

以 BLAT 為基礎的染色體轉位分析平台

鄭煜輝

國立高雄應用科技
大學電子工程系
yuhuei.cheng
@gmail.com

谷德倫

高雄醫學大學生物
醫學暨環境生物系
ed0958719325
@yahoo.com.tw

楊正宏

國立高雄應用科技
大學電子工程系
chyang
@cc.kuas.edu.tw

張學偉

高雄醫學大學生物
醫學暨環境生物系
changhw
@kmu.edu.tw

莊麗月

義守大學
化學工程系
chuang
@isu.edu.tw

摘要

染色體的一個片段由某一個位置移到另一個位置，稱為染色體轉位(Chromosome Translocation)。發生染色體轉位可能會造成慢性骨髓性白血病(chronic myeloid leukemia; CML)、伯奇氏淋巴瘤(Burkitt's lymphoma; BL)、腎臟亮細胞腫瘤(clear cell sarcoma of the kidney; CCSK)或其他疾病等。本研究以 UCSC BLAT 為基礎，建構了一個 web-based 的染色體轉位分析平台，由 UCSC BLAT 來提供快速的序列比對，並利用重組與分析，將單一物種或跨物種(包括 human, mouse 和 rat)序列中的染色體轉位資訊找出。最後，經由單一物種或跨物種不同染色體中的序列組合並測試，測試結果顯示本研究確實能分析出單一物種或跨物種序列中具有染色體轉位的情況。

關鍵字：染色體轉位(Chromosome Translocation)，UCSC BLAT，web-based，跨物種染色體轉位

Abstract

One part of chromosome from somewhere move to another place called 'chromosome translocation'. This condition will cause chronic myeloid leukemia (CML), Burkitt's lymphoma (BL), clear cell sarcoma of the kidney (CCSK) or other diseases possibly. This study was based on UCSC BLAT and constructed a web-based analysis platform for chromosome translocation. It used UCSC BLAT to blast sequence rapidly and employed recombination and analysis to find the information of chromosome translocation at single organism or cross-species including human, mouse and rat organisms. Finally, the test using different chromosome sequences between single organism and cross-species are performed.

The test results shown this system indeed found the conditions of chromosome translocation at single organism or cross-species.

Keyword: chromosome translocation, UCSC BLAT, web-based, cross-species translocation

1、前言

染色體構造上的異常會由物理或化學誘導因子誘導產生，也會自然地發生，而染色體轉位(Chromosome Translocation)便是染色體異常的情形之一，主要是由於染色體的一個片段由某一個位置移到另一個位置稱之。染色體轉位主要分為兩大類，(1)相互轉位(reciprocal translocation)和(2)非相互轉位(nonreciprocal translocation)。相互轉位主要是由於兩條非同源染色體間某一段序列片段互相交換，而非相互轉位主要是染色體中的某一段序列發生轉移，轉移到另一條非同源染色體上，但是並不互相交換。發生染色體轉位會造成細胞學上的變化，正常細胞減數分裂的粗絲期，同源染色體配對是呈二價體，但倘若在兩對同源染色體間發生轉位，轉位的染色體會配對呈四價體(quadrapie)。並且發生染色體轉位可能會造成外表型的變化，若轉位失去基因，外表可能不正常，倘若外表正常，但因減數分裂時會發生配對的困難，其後代外表會不正常。發生染色體轉位可能會造成慢性骨髓性白血病(chronic myeloid leukemia; CML)，而費城染色體(Philadelphia chromosome; PH)是 CML 的疾病指標，出現在90%以上的CML病例中，由於第9條染色體和第22條染色體的相互轉位，導致 Abelson oncogene (ABL) 的 3' 部份從 9q34 轉移到 breakpoint cluster region (BCR) 22q11 的 5' 部份上，產生一個異常的融合基因 (BCR-ABL) [1][2][3]。伯奇氏淋巴瘤(Burkitt's lymphoma)是由於轉位主要發生在染色體 8q24 (c-myc)

和 Ig gene 染色體中的一個，如 14q32 (Ig H gene), 2p12 (Ig kappa) 或 22q11 (Ig lambda)。這種轉位被估計大約有 80% 在 t(8;14), 15% 在 t(2;8) 及 5% 在 t(8;22)[4][5]。腎臟亮細胞腫瘤(Clear cell sarcoma of the kidney)被研究發現有平衡轉位在 t(10;17)(q22;p13) and t(10;17)(q11;p12)[6]。

