

國立臺灣大學醫學院醫學檢驗暨生物技術學所

碩士論文

Department of Clinical Laboratory Sciences and Medical Biotechnology

College of Medicine

National Taiwan University

Master Thesis

非同源性染色體轉位在人類生殖細胞發生機制之探討

Interrogate the mechanism of non-homologous  
chromosomal translocations in human germ cells



Li-Ging Hsieh

指導教授：常蘭陽 博士

Advisor: Lan-Yang Ch' ang, Ph.D.

中華民國 98 年 6 月

June, 2009

## 目錄

目錄 .....	I
緒論附圖目錄 .....	III
表目錄 .....	IV
圖目錄 .....	V
誌謝 .....	VII
中文摘要 .....	VIII
Abstract .....	X
壹、緒論 .....	1
1.1 人類基因體的結構變異 .....	1
1.2 基因體結構變異的偵測技術 .....	2
1.3 基因轉位與疾病的相關性 .....	3
1.4 <i>TCF3</i> (Transcription factor 3)簡介 .....	4
1.5 RECORD(Restriction Enhanced Capturing of Rearranged DNA)技術簡介 .....	4
1.6 研究方向 .....	5
貳、實驗材料與方法 .....	7
2.1 實驗材料 .....	7
2.2 實驗方法 .....	12
參、實驗結果 .....	18
3.1 RECORD 技術偵測敏感度的評估 .....	18
3.2 <i>TCF3</i> 之實驗設計 .....	20
3.3 由人類精蟲細胞偵測 <i>TCF3</i> 結構變異 .....	21
3.3.1 模版有無二級限制酶處理之 5' 及 3' PCR 結果比較 .....	21
3.3.2 3' PCR 結果 .....	22
3.3.3 5' PCR 結果 .....	23
3.3.4 推斷之 <i>TCF3</i> 轉位(putative translocations) .....	23
3.3.5 染色體轉位斷裂點之定位 .....	27
3.3.6 <i>TCF3</i> 染色體轉位斷裂點的分布 .....	29
3.3.7 <i>TCF3</i> 染色體轉位序列的分析 .....	30
3.4 染色體轉位確認試驗之實驗設計 .....	31
3.5 確認試驗之結果 .....	32
3.5.1 <i>MLL</i> 染色體轉位的確認 .....	32
3.5.2 評估生殖細胞 DNA 重排圖譜之可行性 .....	33
3.5.3 生殖細胞 DNA 重排圖譜之真實性 .....	33
3.5.4 確認試驗由 <i>MLL</i> -F2 偵測到的染色體轉位 .....	34

3.5.5 MLL-F2 染色體轉位斷裂點的分布 .....	38
3.5.6 精蟲細胞染色體轉位之對象在各染色體長短臂上的分布 .....	39
3.5.7 精蟲細胞染色體轉位形式為相對互調的可能性 .....	39
<b>肆、討論 .....</b>	<b>41</b>
4.1 哺乳類細胞對 DNA 雙股斷裂之修復機制 .....	41
4.2 生殖細胞染色體轉位發生機制之探討 .....	42
4.3 相對互調存在生殖細胞基因轉位的可能性 .....	43
4.4 生殖細胞的基因轉位與減數分裂 DNA 雙股斷裂修復之關係 .....	44
4.5 實驗結果整合 .....	45
<b>參考文獻 .....</b>	<b>47</b>



## 緒論附圖目錄

緒論附圖一、 體細胞基因轉位產物的形式與所導致的癌症 .....	52
緒論附圖二、 RECORD 偵測技術流程 .....	53
緒論附圖三、 偵測源自同一個體的 $t(11;22)(q23;q11)$ 轉位產物 der(11) 之電泳結果 .....	54



## 表目錄

表一、評估 RECORD 偵測敏感度之引子及 PCR 反應條件 .....	55
表二、用於偵測人類精蟲細胞 <i>TCF3</i> 結構變異之引子及 PCR 反應條件 .....	56
表三、由人類精蟲細胞基因體 DNA 偵測到的 <i>TCF3</i> 染色體轉位總表 .....	57
表四、由人類精蟲細胞基因體 DNA 在 <i>MLL-F2</i> 偵測到的染色體轉位總表 .....	58
表五、用於偵測人類精蟲細胞 <i>MLL-F2</i> 結構變異之引子及 PCR 反應條件 .....	59
表六、 <i>MLL</i> 染色體轉位特異性 PCR 反應之引子及 PCR 反應條件 .....	60
表七、確認試驗時由 <i>MLL-F2</i> 偵測到的人類精蟲細胞染色體轉位總表 .....	61
表八、以 <i>TCF3</i> 及 <i>MLL</i> 為目標基因所偵測出之人類精蟲細胞染色體轉位其轉位對象在染色體長短臂上的分布 .....	62
表九、由人類精蟲細胞偵測出之 <i>MLL-F2-F</i> 染色體轉位總表 .....	63



## 圖目錄

圖一、 K562 細胞株之染色體核型分析 .....	64
圖二、 評估 RECORD 技術偵測敏感度之實驗設計 .....	65
圖三、 評估 RECORD 技術偵測敏感度之實驗結果 .....	66
圖四、 <i>TCF3</i> 之實驗設計 .....	67
圖五、 <i>TCF3-F(3'端)</i> 及 <i>TCF3-B(5'端)</i> 模版有無二級限制酶處理之比較 .....	68
圖六、 <i>TCF3-F(3'端)</i> 之 PCR 電泳結果 .....	70
圖七、 <i>TCF3-B(5'端)</i> 之 PCR 電泳結果 .....	72
圖八之一、 <i>TCF3-F(3'端)</i> 所偵測出的染色體轉位 t(1;19)(q23.3;p13.3) .....	74
圖八之二、 <i>TCF3-F(3'端)</i> 所偵測出的染色體轉位 t(1;19)(q32.1;p13.3) .....	75
圖八之三、 <i>TCF3-F(3'端)</i> 所偵測出的染色體轉位 t(10;19)(q22.1;p13.3) .....	76
圖八之四、 <i>TCF3-F(3'端)</i> 所偵測出的染色體轉位 t(18;19)(q21.1;p13.3) .....	77
圖八之五、 <i>TCF3-F(3'端)</i> 所偵測出的染色體轉位 t(3;19)(p22.2;p13.3) .....	78
圖八之六、 <i>TCF3-F(3'端)</i> 所偵測出的染色體轉位 t(17;19)(q11.2;p13.3) .....	79
圖八之七、 <i>TCF3-F(3'端)</i> 所偵測出的染色體轉位 t(3;19)(p21.31;p13.3) .....	80
圖八之八、 <i>TCF3-F(3'端)</i> 所偵測出的染色體轉位 t(5;19)(q35.2;p13.3) .....	81
圖九之一、 <i>TCF3-B(5'端)</i> 所偵測出的染色體轉位 t(12;19)(q21.31;p13.3) .....	82
圖九之二、 <i>TCF3-B(5'端)</i> 所偵測出的染色體轉位 t(8;19)(q11.23;p13.3) .....	83
圖九之三、 <i>TCF3-B(5'端)</i> 所偵測出的染色體轉位 t(11;19)(p13;p13.3) .....	84
圖九之四、 <i>TCF3-B(5'端)</i> 所偵測出的染色體轉位 t(2;19)(p24.3;p13.3) .....	85
圖九之五、 <i>TCF3-B(5'端)</i> 所偵測出的染色體轉位 t(11;19)(q13.3;p13.3) .....	86
圖九之六、 <i>TCF3-B(5'端)</i> 所偵測出的染色體轉位 t(8;19)(q22.2;p13.3) .....	87
圖九之七、 <i>TCF3-B(5'端)</i> 所偵測出的染色體轉位 t(6;19)(q21;p13.3) .....	88
圖九之八、 <i>TCF3-B(5'端)</i> 所偵測出的染色體轉位 t(10;19)(p12.1;p13.3) .....	89
圖九之九、 <i>TCF3-B(5'端)</i> 所偵測出的染色體轉位 t(4;19)(q22.1;p13.3) .....	90
圖九之十、 <i>TCF3-B(5'端)</i> 所偵測出的染色體轉位 t(1;19)(q32.3;p13.3) .....	91
圖九之十一、 <i>TCF3-B(5'端)</i> 所偵測出的染色體轉位 t(2;19)(q33.1;p13.3) .....	92
圖十、 推測生殖細胞 DNA 雙股斷裂的可能形式及斷裂點之定位方式 .....	94
圖十一、 <i>TCF3</i> 之染色體轉位斷裂點分布 .....	95
圖十二、 <i>TCF3</i> 之染色體轉位斷裂點周邊序列 SINE 分布 .....	96
圖十三、 T-element 之分析 .....	97
圖十四、 T-element 與 MER20 在人類及黑猩猩 19p13.3 的分布圖 .....	98
圖十五、 MER20 在人類基因體中之分布圖 .....	99
圖十六、 <i>TCF3-B</i> 之染色體轉位斷裂點周圍序列 MER20 分布 .....	100
圖十七、 在人類生殖細胞偵測到的 <i>MLL</i> 染色體轉位斷裂點分布位置 .....	101
圖十八、 <i>MLL-F2</i> 確認實驗之實驗設計 .....	102

