.1	7.91
ehk	c1.loc89	1	92.3	3.74
	D4Mit164	4	37.3	8.06

الاشكال (3) و(4) و(5) و (6) تبين الرسم البياني لطرائق تحليل (QTL) لقيم (LOD Score) لكل كروموسوم والتي تبين منها وفي جميع الاشكال عدم معنوية تشخيص الصفات الكمية على الكروموسومات ما عدا تلك التي تم استعراضها في الجداول؛ وذلك لكون قيمة (LOD Score) بالنسبة للكروموسوم الاول والخامس بالنسبة للطرائق الثلاث الاولى، والاول والرابع بالنسبة للطريقة الاخيرة كأتت اكبر من القيمة الحرجة للعتبة (3) لذلك تعد قيمة إحصاءة الاختبار المتعلقة بالكروموسومات المشار اليها في اعلاه معنوية تدل على أن طرائق التقدير المختلفة استطاع كل منها أن يكتشف اثنين من مواقع الصفات الكمية من المجموع الكلي (200) واسمة، تشير الرسوم البيانية أيضاً الى عدم معنوية إحصاءة الاختبار (LOD Score) الخاصة بكل كروموسوم من الكروموسومات المتبقية لكونها اقل من القيمة الحرجة للعتبة (3).

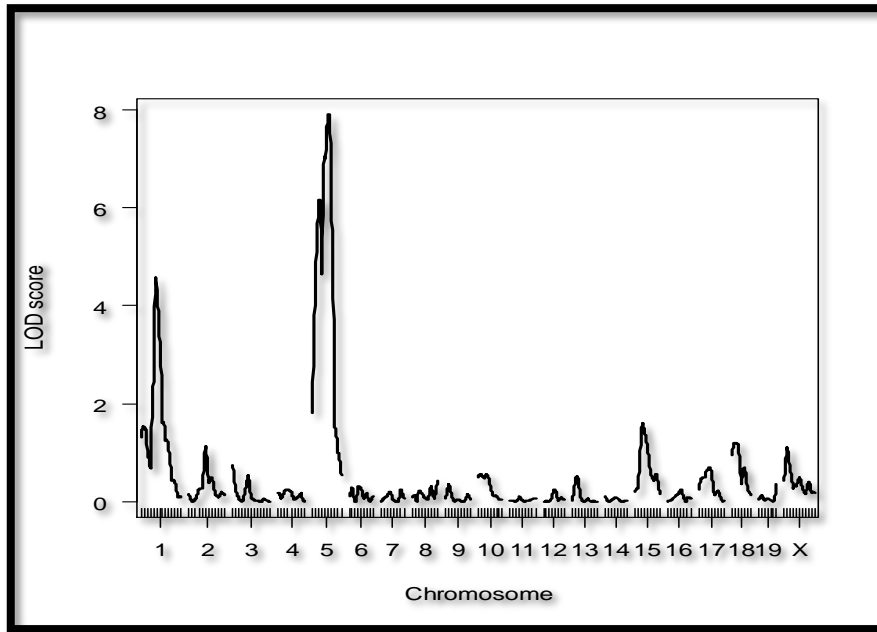
الشكل (3)

الرسم البياني لقيم (LOD Score) بطريقة Mr لتحليل (QTL) في التضارب العكسي Backcross