由於染色體轉位跟許多疾病息息相關，為了幫助生物學家分析序列中的染色體轉位資訊，本研究以 UCSC BLAT[8]為基礎，建構了一個 web-based 的染色體轉位分析平台，它能夠將使用者輸入的序列切成多個序列片段，然後將這些序列切片段利用 UCSC BLAT 快速的比對相關物種序列，取得序列片段的相關資訊，並利用重組與分析將序列中的染色體轉位資訊找出。

2、研究方法

(1) 系統架構

本系統利用 JAVA 程式語言實作，整個系統架構以 MVC (Model-View-Controller) 架構設計，如圖 1 所示。系統分為以 Model 為主的(1)輸入序列切片處理、(2)UCSC BLAT 搜尋和(3)轉位分析；(1)輸入序列切片處理主要是將使用者輸入的序列依照系統或使用者設定的長度做分割。(2)UCSC BLAT 搜尋則利用 BLAT 演算法快速比對的特性，將分割的序列片段位置比對出來。(3)轉位分析則是用來判斷是否發生染色體轉位。以 View 為主的(1)序列輸入及物種選擇介面和(2)輸出分析結果；(1)序列輸入及物種選擇介面主要提供使用者選擇物種、資料庫版本以及輸入欲分析的序列。(2)輸出分析結果主要呈現最後分析後的資訊，包括序列是否發生染色體轉位、序列比對到的染色體位置及 cytoband 位置等。

以及以 Controller 為主的切片段序列連續性判別及重組；切片段序列連續性判別及重組主要是用來判別分割序列所比對到的染色體位置是否具有連續性並重新

做排列組合。

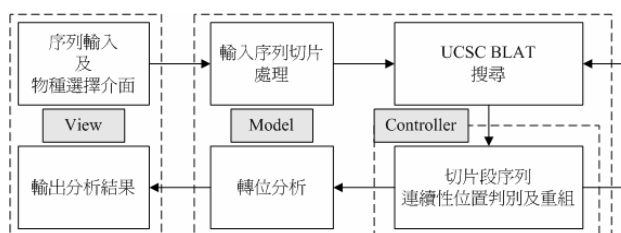


圖 1. 以 MVC (Model-View-Controller) 為主之系統架構圖

(2) 系統流程圖

染色體轉位分析流程如下圖 2 所示，主要分為兩部分，第一部分為單一物種之染色體轉位分析，第二部份為跨物種之染色體轉位分析。單一物種之染色體轉位分析主要經由序列輸入將此輸入序列切成多個序列片段，然後將這些序列切片段利用 UCSC BLAT 去比對單一物種序列，取得序列片段的相關資訊，包括 chromosome, chromosome position, cytoband 等，接著判斷這些片段當中是否具有連續性的 chromosome position，如果有的話就依照 chromosome position 將片段序列重組，沒有的話則不必重組的動作。接著找出片段序列在輸入序列中具有連續的部份，並合併此連續的序列，將其利用 UCSC BLAT 重新比對，找出專一性更高的序列，最後判別所有片段序列中是否存在多條 chromosomes，如果有的話即發生染色體轉位，沒有的話即是染色體正常，最後將所有比對到的片段及判別的結果輸出。而跨物種之染色體轉位分析主要利用單一物種之染色體轉位分析來一一分析比對所有物種之切片段，並且混合各個物種之切片段，接著從這些混合的序列片段中，找出較長序列的片段，並將小片段序列一一吸收，以獲得專一性更高的序列，最後重組所有物種的連續片段，並將結果資訊輸出。