圖十九、 <i>MLL</i> 染色體轉位的確認 .....	103
圖二十、 生殖細胞 DNA 重排圖譜可信度之評估 .....	104
圖二十一、 生殖細胞 DNA 重排圖譜之真實性 .....	105
圖二十二、 <i>MLL-F2-B</i> 之 PCR 電泳結果 .....	106
圖二十三之一、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(10;11)(p12.1;q23.3)$ .....	107
圖二十三之二、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(3;11)(p22.1;q23.3)$ .....	108
圖二十三之三、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(1;11)(p32.1;q23.3)$ .....	109
圖二十三之四、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(9;11)(p24.3;q23.3)$ .....	110
圖二十三之五、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(4;11)(q35.1;q23.3)$ .....	111
圖二十三之六、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(4;11)(q25;q23.3)$ .....	112
圖二十三之七、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(3;11)(q23;q23.3)$ .....	113
圖二十三之八、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(5;11)(p12;q23.3)$ .....	114
圖二十三之九、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(11;18)(q23.3;q12.2)$ .....	115
圖二十三之十、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(11;14)(q23.3;q31.1)$ .....	116
圖二十三之十一、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(3;11)(q22.3;q23.3)$ .....	117
圖二十三之十二、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(10;11)(q26.2;q23.3)$ .....	118
圖二十三之十三、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(1;11)(q21.1;q23.3)$ .....	119
圖二十三之十四、 <i>MLL-F(5'端)</i> 所偵測出的染色體轉位 $t(11;15)(q23.3;q22.2)$ .....	120
圖二十四之一、 <i>MLL-B(3'端)</i> 所偵測出的染色體轉位 $t(11;18)(q23.3;q12.2)$ .....	121
圖二十四之二、 <i>MLL-B(3'端)</i> 所偵測出的染色體轉位 $t(4;11)(q26;q23.3)$ .....	122
圖二十四之三、 <i>MLL-B(3'端)</i> 所偵測出的染色體轉位 $t(4;11)(p12;q23.3)$ .....	123
圖二十四之四、 <i>MLL-B(3'端)</i> 所偵測出的染色體轉位 $t(2;11)(p24.3;q23.3)$ .....	124
圖二十四之五、 <i>MLL-B(3'端)</i> 所偵測出的染色體轉位 $t(11;14)(q23.3;q31.1)$ .....	125
圖二十四之六、 <i>MLL-B(3'端)</i> 所偵測出的染色體轉位 $t(11;16)(q23.3;q21)$ .....	126
圖二十五、 <i>MLL-F2</i> 之染色體轉位斷裂點分布 .....	127
圖二十三、 MMIT(microhomology mediated interchromosomal translocation).....	128
圖二十四、 生殖細胞染色體轉位產物形式之分類 .....	129

## 誌謝

猶記得剛來到實驗室的不安與忐忑，在大家的友善幫忙下，使我逐漸融入這個新環境。感謝林文昌老師及張淑媛老師對於論文的建議及指導。感謝常蘭陽老師在實驗上的指導及給予我思考上的啟發，感謝老師在生活上的關心與鼓勵，使我有勇往向前的勇氣。謝謝老師的耐心與包容，讓我擁有自由發揮的空間，並體認到對科學研究應保持大膽假設，嚴謹求證的態度。

感謝碧玉阿姨對於實驗及生活上的關心。謝謝小支學姐及吟紋學姐在實驗上的協助與鼓勵。謝謝宗呈學長及建勛學弟平時的幫忙。感謝素雯姐的幫忙與照顧。感謝身邊的朋友們及曾幫助或鼓勵過我的人，謝謝你們的聆聽與安慰，使我有面對挫折與克服困難的勇氣。

最後，感謝我的家人在我無助時給予的支持與鼓勵，也謝謝你們，使我能無後顧之憂完成此論文。



## 中文摘要

人類基因體計畫在 2003 年完成後，許多研究開始探討存在人類基因體的變異。國際人類基因體圖單型體計畫於 2002 年底至 2007 年底，已定出超過三百萬個單一核苷酸多型性(SNPs)位點。2004 年開始，研究團隊以不同的策略，在表現型正常的個體間偵測 $>1\text{kb}$  之變異(結構變異)，到目前為止，人類基因體變異數據庫(Database of Genomic Variants, DGV)中的結構變異資料，總數已達將近四萬個，顯示出結構變異在構成人類基因體變異的重要性。

對結構變異做 genome-wide 的偵測，主要以 array-based 的策略進行，受限於方法的特性與解析度，故無法對平衡結構變異做偵測或將結構變異精確定位。實驗室先前建立了 Restriction Enhanced Capturing of Rearranged DNA(RECORD)，此技術以 Inverse PCR 為基礎，用於目標基因結構變異的偵測，配合序列分析，進而可推測其可能的發生機制。

實驗室先前已利用 RECORD 技術對在新生兒及幼兒型血癌常見發生基因轉位的 *MLL*，及在幼兒型骨癌常見發生基因轉位的 *EWSR1*，在人類精蟲細胞中偵測到二種基因之轉位。本論文的目的是除選用已知基因轉位與兒童型淋巴性白血病相關的 *TCF3* 當作目標基因，探討在人類生殖細胞當中是否有類似結構變異牽涉到此基因；若有，則其可能的發生機制為何。另依據先前的實驗結果，顯示 *MLL*-F2 (*MLL*-fragement 2) 偵測到的染色體轉位斷裂點呈現明顯叢聚分布，故以此片段之環化模板利用 RECORD 技術配合基因轉位特異性引子進行確認試驗。

首先進行以 HEK293 與 K562 細胞株基因體混合 DNA 作為模版，檢測 RECORD 技術的偵測敏感度，約為  $5 \times 10^{-4}$ 。接著在 12 個不同的精蟲細胞單套基因體 DNA，共偵測到 19 個 *TCF3* 的結構變異，均為非同源性染色體轉位。經由序列對比判定 4 個轉位對象座落於基因內(geneic region)，另 15 個則位於基因間(intergenic region)。而就染色體轉位之機制而言，斷裂點周邊重複序列並未參與，我們可排除

其為同源性重組(homologous recombination, HR)造成之可能。而在染色體轉位之斷裂點交界處均有 1-13 個核苷酸的微同源性序列(microhomology sequence)，與在體細胞由 NHEJ 及 MMEJ(microhomology mediated end-joining)所造成的基因轉位類似。由於 NHEJ 核心因子 Ku70/Ku80 在減數分裂時並未表現，故並非 NHEJ 所造成。推測這些染色體轉位可能由類似 MMEJ 的機制所產生。

針對 3 個先前由同一捐贈者之精蟲細胞基因體 DNA 在 *MLL-F2* 偵測到的染色體轉位進行確認，但均未能被確認。故知其發生頻率可能低於或恰位於 RECORD 技術偵測之極限，應以敏感度更高的取代方法來驗證。此外，由十二個不同捐贈者的精蟲細胞基因體 DNA，在 *MLL-F2* 偵測到 20 個染色體轉位。轉位對象 8 個位於基因內(geneic region)，12 個位於基因間(intergenic region)。在轉位斷裂點交界處均具有 2-10 個核苷酸的微同源性序列(microhomology sequence)。

若將目標基因序列調整為正股時，依轉位對象位於染色體長臂或短臂、microhomology 位於正股或負股及目標基因以 5' 端或 3' 端序列發生轉位，產物形式可分為八種，以 5' 或 3' PCR 分別可偵測到四種。我們發現偵測到的染色體轉位均可依上述規則分類，由 5' 及 3' PCR 均偵測到預期形式的產物，且屬於平衡互換(balanced exchange)；此外，亦顯示出這些轉位的發生需遵循微同源性序列(microhomology sequence)間的 base-pairing rule，應為微同源性序列媒介的機制。為了與在體細胞發生的 MMEJ 有所區隔，我們將之命名為 MMIT(microhomology mediated inter-chromosomal translocation)。

生殖細胞進行減數分裂重組的重要目的即是在修復減數分裂時產生的DNA雙股斷裂(double strand break)。由於我們在不同來源的精細胞基因體 DNA 中，普遍均能偵測到基因轉位的存在，且這些序列之特性，明顯異於由同源性重組(HR)修復者，故我們推測 MMIT 應屬於人類生殖細胞中修復 DNA 雙股斷裂的機制之一。

## Abstract

Since the human genome project finished in 2003, many studies focused on the variations in the human genome. From 2002 to 2007, the international HapMap project mapped over 3.1 million single nucleotide polymorphisms (SNPs). Using different approaches, until recently, about 40000 >1kb structural variations (SVs) had been detected between non-disease phenotype individuals. These observations reflected the important contribution of SVs to human genome variation.