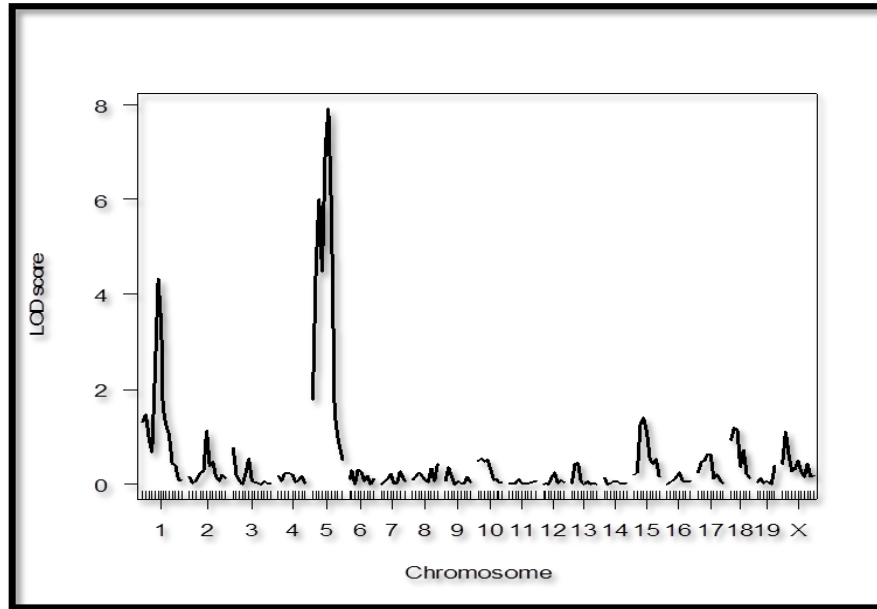
الشكل (4)

الرسم البياني لقيم (LOD Score) بطريقة (EM) لتحليل (QTL) في التضارب العكسي Backcross



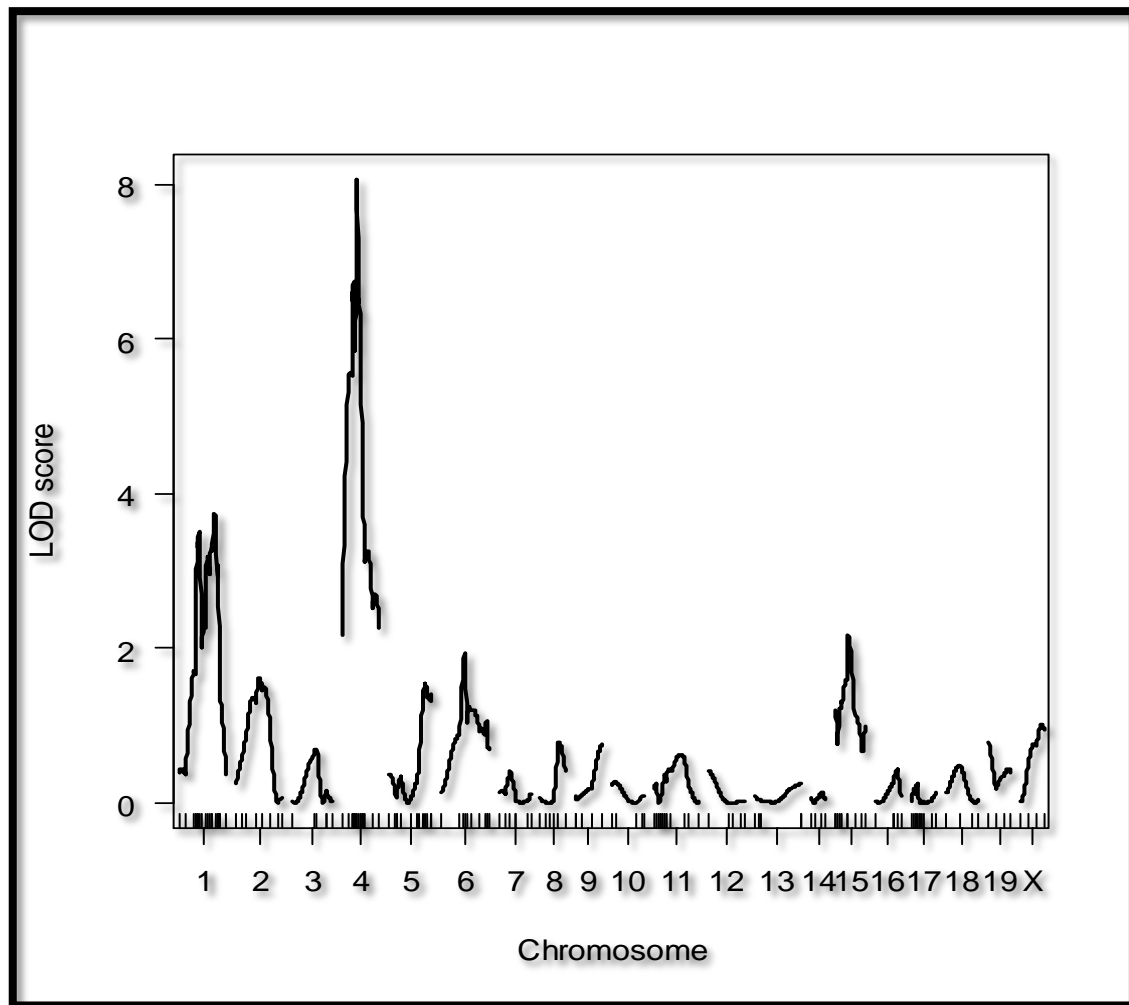
الشكل (5)