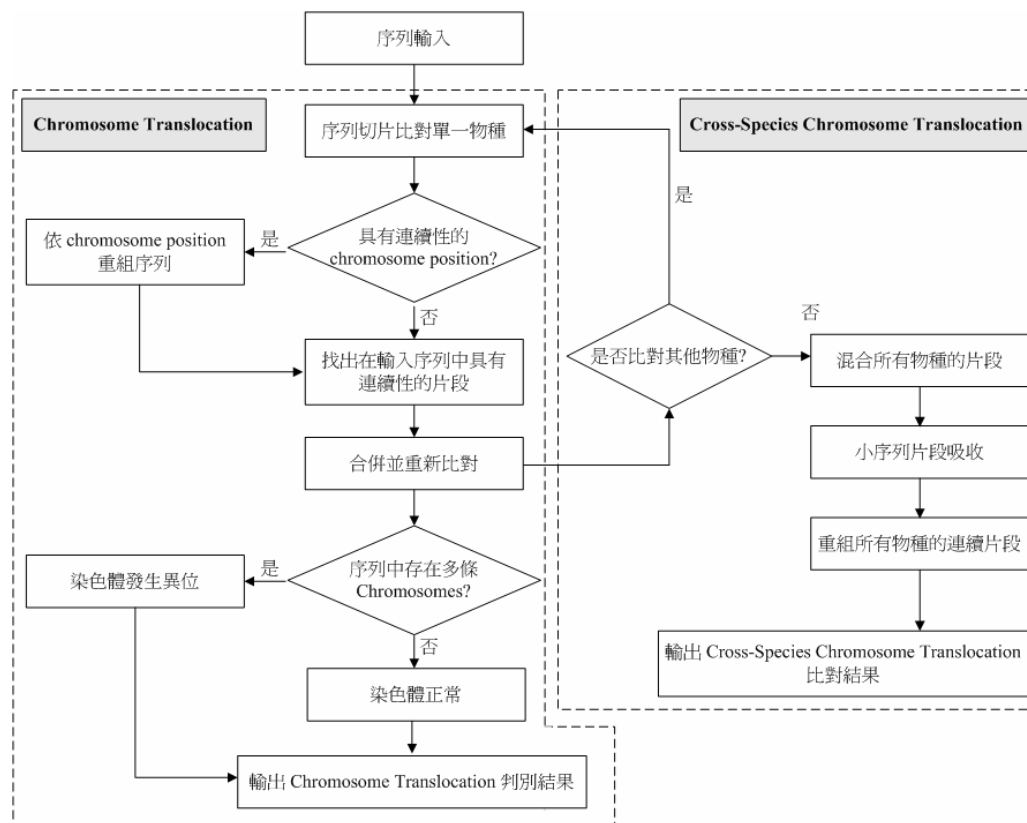


圖 2. 染色體轉位分析流程圖

(3) 序列切片段的選取

由於發生染色體轉位，序列並無法完整被比對到，因此，我們將序列切成多個片段來一一比對，以減少無法比對到的風險。然而，切成多個序列片段有個需要考量的因素，即太長或太短的序列切片段都會影響其比對的過程及結果，太長的切片段雖然其比對專一性較高，然而會降低其比對速度；太短的切片段會造成比對專一性不足，而比對到多條染色體，因而造成誤判，在此，本研究序列切片段的長度預設為 100bps，使用者可自行調整。

(4) 序列切片段連續性的判別

倘若，序列切片具有染色體位置的連續性，則必須將這些序列片段重組合併，對於沒有染色體轉位的序列而言，利用序列切片比對出來的結果會呈連續性的序列，如圖 3 所示，可經由有比對到的序列切片段且對應輸入序列的位置來判別其連續性，公式如下：

$$\text{segment}_{N_end} + 1 = \text{segment}_{N+1_start},$$

$$\text{where } N < M, M = \text{segment size} \quad (\text{公式 1})$$

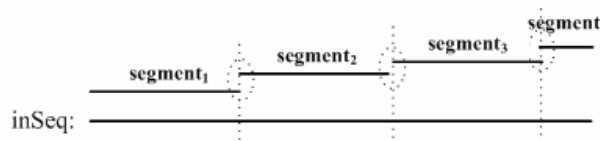


圖 3. 沒有染色體轉位的序列

(4) 輸入序列中的連續性及合併

由於本研究方法將序列切成多個序列片段，在輸入序列中連續的序列片段不一定是連續的染色體位置，因此，比對之後必須將輸入序列中連續的切片段重新合併，以提高其專一性，然後再重新送至 Blat 比對，提高專一性的結果將比對到更相似的序列，最後針對染色體位置以及切片段在序列中的位置來做分析。

(5) 染色體轉位的判別

然而，當利用序列切片比對時，序列切片段有部份序列沒有被比對到的話，很可能發生了染色體轉位，如圖 4 所示，在序列切片段 1 (segment1) 有部份序列沒有被比對到，因此，序列切片段 2 (segment2) 延續沒有比對到的序列部份繼續比對，以完成整個的比對，倘若比對的結果發現，有兩個或以上的染色體序列被找出，即可斷定此序列發生了染色體轉位，否則仍只是正常的染色體序列。