To interrogate which mechanisms involve in SVs formation in human genome, PCR-based approach is the most sensitive one. Previously, an inverse PCR-based approach, RECORD (Restriction Enhanced Capturing of Rearrange DNA), was used to identify *MLL* and *EWSR1* translocations in human germline. *MLL* and *EWSR1* are frequently involved in recurrent translocation associated with infant and childhood leukemia and childhood sarcoma, respectively.

The aim of my project is to use human *TCF3*, which is a known target of recurrent translocation associated with childhood acute lymphoblastic leukemia (ALL), as an anchor gene to examine whether similar event occurs in human sperm and to explore the underlying mechanisms involved. Moreover, we will use RECORD and translocation specific PCR to validate germline translocations which identified previously.

To assess the detection limit of RECORD, we used HEK293 spiked with K562 genomic DNA as template, and examined the recovery of *BCR-ABL* fusion from K562 cells. We found that the detection limit by RECORD is about  $5 \times 10^{-4}$ . By using *TCF3* as an anchor gene, we identified 19 non-homologous interchromosomal translocation events from 12 donors. These 19 translocation events include four translocation partners in geneic region and 15 in intergenic region. Sequence analysis of the 2 kb regions flanking both sides of breakpoints excluded the involvement of non-allelic homologous

recombination (NAHR). A microhomology of 1-13 base pairs at the breakpoint junction indicates that either non-homologous end joining (NHEJ) or microhomology mediated endjoining (MMEJ) may be responsible for germline translocation. though the core factors, Ku70/Ku80 of NHEJ are not expressed during meiotic recombination.

Using *MLL* as an anchor gene, we identified 20 additional *MLL*-F2 translocations from the haploid genomes of 12 sperm donors. Eight of which were mapped in the genic region and 12 in intergenic region of partners. At the breakpoint junctions, there are microhomologies of 2-10 nucleotides.

When microhomology of the anchor gene is considered always on the top strand, the translocation partners can be classified into four groups, based on its location on the p or q arm, and its sequence from top or bottom strand. Thus 8 derivatives of translocation products can be defined by using 5' and 3' iPCR in our studies. This germline balanced exchange between non-homologous chromosomes is mediated by microhomology. To distinguish from MMEJ reported only in somatic cells, we proposed this putative mechanism as microhomology mediated interchromosomal translocation (MMIT). As the important purpose of meiotic recombination is to repair DNA double-strand breaks (DSBs), we suggest that MMIT may be another meiotic DSBs repair pathway in male germ cells.

# 壹、緒論

## 1.1 人類基因體的結構變異

人類基因體計畫在 2003 年完成人類基因體定序後，許多研究開始探討存在人類基因體的變異。人類基因體的變異可分為幾個層次；以單一核苷酸的層次而言，稱為單核苷酸多型性(single nucleotide polymorphisms, SNPs)；數十個至數百個核苷酸序列的變異，包含小片段的插入(insertion)及刪除(deletion)、短片段串連重複序列(short tandem repeats, STRs)重複次數的差異等；牽涉到大片段(kb~Mb)序列的變異，則稱為結構變異(structural variations, SVs)；而結構變異可再依據是否能以 cytogenetic analysis 以顯微鏡被觀察到，區分為 microscopic SVs(>3 Mb)(如：染色體數目異常, aneuploidy；大片段重排, gross rearrangement 等)及 submicroscopic SVs (1 kb~3 Mb)(如：轉位、倒置、重複、刪除等)。此外，又依據是否會影響 DNA 拷貝數的改變，區分為平衡的變異(如：倒置及染色體轉位)與會改變 DNA 拷貝數的變異(copy number variations, CNVs)。

於 2002 年底開始的另一項跨國型計畫—國際人類基因體圖單型體計劃 (International HapMap Project)，藉由不同個體間基因體序列的比對，定出人類染色體上單核苷酸多型性(SNPs)的位點。國際人類基因體圖單型體計劃於 2002 年 10 月開始，至 2007 年底 Phase II HapMap 所發表的資料，已定出超過三百萬個 SNPs 位點<sup>(1)</sup>。此結果顯示出 SNPs 在構成人類基因體變異中的重要性，且在後續的研究當中，也證實部分 SNPs 會影響表現型(phenotype)，而與疾病的發生有關。

在 2004~2005 年的幾篇突破性研究中，分別針對表現型正常的不同個體(non-disease phenotype individuals)，對其基因體內的 SVs 做 large-scale(>100 kb)<sup>(2, 3)</sup> 及 fine-scale(>8 kb)<sup>(4)</sup> 兩種不同精細度的偵測，共偵測到數百個 submicroscopic SVs。此結果似乎意味著 SVs 在構成人類基因體變異中所扮演的角色，可能與 SNPs、短

片段串連重複序列(STRs)及其他小片段變異同樣重要，不容小覷<sup>(5)</sup>。此後，不同的研究團隊，陸續進行人類基因體結構變異的偵測<sup>(6, 7)</sup>，實驗結果呼應了先前的推測，證實 SVs 在構成個體間基因體之差異相當重要。根據人類基因體變異數據庫(Database of Genomic Variants, DGV)的統計資料，從 2004 年到目前為止，由表現型正常的不同個體所偵測出大於 1 kb SVs，總數已達將近四萬個，涵蓋範圍約為基因體的 30%，其中又以 CNVs 的數量最多。

## 1.2 基因體結構變異的偵測技術

早期對基因體 SVs 的偵測方式，主要以 cytogenetic analysis 為主，利用染色體帶型(chromosome banding)技術配合光學顯微鏡的觀察，解析度約為百萬核苷酸(Mb)的程度；此方式除了無法得知涉及的基因为何之外，也無法偵測到較小片段的結構變異。

目前所用的偵測策略可分為三種：(1)Array-based approach (2)Sequence-based approach (3)Targeted scans。Array-based approach(如：CGH)可進行高通量的 genome-wide 偵測，其結果可反應出 CNVs，但無法反應出平衡的染色體結構變異，例如：轉位(translocation)及倒置(inversion)。Sequence-based approach(如：paired-end sequencing, *in silico* sequence comparisons)可進行 genome-wide 偵測，解析度可達單一核苷酸(bp)的程度，且可偵測出平衡的結構變異，但無法做高通量分析；由於解析度達 bp 的程度，故可藉由序列分析，推測其可能的發生機制。Targeted scans 主要有兩種方式；針對目標基因設計探針(probe)，進行螢光原位雜交(fluorescence *in situ* hybridization, FISH)或南方墨點法(Southern blotting)，解析度為 Mb 或 kb 的程度；另一種方式是針對目標基因(target gene)設計引子(primer)，進行聚合酶連鎖反應(polymerase chain reaction, PCR)，再針對產物做定序分析，即可得知發生結構變異的序列為何，解析度可達 bp 的程度，且對於平衡的結構變異也可偵測出來。

從 2004 年到目前為止，人類基因體變異數據庫(DGV)的資料來源，主要由 Array-based approach 的方式獲得，但此方法的解析度受限於探針分布影響，無法提供 bp 解析度的資訊，故不能精確定出 SVs 的分布位置及其大小，只能做約略估計；此外，也無法提供平衡結構變異的資訊，而此限制可能是造成目前 DGV 的統計資料，以 CNVs 為最多的原因，這也暗示著到目前為止，人類基因體中仍有許多 SVs(如：平衡的結構變異，小片段的插入或刪除)尚未被偵測出來<sup>(8)</sup>。