الرسم البياني لقيم (LOD Score) بطريقة (hk) لتحليل (QTL) في التضارب العكسي Backcross



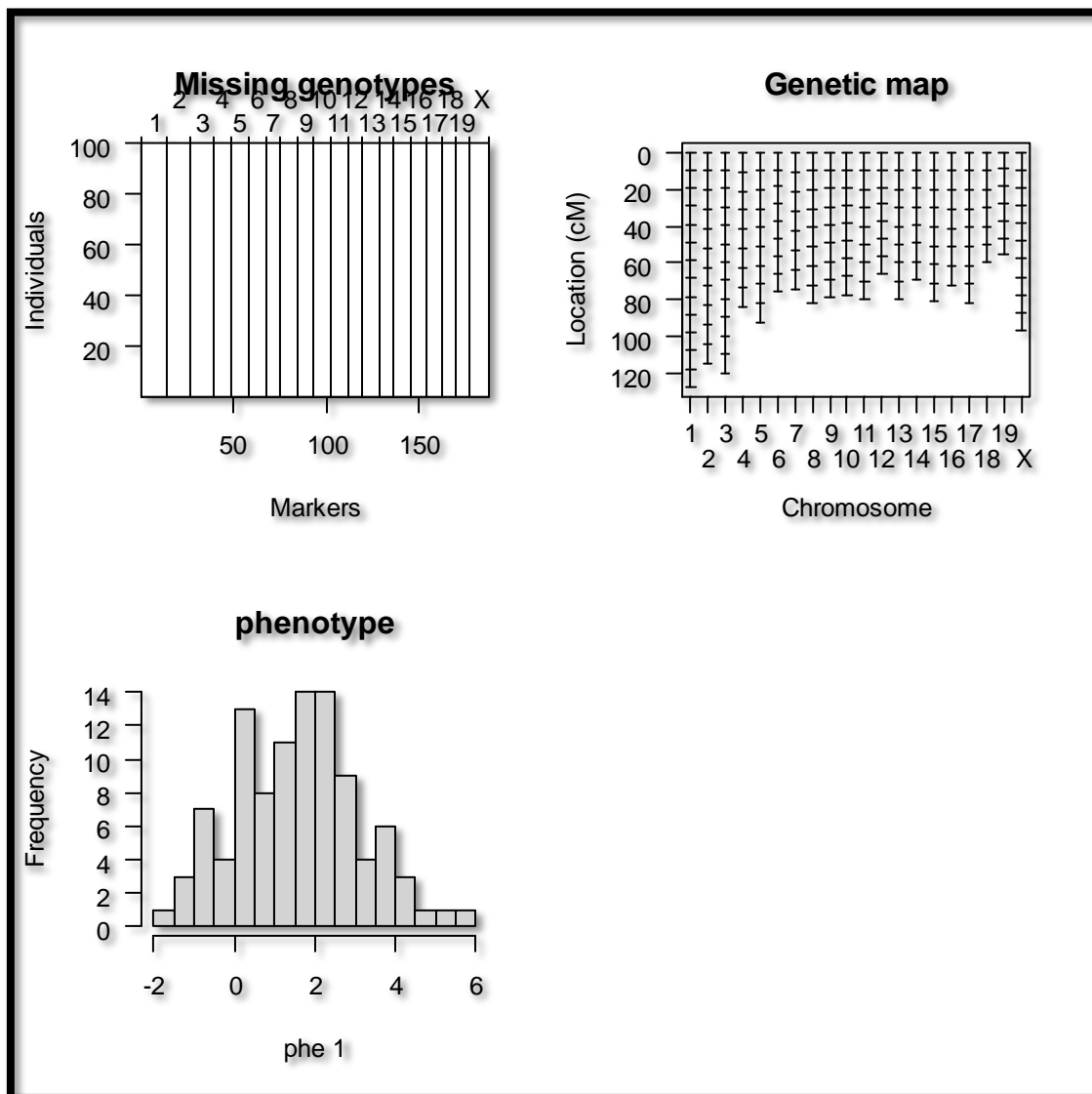
الشكل (6)