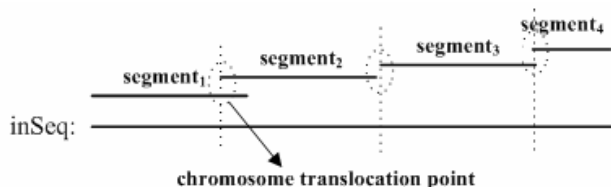


圖 4. 發生染色體轉位的序列

(6) Cross-Species 小序列片段吸收原理

當一條序列拿到多個物種 genome 做比對時，對於不同物種所比對到的序列長短都不盡相同。對於大部分物種只會比對到部份序列而不是完整比對到此序列，或許其中會因變異而有部分比對不正確，但大體來說，在與此序列相同物種的比對下，所比對到的序列一定會比在其他物種比對到的序列來的完整。為了找出片段序列是屬於何物種的序列，我們必須從同物種間具有較高比對專一性著手，因此，在混合物種的序列當中，長序列片段一定比對短序列片段具有更高的專一性，如果短序列片段是屬於長序列片段的子序列，那麼短序列片段將被長片段序列吸收，如圖 5 所示，便可以從中分離出物種。

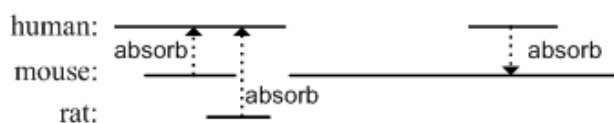


圖 5. Cross-Species 小序列片段吸收

3、結果與討論

本系統輸入介面如圖 6 所示，單一物種之染色體轉位分析提供 human、mouse 和 rat 三個物種單選及 UCSC Assembly 和切片斷長度的大小供使用者選擇，並提供一序列輸入方塊，讓使用者輸入要分析之序列。而跨物種之染色體轉位分析提供 human、mouse 和 rat 三個物種多選及各物種之 UCSC Assembly 和切片斷長度的大小供使用者選擇，同樣也提供了一序列輸入方塊，讓使用者輸入要分析之序列。

圖 6. 單一物種和跨物種染色體轉位分析輸入介面

我們利用本系統的單一物種之染色體轉位分析測試 ACCESSION AB118156 的序列，可得 AB118156 序列無染色體轉位，比對到的染色體為 chr17，且位置為 7519096 ~ 7519279，cytoband 為 p13.1，如圖 7 所示。

Organism: human				
UCSC Assembly: Mar. 2006				
Least length for traslocation: 100 bps				
This sequence is not translocation.				
Input Sequence				
000001 TACTGCGCTG CCTCAAGAA GATGTTTTCG CAAGTGGGCA AGAGCTGGCC				
000051 TGTGAGGCTG TGGGTTGATT CCACAGCCGC GCGCGGAC CCAGTTCGGG				
000101 CCATGGGCAT CTACAGCAG TCACAGACCA TGACGGAGGT TGTGAGGCGC				
000151 TGCCGCCACC ATGAGGCGCT CTCAGATAGC GATG				
Sequence	Sequence Position	Chr Position	Chr	Cytoband
000001 TACTGCGCTG CCTCAAGAA GATGTTTTCG CAAGTGGGCA AGAGCTGGCC				
000051 TGTGAGGCTG TGGGTTGATT CCACAGCCGC GCGCGGAC CCAGTTCGGG	1-184	7519096-7519279	chr17	p13.1
000101 CCATGGGCAT CTACAGCAG TCACAGACCA TGACGGAGGT TGTGAGGCGC				
000151 TGCCGCCACC ATGAGGCGCT CTCAGATAGC GATG				

圖 7. AB118156 無染色體轉位

接著，我們將不同染色體的序列組合成一條序列，然後送至我們系統的單一物種之染色體轉位分析測試，測試的情況如下所述：

(1) 首先，我們隨機組合 human 的 chr8、chr21、chrY、chr13 和 chr15 五條染色體序列成一條序列。測試結果我們發現，本系統找出 human 染色體轉位分別為(1) chr8，染色體位置在 19855098 ~ 19855166，cytoband 為 p21.3，(2) chr21，染色體位置在 15817047 ~ 15817124，cytoband 為 q21.1，(3) chrY，染色體位置在 4138017 ~ 4138090，cytoband 為 p11.2，(4) chr13，染色體位置在 31347260 ~ 31347327，cytoband 為 q13.1，和(5) chr15，染色體位置在 56621856 ~ 56621934，cytoband 為 q22.1，如圖 8 所示，與我們輸入的組合不謀而合。

Organism: human
UCSC Assembly: Mar. 2006
Least length for traslocation: 100 bps

This sequence is translocation.