先前本實驗室建立了一個以 targeted scans 為策略的技術，命名為 RECORD (Restriction Enhanced Capturing of Rearranged DNA)，能針對目標基因做其 SVs 的偵測。此技術是以 inverse PCR-based 的方式進行，配合產物的定序分析，除了可精確定出該 SVs 斷裂點在染色體上的分布位置之外，亦可推測其可能的發生機制。

### 1.3 基因轉位與疾病的相關性



人類基因體的 SVs 可藉由多種方式造成表現型改變而導致疾病發生，例如：劑量敏感性基因(dosage-sensitive genes)發生 CNVs；因 SVs 導致劑量敏感性基因在基因體中的位置效應(positional effect)而影響其表現量，及經由因基因轉位產生功能不同於原本蛋白質產物的融合蛋白(fusion protein)等。根據先前的研究，偶發性疾病(sporadic diseases)、常見疾病(common diseases)、罕見的遺傳性疾病及與臨床用藥代謝相關的表現(pharmacogenetics)，均與結構變異有關<sup>(8)</sup>。

根據過去的研究，體細胞內之基因轉位主要藉由兩種方式引發癌症；基因轉位產生致癌性的融合蛋白質產物(序論附圖一之 A)，或是某些基因(如：oncogenes)經轉位而被帶至鄰近另一個基因的啟動子(promoter)或增強子(enhancer)，進而影響該基因的表現量(序論附圖一之 B)<sup>(9)</sup>。過去已報導許多因體細胞基因轉位而導致癌症發生的例子<sup>(10)</sup>，根據美國國立癌症研究所 Mitelman Database of Chromosome Aberrations in Cancer 資料庫所整合的資料，已知與人類癌症相關的基因轉位，以

出現於血液細胞的惡性疾病(hematopoietic malignancies)的比例最高(約 75%)，而在骨肉瘤(sarcoma)及上皮源惡性疾病(carcinomas)比例較少(約 25%)<sup>(10)</sup>。此外，癌細胞的基因轉位常具有以下特性：(1)復發性(recurrent)轉位。有相同癌症的不同病患，發生轉位的基因具有特定的轉位熱處(hotspots)，稱為斷裂點叢聚區域(breakpoint cluster region, BCR)。(2)具有共同的轉位對象基因(common translocation partner genes, common TPGs)。

#### 1.4 TCF3 (Transcription factor 3) 簡介

TCF3 位於第十九對染色體短臂 13.3(19p13.3)的位置，基因大小約為 41.3 kb，具 19 個外顯子(exon)。其蛋白質產物屬於 E-protein family 的一員，在結構上具有 basic helix-loop-helix motif(bHLH)，可與 E-box(5'-CANNTG-3')結合，再藉由 trans-activation domain(TAD)以調控基因表現，在功能上屬於轉錄因子(transcription factor)。TCF3 可與其他具有 bHLH 的蛋白質(如：HEB)組成二聚體(heterodimer)，調控組織特異性基因的表現。而 TCF3 本身組成的二聚體(homodimer)，對於調控 B 細胞發育過程相關的基因表現扮演重要的角色<sup>(11)</sup>。

根據先前的報導，在兒童型 B 細胞譜系的急性淋巴球性白血病(B-lineage ALL)患者中，TCF3 常與其他轉錄因子，如：*PBX*<sup>(12)</sup>，*HLF*<sup>(13)</sup>及 *TFPT*<sup>(14)</sup>發生基因轉位，產生致病性的融合蛋白產物。

#### 1.5 RECORD (Restriction Enhanced Capturing of Rearranged DNA) 技術簡介

RECORD 偵測技術是以反向聚合酶連鎖反應(Inverse PCR)為基礎，再經過小部份修改，進而達到提升偵測敏感度的效果。

RECORD 偵測技術的流程(序論附圖二)如下：定出目標基因的結構好變熱處

(hotspots)後，隨之分析該基因的限制酶之切位點，找尋適當的初級限制酶(primary restriction enzyme)，以擷取可涵蓋該基因之結構變異好變處。利用所選用之初級限制酶，將基因體 DNA 片段化(fragmentation)後，再進行片段環化的步驟。為了提升目標基因發生結構變異的 DNA 分子(rearranged form)被偵測到之機率，故再依據欲分析之區域位於誘餌基因的 3' 端或 5' 端，選擇另一個二級限制酶(secondary restriction enzyme)切割未發生結構變異(unrearranged form)的環化目標基因片段，以降低其被擴製的機率。經上述方式處理後的模板，以設計於目標基因 3' 端或是 5' 端的引子進行反向聚合酶連鎖反應後，將產物選殖並定序，即可得知與目標基因發生結構變異的對象為何。

由於此技術使用二級限制酶切割 unrearranged form 模板，以提升 rearranged form 模板的檢出率，導致此技術產生兩個限制：第一，與目標基因發生結構變異的 DNA 序列當中不能存在二級限制酶的切位，若有，則無法被 PCR 反應放大出來，因此無法被偵測到。第二，目標基因發生結構變異的斷裂點需位於二級限制酶切位內側，若位於外側，則二級限制酶可將此模版切割，亦無法被 PCR 反應放大出來而不能被偵測到。

實驗室先前已利用 RECORD 技術，分別針對在新生兒及幼兒型血癌常見發生轉位的 *MLL*，及在幼兒型骨癌常見發生轉位的 *EWSRI*，在人類精蟲細胞中均偵測到二種基因的轉位。

## 1.6 研究方向

當構成基因體的 DNA 在組成上有異常時，即可能會導致遺傳性疾病(genetic diseases)。根據過去的研究，大於 97% 遺傳性疾病並非由親代遺傳而來，而是偶發的(sporadic)，且大部分與 CNVs 有關<sup>(15)</sup>。

到目前為止，因體細胞的復發性基因轉位(recurrent translocation)而導致癌症發

生的例子已有許多<sup>(10)</sup>，但對於生殖細胞基因轉位的相關研究仍然相當缺乏。目前已知發生在生殖細胞的復發性基因轉位，t(11;22)(q23;q11)是唯一的例子<sup>(16)</sup>。t(11;22)(q23;q11)平衡轉位的帶因者(carriers)，表現型正常，但若此帶因者的生殖細胞在減數分裂 I 期(meiosis I)同源染色體未發生正常分離，導致配子細胞的染色體數目分布不均，受精後，子代染色體核型異常，呈現 47,XX,t(11;22)(q23;q11),+der(22) 或 47,XY,t(11;22)(q23;q11),+der(22)，導致 supernumerary-der(22)t(11;22) syndrome。此篇研究針對四名健康捐贈者的精細胞基因體 DNA，以 PCR 反應分別對 t(11;22)(q23;q11)產生的兩種衍生染色體(derivative chromosomes)做偵測，每個檢體均進行多次重複實驗。實驗結果顯示，4 名捐贈者檢體的不同 aliquots 中，均有部分呈現 t(11;22)(q23;q11)陽性反應，預估發生頻率為  $1.24\text{-}9.46 \times 10^{-5}$ ，且由相同檢體偵測出的兩種衍生染色體發生頻率近乎相等。此外，四名捐贈者咽喉拭子(buccal swab)檢體的 DNA，均呈現 t(11;22)(q23;q11)陰性反應，顯示此基因轉位為在精細胞 *de novo* 發生的。源自相同捐贈者的 PCR 產物經電泳及定序分析，顯示在相同捐贈者的精細胞中，t(11;22)(q23;q11)的 PCR 產物片段大小仍有差異(序論附圖三)，且基因轉位的斷裂點也有所不同。

上述結果似乎暗示著人類生殖細胞具有 *de novo* 產生相對互調形式的基因轉位(reciprocal translocation)之潛力，且在同一個體中，即使是相同形式的基因轉位，也可能有差異性存在。因此，本論文先對 RECORD 技術用於偵測精蟲細胞染色體轉位的偵測敏感度進行評估，接著，再以 TCF3 當作目標基因，探討在人類精蟲細胞中是否有結構變異涉及此基因；若有，則其可能的發生機制為何。另外，依據先前的實驗結果，顯示 MLL-F2(MLL-fragement 2)偵測到的基因轉位斷裂點呈現明顯叢聚分布，故以此片段之環化模板利用 RECORD 技術配合轉位特異性引子進行確認試驗。

## 貳、實驗材料與方法

### 2.1 實驗材料

#### 2.1.1 試藥

限制酶(restriction enzyme)(New England BioLabs Inc.)