الرسم البياني لقيم (LOD Score) بطريقة (ehk) لتحليل (QTL) في التضارب العكسي Backcross



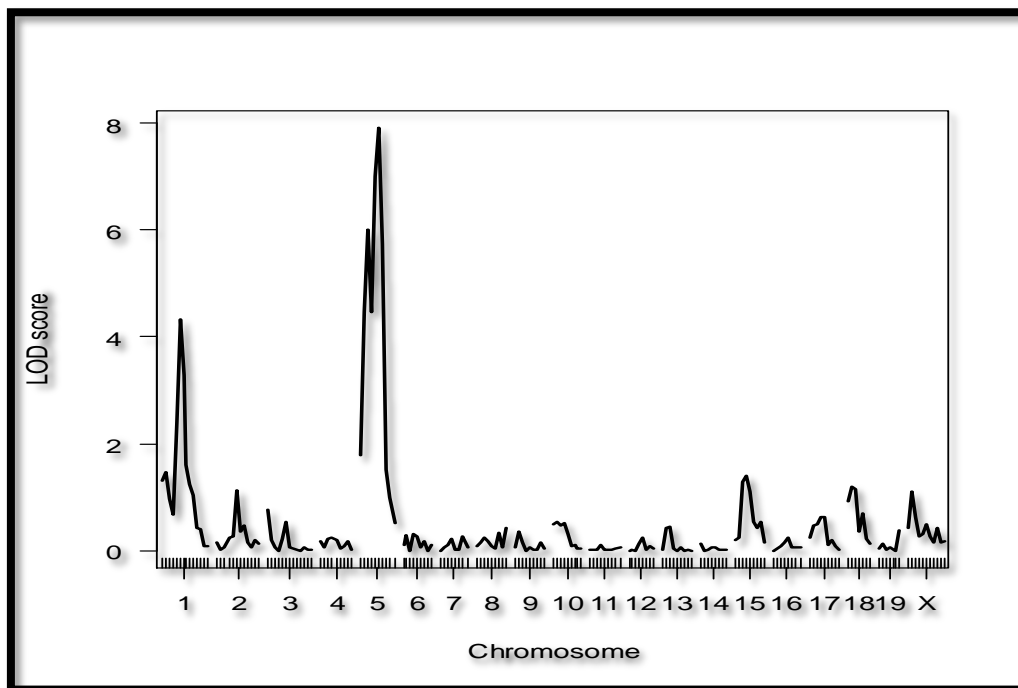
(7) الشكل

الرسم البياني لتوليد الخارطة الجينية لانحدار السمات