Input sequence

```
000001 AAGGAGGAGT TTAACATACC TCTGGACAAT GTCCATCTCT TGGGATACAG
000051 CTTTGAGGCC CATCTGCTTA ATAAAGACCC AAAATCTTAT TTATATATA
000101 TTTTATGAAA AACTGGTTAG CACACAGATA ATTACATAG AGCTGGTTGG
000151 GTTATATGTA AAGGAGGACA TCCAAAGCCT TAGGCAATC GGAATGTTGG
000201 ATTATGACTC TTAGACCTAC TAGTATAAGT TGTTCATCTG TGAATCATAT
000251 CGATGGCTGA ATGATATTCG AAATATATCT ATATTTTAC CAGACAGAGT
000301 CAGCAGCTTG AGAGATGATG TTACGTTTGG AGTGTGAATA AGCAGCTTCA
000351 TTTTGAAGCA GGCAGGAA
```

Translocation Chromosomes

chr8 - chr21 - chrY - chr13 - chr15 -

Translocation Sequences	Sequence Position	Chr Position	Chr	Cytoband
000001 AAGGAGGAGT TTAACATACC TCTGGACAAT GTCCATCTCT TGGGATACAG	1 ~ 69	19855098 ~ 19855166	chr8	p21.3
000051 CTTTGAGGCC CATCTGCTTA ATAAAGACCC AAAATCTTAT TTATATATA	70 ~ 147	15817047 ~ 15817124	chr21	q21.1
000001 AATAAAGACC GAAATCTTIA TTTTATATAT ATTTTATGGA AAACGTGTTA	148 ~ 221	4138017 ~ 4138090	chrY	p11.2
000001 TGTGTTTAAAT CTAAAGGAGG GGAATCCAAAG GCTTAGGAGG ATTGGAGTGG	222 ~ 289	31347260 ~ 31347327	chr13	q13.1
000001 AATATACTCT TAATTTTAA GTATCATGTC GATGGCTGAA TGTATTTCCA	290 ~ 368	56621856 ~ 56621934	chr15	q22.1
000001 CAGCAGGAGG TCAGCAGCTG GAGAGATTAG TTTCAGTTGG GAGTGTGAAT				
000051 AAGCAGCTCT ATTTCTGAGC AGGACAGAA				

圖 8. human 染色體轉位測試

(2) 接著，我們隨機組合 mouse 的 chr11、chr2 和 chr18 三條染色體序列成一條序列。由單一物種之染色體轉位分析測試結果我們發現，本系統找出 mouse 染色體轉位分別為(1) chr11，染色體位置在 63827710 ~ 63827773，cytoband 為 qB3，(2) chr2，染色體位置在 172188284 ~ 172188342，cytoband 為 qH3 和(3) chr18，染色體位置在 27717934 ~ 27717998，cytoband 為 qA2，如圖 9 所示，此測試結果仍舊與我們輸入的組合不謀而合。

Organism: mouse
UCSC Assembly: Feb. 2006
Least length for traslocation: 100 bps

This sequence is translocation.

Input sequence

```
000001 AAGCTAATCT ATTAGTTTIA AATGCTCTCT GGCATTACTT AAAACAGAAA
000051 AATATAAAGG CTAGTCTTAA AGCTTGGAGG CACTGACCTA CTCTCTGCTG
000101 AAGCTATCTT GTGGAGCCTC CACACAGCTC AAGCTAGAGC AGACTGATAA
000151 GTCTGTTTGA AGGGAGACTC CTGTGAGGCC GAATTTCA
```

Translocation Chromosomes

chr11 - chr2 - chr18 -

Translocation Sequences	Sequence Position	Chr Position	Chr	Cytoband
000001 AAGCTAATCT ATTAGTTTIA AATGCTCTCT GGCATTACTT AAAACAGAAA	1 ~ 64	63827710 ~ 63827773	chr11	qB3
000051 AATATAAAGG CTAGTCTTAA AGCTTGGAGG CACTGACCTA CTCTCTGCTG	65 ~ 123	172188284 ~ 172188342	chr2	qH3
000001 AAGCTAATCT ATTAGTTTIA AATGCTCTCT GGCATTACTT AAAACAGAAA	124 ~ 188	27717934 ~ 27717998	chr18	qA2