BamHI, BglII, KpnI, XbaI, XhoI, NsiI, EcoRI

限制酶(restriction enzyme)(Fermentas)

BglII, AfeI, Mva1269I(Bsm I)

T4 DNA ligase(3U/ $\mu$ l)(Promega)

ExPrime Taq<sup>TM</sup> DNA polymerase(2/ $\mu$ l)(GENET BIO)

dNTP(Invitrogen)

primers(PURIGO BIOTECH, Inc.)

NaCl(MERCK)

MgCl<sub>2</sub>(MERCK)

Tris base(J.T. Baker)

Boric acid(MERCK)

EDTA(MERCK)

100% Ethanol(MERCK)

Yeast tRNA(實驗室提供)

DTT(Sigma)

Ultra Pure Agarose(Invitrogen)

DMSO(Sigma)

Ethidium bromide(Sigma)

Ampicillin(Sigma)

Equilibrated phenol(pH7.8)(Invitrogen)

Chloroform(MERCK)

Isoamylalcohol(MERCK)

Isopropanol(MERCK)

### 2.1.2 Kit

DNA Purification Kit(PUREGENE)

PCR-M<sup>TM</sup> system(Viogene)

Plasmid Miniprep Purification Kit(GeneMark)

pGEM-T Easy Vector System(Promega)

CloneJET<sup>TM</sup> PCR Cloning Kit(Fermentas)

ECOS<sup>TM</sup> 101 competent cell(Sigma)

### 2.1.3 模板(Templates)

HEK293 基因體去氧核醣核酸(genomic DNA)

K562 基因體去氧核醣核酸(genomic DNA)

S3 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Donor 1)

S4 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Donor 2)

S5 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Donor 3)

S6 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Donor 4)

S7 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Sample A)

S8 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Donor 5)

S13 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Sample B)

S14 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Sample C)

S15 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Sample D)

S18 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Donor 6)

S19 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Donor 7)

S23 精蟲細胞基因體去氧核醣核酸(sperm genomic DNA)(Sample E)

#### 2.1.4 引子(Primers)

用於評估 RECORD 技術偵測敏感度之引子

<i>BCR-i-1F</i>	TCCAGCCACTCATTCTCAGACC
<i>BCR-i-1R</i>	GGAAGTGCCCAGAGAAGGAAGG
<i>BCR-i-2F</i>	ATGCTCTGTGCCTTGGATCTGG
<i>BCR-i-2R</i>	TCAGGACAGAGGATACTCACACTTC
<i>BCR-i-C</i>	TGTACACGCCATTGTTACACAGTG
<i>BCR/ABL-R</i>	ACTATCAAGGGACAAGTACTGAGTG

用於偵測精蟲細胞 *TCF3* 結構變異之引子

<i>TCF3-F-1F</i>	CGTCCAGGTTAGGTACAGCACAGG
<i>TCF3-F-2R</i>	GCCCTCCCTCACTGCCTGGTTATG
<i>TCF3-F-3F</i>	ATGTGTGCGTGTGCCATTGCTC
<i>TCF3-F-4R</i>	CGGTTCTGACCAGCGATGCC
<i>TCF3-F-Lig.</i>	TGAGTGATCCGGGGAGTAGATCC
<i>TCF3-B-1F</i>	ACTGTGCCCATCCGAAGTGTTC
<i>TCF3-B-2R</i>	GCACTCTGGTGTGACAGGGTGAGAC
<i>TCF3-B-3F</i>	TGTGGTCCTTAATCTGTCTGGCATTGG
<i>TCF3-B-4R</i>	ACACCCACCAACAGATGTAGCCAACC
<i>TCF3-B-Lig.</i>	ACACATGGCGTGAGGTAGGGATCT

用於由 *MLL-F2* 偵測到之染色體轉位確認實驗之引子

<i>MLL-F2-1F</i>	AACATTCCATCCATCCTGAGC
------------------	-----------------------

<i>MLL</i> -F2-2R	ATTCCTCTTCAACACCTTG
<i>MLL</i> -F2-2F	CCATGATGATTCCCTGAGTCAG
<i>MLL</i> -F2-3F	TATGAATTGAACAACTAGGTGAGC
<i>MLL</i> -F2-4R	TTGCTTAGAACTATTGCCATTGG
<i>MLL</i> -F2-Lig.	TGCAAATGACAAATTTATTAATAGA
<i>MLL</i> -F2-B-1F	TGGAGTGTAAATAAGTGCCAACAG
<i>MLL</i> -F2-B-6R	ACCCGTGAGATTGCTATGTACAAC
<i>MLL</i> -F2-B-3F	TGCCTCATTACTAGGAAATCATCTCAG
<i>MLL</i> -F2-B-8R	TGTGGTAAGTAATAAGCCAAAGAGG
<i>MLL</i> -F2-B-Lig.	CTATAAATGGATGCATTTAAGA
<i>MLL</i> -F2-B-10R	AACTGTAAAGGATGTAGGATATG
<i>NOTCH2NL</i> -V	ATCACCTACTGTTAATGTC
8q11.21-V	AGATGTTGAACCAGGGAATG
<i>PIK3R4</i> -V	ACCTTCACAAACGACCAG



### 2.1.5 儀器

核酸暨蛋白質分光測定儀	Eppendorf Biophotometer 6131
水浴槽	Hotech Waterbath 830
乾浴器	Fisher Scientific
離心機	Eppendorf Centrifuge 5424
電泳槽	Pharmacia Blotech SE250-10A-75
PCR 機器	ABI GeneAmp PCR System 9700
照膠系統	UVP GDS-7900 system

### 2.1.6 溶液之配製

### TE buffer

1M Tris-HCl(pH8.0)10ml 加入 0.5M EDTA 2ml，補二次水至 1 公升，混合均勻後滅菌，室溫保存。

### TEN buffer

1M Tris-HCl(pH8.0)10ml 加入 0.5M EDTA 10ml 及 5M NaCl 20ml，補二次水至 1 公升，混合均勻後滅菌，室溫保存。

### 10X ligation buffer

1M MgCl<sub>2</sub> 500μl，0.5M DTT 250μl，100mM ATP 500μl，1M Tris-HCl(pH7.8) 2.5ml，補滅菌水至最終體積為 5ml，混合均勻。適量分裝保存於-20°C 備用，避免重複解凍太多次。

### 5X TBE

取 54g Tris Base 及 27.5g Boric acid，加入 0.5M EDTA(pH8.0)0.2ml，補二次水至最終體積為 1 公升，儲存備用。使用前，以二次水稀釋十倍(0.5X TBE)使用。

### 0.8% agarose gel

0.8g 洋菜膠粉加入 100ml 0.5X TBE，於微波爐加熱至透明完全溶解，稍微冷卻之後加入 Ethidium bromide，混合均勻後倒入製膠盤中，膠體凝固後即可用於分離 DNA 片段。

### LB broth

取 Bacto tryptone 10g，Bacto yeast extract 5g 及 NaCl 10g，加入 5N NaOH 0.2ml 後，補二次水至最終體積為 1000ml，滅菌後保存於 4°C 備用。

### LB plate

取 Bacto tryptone 4g，Bacto yeast extract 2g 及 NaCl 4g，加入 5N NaOH 0.08ml 後，再加入 Agar 3g，補二次水至最終體積為 400ml，經高溫高壓滅菌後，待其冷卻至 55°C，加入 Ampicillin，使最終濃度為 100μg/ml，混合均勻後倒入養菌之培養盤(Petri-dish)，待其凝固風乾後，保存於 4°C 備用。

## Chloroform(CIA)

Chloroform 與 Isoamylalcohol 以體積比為 23:1 的比例混合，避光保存於室溫。

## Phenol: Chloroform(PCIA)(pH8.0)

依體積比 Equilibrated phenol: Chloroform: Isoamylalcohol = 25: 24: 1 的比例混合，避光保存於 4°C。使用時取下層。