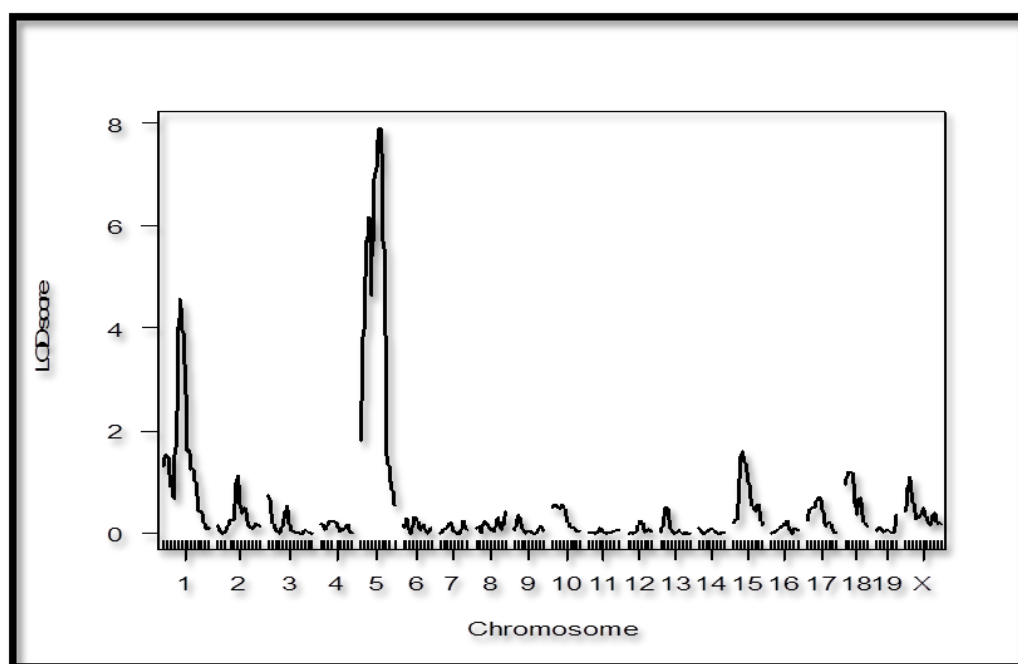
الشكل (8)

الرسم البياني لقيم (LOD Score) بطريقة Mr لتحليل (QTL) في انحدار السمات



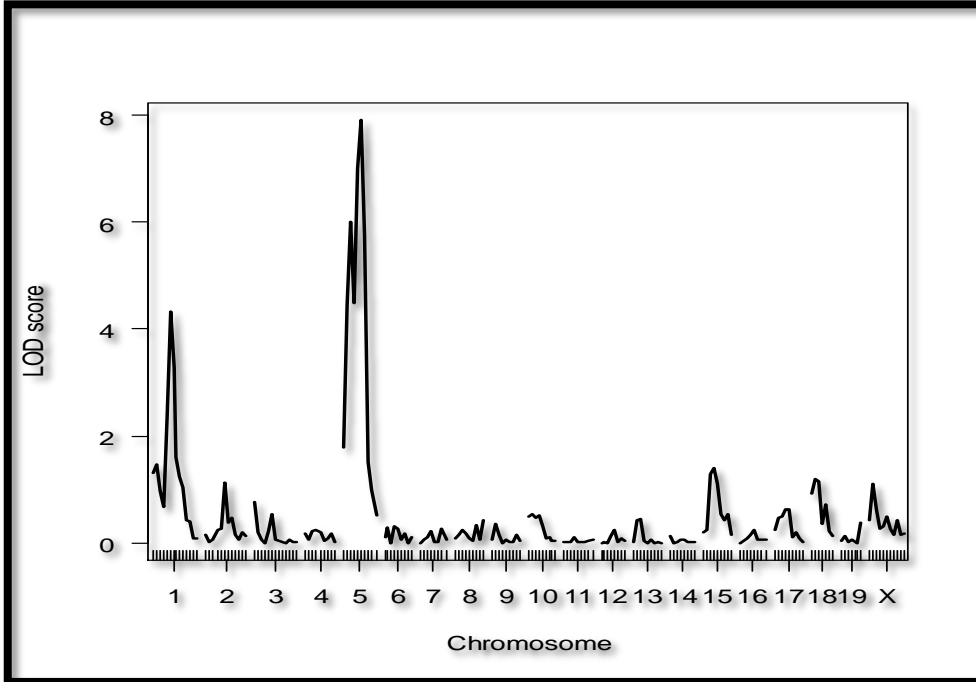
الشكل (9)