圖 9. mouse 染色體轉位測試

(3) 然後，我們隨機組合 rat 的 chr1、chr10、chr5、chr16、chr8 和 chr3 六條染色體序列成一條序列。測試結果我們發現，本系統找出染色體轉位分別為(1) chr1，染色體位置在 82569087 ~ 82569138，cytoband 為 p21，(2) chr10，染色體位置在 40384150 ~ 40384211，cytoband 為 q22，(3) chr5，染色體位置在 169501741 ~ 169501803，cytoband 為 q36，(4) chr16，染色體位置在 71526726 ~ 71526796，cytoband 為 q12.4，(5) chr8，染色體位置在 30184743 ~ 30184804，cytoband 為 q13，和(5) chr3，染色體位置在 145885831 ~ 145885904，cytoband 為 q42，如圖 10 所示，與我們輸入的組合完全一致。

Organism: rat
UCSC Assembly: Nov. 2004
Least length for traslocation: 100 bps

This sequence is translocation.

Input sequence

```
000001 AACACCATCT CTGTCTCTGG GTCCCGTCAT TTCACCACTG GAGCTTATT
000051 TCGACATCTG ACACAGCTTA AATCAGGAGC GGAATCTGTC GAGGATGAGA
000101 AGCTAGGCTG GCGGAGGAGT ATAAATAAG CTTTGGTACA CAGGTTTGG
000151 GAGAGGCA CAGCTACAC TCACCACAC AAATATATAT TTTTAAAGAC
000201 TAATGAATAT CAGTTAATAT ACCATTATTT AAATCTTTT ATCAAGAAC
000251 CTTAATTCAC GTGGATCCCG TAGTTTACTG TAGCCAGGCC AGAGCTACAC
000301 AGGCTTACTG CACTGAGCC AGTCTCAACA CAACACACCC AACAGAGATG
000351 CACTCAACTT CTGTGTTTCA TTGGAACCTA GGTG
```

Translocation Chromosomes

chr1 - chr10 - chr5 - chr16 - chr8 - chr3 -

Translocation Sequences	Sequence Position	Chr Position	Chr	Cytoband
000001 AACACCATCT CTGTCTCTGG GTCCCGTCAT TTCACCACTG GAGCTTATT	1 ~ 52	82569087 ~ 82569138	chr1	q21
000051 TCGACATCTG ACACAGCTTA AATCAGGAGC GGAATCTGTC GAGGATGAGA	53 ~ 114	40384150 ~ 40384211	chr10	q22
000001 AGGATATATA ATAAAGCTTT GGTGACAGAG TTTTGGAGAC AGGCCACAGG	115 ~ 177	169501741 ~ 169501803	chr5	q36
000001 AACAAATAT TATTTTAAAA AACTAATGAA ATTCACTTAA TATACCTATT	178 ~ 248	71526726 ~ 71526796	chr16	q12.4
000001 ACCTCAATTC AGCTGATCCG GGTAGTTTAG TGTAGGAGG CAGAGCTAC	249 ~ 310	30184743 ~ 30184804	chr8	q13
000001 CACTTCAGCC AGTCTCAACA CAACACACCC AACAGAGATG CACTCAACTT	311 ~ 384	145885831 ~ 145885904	chr3	q42
000051 GTTGTGTTCA TTGGAACCTA GGTG				

圖 10. rat 染色體轉位

最後，我們將不同物種之染色體的序列組合成一條序列，然後送至我們系統的跨物種之染色體轉位分析測試，測試的情況如下所述：

(1) 我們將 human 和 mouse 混合的序列送至我們的系統分析，分析的結果如下圖 11 所示，然而此分析結果卻尚未將比對的結果做整合性的分析，藉由“Analyze Cross Translocation”按鈕，我們可以得到跨物種轉位整合性分析的結果，如圖 12 所示，找出跨物種染色體轉位分別(1) human chr21，染色體位置在 15817004 ~ 15817098，cytoband 為 q21.1 和 (2) mouse chr19，染色體位置在 24967089 ~ 24967167，cytoband 為 qB。

Cross organisms

1. Organism: human
2. Organism: mouse
Least length for traslocation: 100 bps

UCSC Assembly: Mar. 2006
UCSC Assembly: Feb. 2006

Input sequence

```
000001 TTCCTATCTC CCATCAGCTT TATGAATAC TGAATCATTT AAAAATAAAA
000051 ACCGAAATCT TTATTTATTA TATATTTTAT TGAATCATTT TTAGCCCGCC
000101 CCCTGATCTC CCATCAGCTT CACTACCTCT TCTGCGGGCC CCGGCGCATG
000151 GCGACACCC TAGAAGAGCG GAGGCGCGCC ACC
```

This segment sequence information for human.