## 2.2 實驗方法

### 2.2.1 細胞株基因體 DNA(genomic DNA)萃取

- A. 貼附型細胞株(如：HEK293)先以 1X PBS 洗 2 次後，加入 0.5% trypsin 將細胞打下，再用含 10% FBS 的 DMEM 中止 trypsin 作用。計數細胞。  
懸浮型細胞株(如：K562)接取細胞懸浮液計數細胞。
- B. 將細胞懸浮液移至 15ml 離心管，離心 1000rpm，5 分鐘，去除上清液。
- C. 以 1X PBS 將細胞濃度調整至  $1-2 \times 10^6/\text{ml}$ ，分裝至 1.5ml 離心管中。離心 14000rpm，5 分鐘，去除上清液。
- D. 加入 300 $\mu\text{l}$  Cell Lysis Solution，混合均勻。置於 37°C 作用 1 小時。
- E. 加入 4mg/ml RNase A Solution 1.5 $\mu\text{l}$ ，混合均勻。置於 37°C 作用 1 小時。
- F. 冰上靜置 1 分鐘，100 $\mu\text{l}$  Protein Precipitation Solution，vortex 20 秒後，離心 14000rpm，1 分鐘。取上清液至新的 1.5ml 離心管，加入等體積 phenol: chloroform(PCIA)(pH8.0)，混合均勻後，離心 13000rpm，5 分鐘。
- G. 取上層至新的 1.5ml 離心管，加入等體積 chloroform(CIA)混合均勻後，離心 13000rpm，2 分鐘。
- H. 取上層至新的 1.5ml 離心管，加入 300 $\mu\text{l}$  100% Isopropanol 混合均勻，離心 14000rpm，5 分鐘。去除上清液。
- I. 加入 300 $\mu\text{l}$  70% Ethanol 清洗 DNA pellet，離心 14000rpm，1 分鐘，去上

清液後，開蓋風乾，加入適量 TE buffer，置於 65°C 1 小時或室溫 16 小時，使 DNA 均勻回溶。

J. 以測光儀測定濃度，保存於-20°C 備用。

### 2.2.2 模板之前處理

#### 1. 基因體 DNA 初級限制酶水解(primary restriction enzyme digestion)

A. 5μg 基因體 DNA 加入 20μl 10X buffer，8μl BamHI 及 8μl BglII，補水至最終體積為 200μl，置於 37°C 作用 16 小時。

B. 取 5μl 跑 0.8% agarose gel，確認基因體 DNA 已被片段化。

#### 2. 酒精沉澱

A. 在已片段化之基因體 DNA 反應物(200μl)中加入  
5M NaCl                  20μl(1/10 體積)  
100% Ethanol            440μl(2 倍體積)  
翻轉數次後，-20°C 靜置 30 分鐘。

B. 離心 14000rpm，20 分鐘。去除上清液。

C. 加入 70% Ethanol 200μl(1 倍體積)，離心 14000rpm，5 分鐘。去除上清液。

D. 開蓋置於 65°C 乾浴槽 10 分鐘，使剩餘酒精揮發。

E. 加入 100μl TE buffer 回溶 DNA。

F. 以測光儀測定 DNA 濃度。

#### 3. 片段化之基因體 DNA 的環化(circularization)

需先分別以 DNA 濃度為 5ng/μl, 1ng/μl 及 0.2ng/μl 測試最佳環化條件。配製方式如下：

A. 先配製環化條件為 DNA 濃度 5ng/μl 之 mastermix。取 3μg 片段化之基因體 DNA，加入 60μl 10X ligation buffer 及 3μl T4 DNA ligase(3U/μl)，補水至最終體積為 600μl。

- B. 將配製完成的 mastermix 分裝至 3 管 1.5ml tube 中(每管 200 $\mu$ l mastermix)。
- C. Tube 1: 200 $\mu$ l mastermix，總體積為 200 $\mu$ l(環化條件為 DNA 濃度 5ng/ $\mu$ l)。
- Tube 2: 200 $\mu$ l mastermix，再加入 80 $\mu$ l 10X ligation buffer 及 720 $\mu$ l H<sub>2</sub>O，總體積為 1000 $\mu$ l(環化條件為 DNA 濃度 1ng/ $\mu$ l)。
- Tube 3: 將 200 $\mu$ l mastermix 再分成 5 管(40 $\mu$ l/tube)，每管再加入 96 $\mu$ l 10X ligation buffer 及 864 $\mu$ l H<sub>2</sub>O，每管總體積為 1000 $\mu$ l(環化條件 DNA 濃度 0.2ng/ $\mu$ l)。

D. 置於 16°C 作用 16 小時。

#### 4. 酒精沉澱

- A. 分裝環化產物，每管 200 $\mu$ l。
- B. 每管加入 1 $\mu$ l 10mg/ $\mu$ l yeast tRNA(co-carrier), 20 $\mu$ l 5M NaCl(1/10 體積), 440 $\mu$ l 100% Ethanol(2 倍體積)。翻轉數次後，-20°C 靜置 30 分鐘。
- C. 離心 14000rpm，20 分鐘。去除上清液。
- D. 開蓋置於 65°C 乾浴槽 10 分鐘，使剩餘酒精揮發。
- E. 加入 TE buffer 回溶 DNA(將最終濃度調整成 20 ng/ $\mu$ l)。

#### 5. 環化之基因體 DNA 二級限制酶水解(secondary restriction enzyme digestion)

- A. 取 25 $\mu$ l 環化之基因體 DNA(500ng)，加入 10 $\mu$ l 10X buffer(依限制酶不同選用不同 buffer)及 5U 限制酶，補水至最終體積為 100 $\mu$ l。
- B. 置於所選用之限制酶的反應溫度，作用 16 小時。

#### 6. 酒精沉澱

- A. 在環化之基因體 DNA 限制酶水解產物中加入 1 $\mu$ l 10mg/ $\mu$ l yeast tRNA，5M NaCl 10 $\mu$ l(1/10 體積)，100% Ethanol 220 $\mu$ l(2 倍體積)，翻轉數次後，-20°C 靜置 30 分鐘。
- B. 離心 14000rpm，20 分鐘。去除上清液。

- C. 加入 70% Ethanol 100 $\mu$ l(1 倍體積)，離心 14000rpm，5 分鐘。去除上清液。
- D. 開蓋置於 65°C 乾浴槽 10 分鐘，使剩餘酒精揮發。
- E. 加入 20 $\mu$ l TE buffer 回溶 DNA(最終濃度為 25 ng/ $\mu$ l)。

### 2.2.3 長片段反向聚合酶連鎖反應

(Long-distance inverse polymerase chain reaction, LDI-PCR)

#### 1. 配製長片段聚合酶連鎖反應之反應物

10X ExPrime buffer 5 $\mu$ l，2.5mM dNTP mixture 7 $\mu$ l，ExPrime Taq<sup>TM</sup> DNA polymerase 1 $\mu$ l(2U/ $\mu$ l)，10 $\mu$ M forward primer 1 $\mu$ l，10 $\mu$ M reverse primer 1 $\mu$ l，模板(template)100ng，視狀況添加 DMSO(去除 primer dimer)，補水至最終體積為 50 $\mu$ l。

#### 2. 長片段反向聚合酶連鎖反應

PCR 反應條件設定如下：

$(94^{\circ}\text{C}, 5\text{min}) \rightarrow [(94^{\circ}\text{C}, 10\text{sec}) \rightarrow (Y^{\circ}\text{C}, 30\text{sec}) \rightarrow (68^{\circ}\text{C}, Z \text{ min})]_{20}$   
 $\rightarrow [(94^{\circ}\text{C}, 10\text{sec}) \rightarrow (X^{\circ}\text{C}, 30\text{sec}) \rightarrow (68^{\circ}\text{C}, Z \text{ min}^*)]_{15}$   
 $\rightarrow (68^{\circ}\text{C}, 7\text{min}) \rightarrow (4^{\circ}\text{C}, \infty)$

以 touch-down PCR 的方式增加產物的特異性。不同片段的 PCR 反應條件詳列於表一、二、五。

\* 每增加一個 cycle，延長 5 秒鐘

#### 3. 產物的純化

取 30 $\mu$ l PCR 產物，以 PCR-M<sup>TM</sup> system(Viogene)做純化，再以 35 $\mu$ l TE buffer 將 DNA 沖提出來。

#### 4. PCR 產物二級限制酶水解

為避免模板前處理時二級限制酶未作用完全的疑慮，故再將產物做一次二級限制酶水解。

- A. 取純化後的 PCR 產物 1 $\mu$ l，加入 1 $\mu$ l 10X buffer(依限制酶選用不同 buffer)及

5U 限制酶，補水至最終體積為 10 $\mu$ l。(相當於 inverse PCR 產物稀釋 10 倍)。

B. 置於所選用之限制酶的反應溫度，作用 16 小時。

### 5. 巢式聚合酶連鎖反應(Nested PCR)

在 Inverse PCR 夾出的誘餌基因序列的內側，再設計一對引子，以提高產物的特異性並增加產率。

A. 將上述經二級限制酶水解的產物以 TE buffer 做適當比例的稀釋( $10^{-2} \sim 10^{-4}$  倍稀釋)，取 1 $\mu$ l 進行 Nested PCR 反應。

B. Nested PCR 反應物之配製：

10X ExPrime buffer 5 $\mu$ l，2.5mM dNTP mixture 7 $\mu$ l，ExPrime Taq<sup>TM</sup> DNA polymerase 0.5 $\mu$ l(2U/ $\mu$ l)，10 $\mu$ M forward primer 1 $\mu$ l，10 $\mu$ M reverse primer 1 $\mu$ l，模板(template) 1 $\mu$ l，視狀況添加 DMSO(避免 primer dimer 形成)，補水至最終體積為 50 $\mu$ l。