الرسم البياني لقيم (LOD Score) بطريقة EM لتحليل (QTL) في انحدار السمات



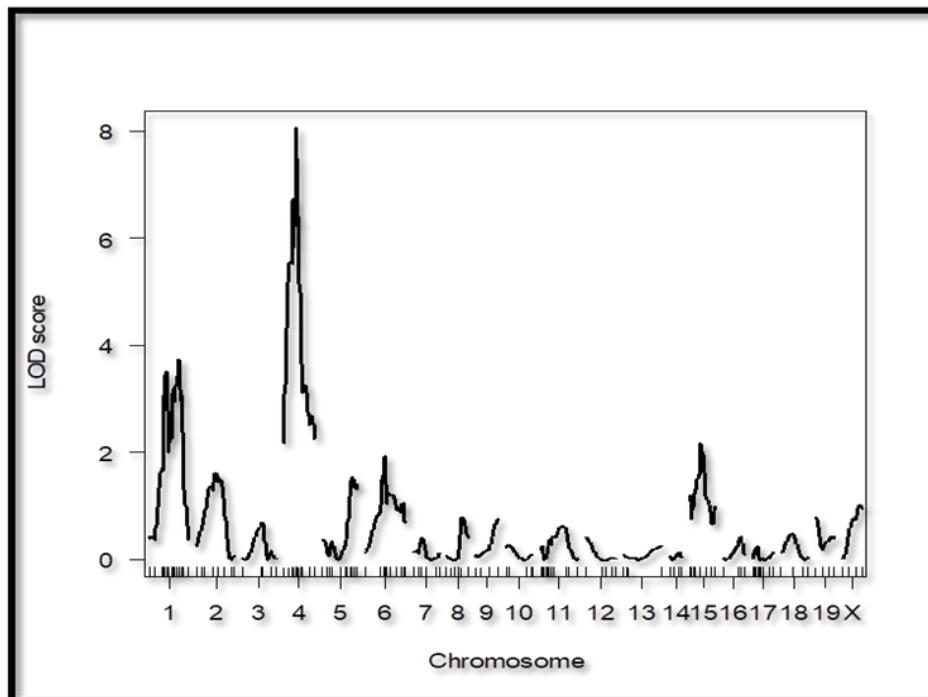
الشكل (10)

الرسم البياني لقيم (LOD Score) بطريقة hk لتحليل (QTL) في انحدار السمات



الشكل (3.41)

الرسم البياني لقيم (LOD Score) بطريقة Ehk لتحليل (QTL) في انحدار السمات



الاستنتاجات

عند استعمال طرائق التحليل لمواقع الصفات الكمية (mr,em,hk,ehk) (QTL) باستخدام التضارب العكسي (Backcross) فان هذه الطرائق لتحليل (QTL) قد تساوت في الكشف عن اثنين (QTL) ومع ذلك يمكن القول بأفضلية طريقة انحدار (ehk) لكونها حققت اعلى قيمة lod .

التوصيات :

1. يمكن استعمال طرائق التحليل لمواقع الصفات الكمية (QTL) (mr,em,hk,ehk) باستخدام التضارب الداخلي (intercross) فان هذه الطرائق لتحليل الـ (QTL) .
2. 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 .

المصادر :

1. East. E.M. 1908. In breeding in corn. P. 414-212.
2. Haley, C.S. and S.A. Knott. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69:315-324.
3. Hill, W. G. 1990. Considerations in the design of animal breeding experiments. In D. Gianola and K. Hammond (eds.), Advances in statistical methods for genetic improvement of livestock, pp. 59-76. Springer-Verlag , Berlin (17).
4. Hill, W.G.1970. Design of experiments to estimate heritability by regression of offspring on selected parents. Biometrics 26: 566-571 [17].
5. Julius Van der Werf,2000, Chapter 9 , "method for QTL analysis" .
6. Kao, 2000,"On the differences between maximum likelihood and regression interval mapping in the analysis of quantitative trait loci". Genetics 156, 855–865.
7. Mskay & Falconer ,1996, "Introduction to Quantitative Genetics" ,fourth edition , Addison Wesley Longman , Harlow ,Essex, UK .
8. Whittaker , J.C. , Thompson ,R ., & Visscher , 1996, "On the mapping of QTL by regression of phenotype on marker Type ".Heredity 77:23-32 .
9. Wright, Sewall (1951). "The genetical structure of populations". Annals of Eugenics. 15: 323–354.