Segment Sequence	Sequence Position	Chr Position	Chr	Cytoband
000001 TTCCTATCTC CCATCAGCTT TATGAATAC TGAATCATTT AAAAATAAAA	1 ~ 95	15817004 ~ 15817098	chr21	q21.1
000051 ACCGAAATCT TTATTTATTA TATATTTTAT TGAATCATTT TTAGCCCGCC	148 ~ 173	80506001 ~ 80506026	chr5	q14.1

This segment sequence information for mouse.

Segment Sequence	Sequence Position	Chr Position	Chr	Cytoband
000001 GAATCATTTA AAAAATAAAA CCGAAATCTT TATTTATTT AT	32 ~ 73	104222336 ~ 104222363	chr6	qE1
000001 CCGGCGCGCT CATCGCATC CCATGCGCTA CTTGCTGCTC GCGGCGCGCC	96 ~ 173	24967089 ~ 24967167	chr19	qB

Analyze Cross Translocation

圖 11. human 和 mouse 混合的序列分析結果

Cross organisms

1. Organism: human
2. Organism: mouse
Least length for traslocation: 100 bps

UCSC Assembly: Mar. 2006
UCSC Assembly: Feb. 2006

Input sequence

```
000001 TTCCTATCTC CCATCAGCTT TATGAATAC TGAATCATTT AAAAATAAAA
000051 ACCGAAATCT TTATTTATTA TATATTTTAT TGAATCATTT TTAGCCCGCC
000101 CCCTGATCTC CCATCAGCTT CACTACCTCT TCTGCGGGCC CCGGCGCATG
000151 GCGACACCC TAGAAGAGCG GAGGCGCGCC ACC
```

This sequence is translocation.

Organism	Segment Sequence	Sequence Position	Chr Position	Chr	Cytoband
human	000001 TTCCTATCTC CCATCAGCTT TATGAATAC TGAATCATTT AAAAATAAAA	1 ~ 95	15817004 ~ 15817098	chr21	q21.1
mouse	000001 CCGGCGCGCT CATCGCATC CCATGCGCTA CTTGCTGCTC GCGGCGCGCC	96 ~ 173	24967089 ~ 24967167	chr19	qB

圖 12. human 和 mouse 跨物種轉位整合性分析結果

(2) 同樣地，我們將 human 和 rat 混合的序列和送至我們的系統分析，分析的結果如下圖 13 所示，然而這個分析結果同樣地也尚未將比對的結果做整合性的分析，因此，藉由“Analyze Cross Translocation”按鈕，我們可以得到跨物種轉位整合性分析的結果分別為圖 14 所示，找出跨物種染色體轉位分別(1) human chr19，染色體位置在 19148860 ~ 19148952，cytoband 為 p13.11 (2) rat chr12，染色體位置在 34434527 ~ 34434906，cytoband 為 q16 和 rat chr17，染色體位置在 55025572 ~ 55025649，cytoband 為 q12.1。

Cross organisms				
1. Organism:	human	UCSC Assembly:	Mar. 2006	
2. Organism:	rat	UCSC Assembly:	Nov. 2004	
Least length for traslocation: 100 bps				
Input sequence				
000001 GTCTAGGTT CCTAGTGATT GGTTCAGAT GGACAGTGA CCAATGAGC 000051 GACCTTTCTT GGGAGTTCTA CCAGGAGCAT CAGGAAGTT GAGTTGACCT 000101 CACAGACAGC TGCTGGCTTC TGTCTCCAGT GTTGTGATTA AAGGTAATTT 000151 CTACTATGGC GGACCAACCT CATGCTTC				
This segment sequence information for human.				
Segment Sequence	Sequence Position	Chr Position	Chr	Cytoband
000001 GTCTAGGTT CCTAGTGATT GGTTCAGAT GGACAGTGA CCAATGAGC	1 ~ 93	19148860 ~ 19148952	chr19	p13.11
000051 GACCTTTCTT GGGAGTTCTA CCAGGAGCAT CAGGAAGTT GAGTTGACCT	104 ~ 129	134957694 ~ 134957708	chr7	q33
This segment sequence information for rat.				
Segment Sequence	Sequence Position	Chr Position	Chr	Cytoband
000001 AGCATCAGGA AGGTTGAGTT GACCT	76 ~ 100	34434527 ~ 34434906	chr12	q16
000001 CACAGACAGC TGCTGGCTTC TGTCTCCAGT GTTGTGATTA AAGGTAATTT 000051 CTACTATGGC GGACCAACCT CATGCTTC	101 ~ 178	55025572 ~ 55025649	chr17	q12.1
Analyze Cross Translocation				