C. PCR 反應條件設定如下：

(94°C, 5min) → [(94°C, 10sec) → (Y°C, 30sec) → (68°C, Z min)]<sub>20</sub>  
→ [(94°C, 10sec) → (X°C, 30sec) → (68°C, Z min<sup>\*</sup>)]<sub>15</sub>  
→ (68°C, 7min) → (4°C,  $\infty$ )

以 touch-down PCR 的方式增加產物的特異性。不同片段的 PCR 反應條件詳列於表一、二、五。

\* 每增加一個 cycle，延長 5 秒鐘

#### 2.2.4 選殖 (cloning)

使用 pGEM-T Easy Vector System(Promega)或 CloneJET<sup>TM</sup> PCR Cloning Kit (Fermentas)，將 PCR 產物接入 vector 中。經轉形作用(transformation)將 vector 送入勝任細胞中，均勻塗在含有 100 $\mu$ g/ml Ampicillin 的 LB 培養基上，於 37°C 培養 12 小時。挑取單一菌落以含有 100 $\mu$ g/ml Ampicillin 的 LB broth 進行培養，於 37°C 震

溫培養 12 小時。

### 2.2.5 抽取質體(plasmid extraction)

使用 Plasmid Miniprep Purification Kit(GeneMark)抽取質體，以 50 $\mu$ l elution buffer 沖提後，保存於-20°C 或 4°C。

### 2.2.6 質體 insert DNA 的確認

使用 EcoRI 或 BglIII 做質體 insert DNA 的確認(依所使用的選殖系統而異)。以 0.8% agarose gel 分析 insert DNA 的片段大小。將含有適當大小 insert DNA 的質體送定序。

### 2.2.7 定序結果分析

- A. 依據 PCR 反應所使用的引子序列，將 insert DNA 序列分離出來。
- B. 檢查 insert DNA 序列中是否含有經初級限制酶切割後的黏合位(ligation site)，且所含黏合位的數目必須為奇數個。若為偶數個，表示此非源自目標基因的序列，為經初級限制酶片段化的基因體 DNA 隨機插入的序列，而並非 DNA 發生重排(rearrange)的產物。
- C. 將符合上述條件的序列送到 UCSC Genome Bioinformatics(<http://genome.ucsc.edu/>)的 Blat 做分析，即可得知與目標基因發生重排(rearrange)的對象為何。檢查與目標基因發生重排的對象，其 RefSeq 在非與目標基因發生融合(fusion)的那端，末端序列是否帶有初級限制酶切位，以確保此 rearranged DNA 片段，確實是由 RECORD 技術所偵測出來的。

## 參、實驗結果

### 3.1 RECORD 技術偵測敏感度的評估

由於本論文所使用的模板為人類精蟲細胞基因體 DNA(gDNA)，與一般臨床所使用的模板(如：血癌病患骨髓細胞 gDNA)，在來源及特性上有明顯的差異，因此，針對此部份設計了以 HEK293 gDNA 模擬精蟲細胞中正常 gDNA 的背景值，外加連續序列稀釋的 K562 gDNA 的實驗，藉由觀察不同稀釋倍數 K562 gDNA 之 *BCR-ABL* 片段被偵測到的狀況，以評估 RECORD 技術在精蟲細胞中對於目標基因結構變異的偵測敏感度。

K562 細胞株是由慢性骨髓性白血病患者所建立出來的細胞株，帶有費城染色體(Philadelphia chromosome)。費城染色體是位於第九對染色體的 *ABL* (v-abl Abelson murine leukemia viral oncogene) 與位於第二十二對染色體的 *BCR* (breakpoint cluster region isoform 1)發生基因轉位所形成的，產生致病性的融合蛋白。K562 細胞株染色體核型分析(karyotype)(圖一之 A)<sup>(17)</sup>顯示其第九及第二十二對染色體所發生的基因轉位屬於相互對調(reciprocal translocation)，產生兩種衍生染色體 der(22)(費城染色體)及 der(9)(圖一之 B)；此外，也具有正常的第九及第二十二對染色體。

K562 細胞株 *BCR-ABL* 斷裂點交界處在 gDNA 的層次已被定序<sup>(18)</sup>，斷裂點分別位於 *BCR* 的 intron 14 及 *ABL* 的 intron 1(圖二之 A)。然而此部份所定出的斷裂點交界處序列为 der(22)者，到目前為止，對於相對互調的另一個產物 der(9)則尚未在 gDNA 的層次被定序。因此，針對於 der(9)的部份，我們先以 der(22)的斷裂點位置作參考，分別還原兩者的序列。在進行實驗設計時，發現 der(9)的 *ABL* 序列有高密度的 repetitive element 分布(圖二之 B)，造成引子設計的困難，故只針對 der(22)

進行實驗，將引子設計在基因轉位斷裂點 5' 端的 *BCR* 序列上(表一)。

以初級限制酶 *Bgl*II 將 gDNA 片段化，在 HEK293 細胞株可切出約 6 kb un-rearranged form *BCR* 片段；在 K562 細胞株則可得到約 6 kb un-rearranged *BCR* 片段及約 7.6 kb 的 *BCR-ABL* 片段(圖二之 B)。使用二級限制酶 *Xho*I 降低 un-rearranged form *BCR* 的分子數。分別進行以 *BCR* 為目標基因的 inverse PCR(iPCR), nested PCR 及 *BCR-ABL* translocation specific PCR(TL PCR)。

以 100 ng K562 gDNA 做十倍連續序列稀釋當作模板，作為引子的敏感度測試(圖三之 A)。由 iPCR 結果，經 *Xho*I 處理的組別在以  $10^{-4}$  μg K562 gDNA 為模板時，仍可見 *BCR-ABL* 片段的產物(7.1 kb)，而未經 *Xho*I 處理者僅在以  $10^{-1}$  μg K562 gDNA 為模板時才可見 *BCR-ABL* 片段的產物(7.1 kb)。且二組在以  $10^{-1}$  μg K562 gDNA 為模板時，未經 *Xho*I 處理者顯現出較強的 *BCR* un-rearranged form(4.6 kb)，*BCR-ABL* rearranged form(7.1 kb)則較弱；而經 *Xho*I 處理者則顯現出相反的結果。雖然在有 *Xho*I 處理的組別仍有未被二級限制處理乾淨的 *BCR* un-rearranged form 被放大出來，但結果仍顯示出二級限制酶處理確實可達到提升 RECORD 技術偵測敏感度的效果。在 Nested PCR，未經過 *Xho*I 處理的組別完全呈現出 *BCR* un-rearranged form(3.3 kb)，推測此結果為 PCR 反應傾向於放大較小片段的模板所導致；而經 *Xho*I 處理後再以  $10^{-2} \sim 10^{-4}$  μg K562 gDNA 為模板時，完全顯現出 *BCR-ABL* rearranged form(5.8 kb)。TL PCR 的結果，顯示出引子的敏感度，以未經 *Xho*I 處理的 K562 gDNA 為模板時，敏感度為  $10^{-3}$  μg，而經過 *Xho*I 處理者則提升至  $10^{-4}$  μg。此部份實驗經過幾次重複，發現在以  $10^{-5}$  μg 經 *Xho*I 處理的 K562 gDNA 為模板時，在 TL PCR 有時會有產物出現，故敏感度應調整為  $10^{-4} \sim 10^{-5}$  μg。

以 100 ng K562 gDNA 做連續十倍序列稀釋外加 HEK293 gDNA 模擬正常 gDNA 的複雜度，來評估 RECORD 技術以精細胞 gDNA 為模板時的偵測敏感度(圖三之 B)。在 iPCR，無論有無 *Xho*I 處理，在以  $10^{-1}$  μg K562 gDNA 為模板時，均可見 *BCR-ABL* rearranged form(7.1 kb)，但產物強弱度因 *Xho*I 的處理而有提升的效果。