圖 13. human 和 rat 混合的序列分析結果

Cross organisms				
1. Organism:	human	UCSC Assembly:	Mar. 2006	
2. Organism:	rat	UCSC Assembly:	Nov. 2004	
Least length for traslocation: 100 bps				
Input sequence				
000001 GTCTAGGTT CCTAGTGATT GGTTCAGAT GGACAGTGA CCAATGAGC 000051 GACCTTTCTT GGGAGTTCTA CCAGGAGCAT CAGGAAGTT GAGTTGACCT 000101 CACAGACAGC TGCTGGCTTC TGTCTCCAGT GTTGTGATTA AAGGTAATTT 000151 CTACTATGGC GGACCAACCT CATGCTTC				
This sequence is translocation.				
Organism	Segment Sequence	Sequence Position	Chr Position	Chr Cytoband
human	000001 GTCTAGGTT CCTAGTGATT GGTTCAGAT GGACAGTGA CCAATGAGC 000051 GACCTTTCTT GGGAGTTCTA CCAGGAGCAT CAGGAAGTT GAGTTGACCT	1 ~ 93	19148860 ~ 19148952	chr19 p13.11
rat	000001 AGCATCAGGA AGGTTGAGTT GACCT	76 ~ 100	34434527 ~ 34434906	chr12 q16
rat	000001 CACAGACAGC TGCTGGCTTC TGTCTCCAGT GTTGTGATTA AAGGTAATTT 000051 CTACTATGGC GGACCAACCT CATGCTTC	101 ~ 178	55025572 ~ 55025649	chr17 q12.1

圖 14. human 和 rat 跨物種轉位整合性分析結果

經由本測試結果，本系統確實能找出單一物種和跨物種序列中具有染色體轉位的序列出來，並提供轉位的序列、染色體、染色體位置及 cytoband 等資訊。

4、結論

染色體轉位可能會造成疾病的發生，為了幫助生物學家分析序列中是否有染色體轉位的發生，並提供染色體轉位的資訊，本研究以 BLAT 為基礎，建構了一個 web-based 的染色體轉位分析平台，將使用者輸入的序列切成多個序列片段，然後將這些序列切片利用 UCSC BLAT 快速的比對物種序列，取得序列片段的相

關資訊，並利用重組與分析將序列中的染色體轉位資訊找出。經由不同染色體中的序列組合和跨物種 (human, mouse 和 rat) 不同染色體序列組合的測試結果，本系統確實能將序列中單一物種和跨物種之染色體轉位情況找出，並提供其相關資訊。

參考文獻

- [1] Deininger MW: Optimizing therapy of chronic myeloid leukemia. *Exp Hematol.* 2007 Apr;35(4 Suppl 1):144-54.
- [2] Buda G, Orciuolo E, Galimberti S, Benedetti E, Caracciolo F, Cervetti G, Carulli G, Papineschi F and Petrini M. “Complex translocation t(3;9;22)(q21;q34;q11) at diagnosis is a negative prognostic index in chronic myeloid leukemia”. *Leuk Res.* 2007 Mar 27.
- [3] Buda G, Orciuolo E, Cecconi N, Galimberti S, Cervetti G and Petrini M. “Poor prognosis chronic myeloid leukemia with a complex variant Philadelphia translocation, t(9;10;22)(q34;q24;q11)”. *Leuk Res.* 2007 Feb 21.
- [4] Liu D, Shimonov J, Primanneni S, Lai Y, Ahmed T and Seiter K. “t(8;14;18): a 3-way chromosome translocation in two patients with Burkitt's lymphoma /leukemia”. *Mol Cancer.* 2007 Jun 4;6(1):35.
- [5] Hecht JL, and Aster JC. *Molecular Biology of Burkitt's Lymphoma.* J. Clin. Oncol. 2000; 18:3707-3721.
- [6] Brownlee NA, Perkins LA, Stewart W, Jackle B, Pettenati MJ, Koty PP, Iskandar SS and Garvin AJ. “Recurring translocation (10;17) and deletion (14q) in clear cell sarcoma of the kidney”. *Arch Pathol Lab Med.* 2007 Mar;131(3):446-51.
- [7] National Center for Biotechnology Information BLAST, <http://www.ncbi.nlm.nih.gov/BLAST/>.