果。在 nested PCR，僅在有 XhoI 處理且單純只用  $10^{-1}\mu\text{g}$  K562 gDNA 為模板時，才稍微看見 *BCR-ABL* rearranged form(5.8 kb)的產物，顯示出正常 gDNA 的複雜度會影響偵測的敏感度。在 TL PCR，顯示出偵測敏感度，在未經 XhoI 處理的 K562 gDNA 為  $10^{-2}\mu\text{g}$ ，在經過 XhoI 處理的 K562 gDNA 則提升至  $10^{-4}\mu\text{g}$ 。

進一步將 gDNA 模板量換算成 amplifiable molecule 數目，一個含雙套(diploid)染色體的體細胞約有 6.6pg gDNA，故 100ng 體細胞 gDNA 含有約 30000 個 *BCR* amplifiable molecules，K562 細胞株一套染色體因發生染色體轉位的緣故，100ng K562 gDNA 則各含有約 15000 個 *BCR* 及 *BCR-ABL* 模板。在引子敏感度的部分，未經 XhoI 處理的情況下，偵測極限為 75 個 amplifiable molecules；經 XhoI 處理，偵測極限可提升至 2~15 個 amplifiable molecules，幾近於單一分子(single molecule)之檢測。在評估 RECORD 技術以精細胞 gDNA 為模板時的偵測敏感度方面，未經 XhoI 的處理，偵測極限為 1500 個 amplifiable molecules，換算成基因轉位發生頻率，約為  $5 \times 10^{-2}$  ( $750/15000$ )；經 XhoI 的處理，偵測極限為 15 個 amplifiable molecules，換算成基因轉位的發生頻率，約為  $5 \times 10^{-4}$  ( $15/30000$ )。

### 3.2 *TCF3* 之實驗設計

以 BamHI 及 BglII 作為初級限制酶在 *TCF3* 擷取約 5kb 的片段，此片段涵蓋 *TCF3* 的斷裂點叢聚區域(BCR)(intron 15)(圖四)。由於我們也想探討生殖細胞的基因轉位是否為相對互調(reciprocal translocation)的型式，故在此 5 kb 片段之 5' 端及 3' 端，依 *TCF3* 的轉錄方向(orientation)，分別設計前置(forward，3' 端)及後置(backward，5' 端)兩組引子(表二)，並將二組引子可分析的區域命名為 *TCF3-F*(3' 端)及 *TCF3-B*(5' 端)；此兩組引子分別包含內外兩對引子，提供 inverse PCR(1F & 2R) 及 nested PCR(3F & 4R) 之用；此外，由於利用 RECORD 偵測到的 rearranged DNA 模板，序列當中必須含有初級限制酶切位黏合的序列，因此另外設計了 3' 端可辨

認此黏合序列的引子(ligation site specific primer, LS primer)，希望能利用 PCR 反應達到篩選效果。並針對 *TCF3*-F(3' 端)及 *TCF3*-B(5' 端)，選用不同的二級限制酶將 unarranged form *TCF3* 自體環化(self-ligated)模板切除，以提升 rearranged form 模板被偵測到的機率。二級限制酶的選擇分別為：*TCF3*-F(3' 端)使用 HindIII 及 KpnI，*TCF3*-B(5' 端)使用 XhoI 及 XbaI。因此，可偵測到的斷裂點範圍座落於引子到二級限制酶切位之間；*TCF3*-F 的偵測範圍為 3.1 kb，*TCF3*-B 的偵測範圍為 2.9 kb，均涵蓋 *TCF3* 的 intron 12-intron 15。為了確保模板的品質，使用前置組引子的 *TCF3*-F-3F 及後置組引子的 *TCF3*-B-4R 作為控制組。

### 3.3 由人類精蟲細胞偵測 *TCF3* 結構變異

#### 3.3.1 模版有無二級限制酶處理之 5' 及 3' PCR 結果比較

此部分先以小量檢體比較二級限制酶作用對 PCR pattern 的影響，以作為之後做較大量檢體的參考。圖五之 A 為 *TCF3*-F(3' 端)及 *TCF3*-B(5' 端)的 PCR 結果。在 iPCR，*TCF3*-F 組(3' 端)因產物太少故看不見產物；*TCF3*-B 組(5' 端)在未經 XbaI 處理的部份，於 4.5 kb 處明顯可見 *TCF3* unarranged form 的產物，而經 XbaI 處理者則看不見此產物，且電泳結果也較乾淨。顯示二級限制酶作用除了可降低 unarranged form 的訊號，也可消除部分非特異性產物。在 nested PCR，顯示產物的特異性及產量均有提升，此外，不同來源的精蟲細胞 gDNA 模板，呈現出的 pattern 均不同。LS PCR 也反應出產物特異性的提升。分別對 nested PCR 及 LS PCR 的產物進行選殖與定序，在 *TCF3*-F 組(3' 端)及 *TCF3*-B 組(5' 端)分別偵測到 2 個染色體轉位。

此外，欲評估使用單一個二級限制酶(KpnI 或 XbaI)處理所得 LS PCR pattern 之真實性，故將 nested PCR 產物再以另一個二級限制酶(HindIII 或 XhoI)處理，比較二種情況的 LS PCR pattern 之異同。推測符合下列條件的 band，反應出 *TCF3*

rearranged form DNA 的真實性較高：(1)異於未經任何二級限制酶處理者所得之 pattern。(2)異於控制組(HEK293)所得之 pattern。(3)額外使用另一個二級限制酶處理後，仍持續存在的 band。圖五之 A 的 Sample A 其 TCF3-F 及 TCF3-B 的 LS PCR，以三角形標示處的 band 均符合上述條件，且經選殖與定序，確實可真實反應出 TCF3 rearranged form DNA。

在 Sample B TCF3-B(5'端)的 LS PCR，並未清楚看見特異性產物(圖五之 A)，進一步將方框部分的 DNA 片段分離出來(去除與未經二級限制酶處理者相同的 band)，稀釋後再進行一次 LS PCR(圖五之 B)，呈現出 2 個先前未看見的產物，將此產物進行選殖與定序，偵測到 2 個染色體轉位。

綜合以上結果，再次證實二級限制酶作用的重要性，故之後在做較大量檢體的分析時，將模板同時以二個限制酶做前處理(KpnI & HindIII，XbaI & Xhol)，以提升 PCR pattern 反應出 rearranged form DNA 的真實性；此外，假若模板在其中一個二級限制酶切位序列發生點突變而導致該限制酶不能作用時，仍可藉由另一個二級限制酶的作用來降低 unrearranged form 模板的訊號。

### 3.3.2 3' PCR 結果

圖六為 TCF3-F(3'端)之 PCR 結果。Sample A 與 Donor 3 在以  $10^{-3}$  倍稀釋的 iPCR 產物當模板時，nested PCR 未有明顯的 band 出現(實心倒三角形標示處)，推測應為模板濃度太高所致，故將 iPCR 產物再往下做 10 倍序列稀釋後做 nested PCR，即有明顯的 band 出現(空心倒三角形標示處)。此部分結果顯示出不同來源的精蟲細胞 gDNA 模板所得的 LS PCR pattern 有明顯的差異性存在，反應出個體間的不同(between individual difference)。

將 LS PCR 產物做選殖與定序，共偵測到 6 個染色體轉位。在 Sample D 偵測到 3 個大小不同的片段，經序列分析後，發現除了發生基因轉位的對象均相同之外，位於 TCF3 及基因轉位對象上的斷裂點位置亦相同，故為同一個基因轉位，推

測可能是因此 rearranged form DNA 模板的 3' 端在進行 PCR 時發生不同 mispriming 所導致。將此 6 個基因轉位的 PCR 產物大小(三角形標示處)與 LS PCR pattern 做比較，以評估 LS PCR pattern 反應出 *TCF3* rearranged form DNA 的真實性；雖然少部份 pattern 的大小與偵測到的基因轉位 PCR 產物大小符合，但大部分 pattern 選殖到的產物均為 PCR artifact，且尚有一部分 pattern 未被選殖到。由 Sample B 偵測到 3 個染色體轉位(圖六)，與其發生轉位的對象均不同，分別為 1q23.3，1q32.1 及 17q11.2，且位於 *TCF3* 上的斷裂點亦不同，反應出在同一個體內，與 *TCF3* 發生轉位的對象亦有差異性存在(within individual difference)；此結果似乎也暗示著在同一個體內，不同細胞發生的染色體轉位均有差異(cellular level difference)。

### 3.3.3 5' PCR 結果

圖七為 *TCF3-B*(5' 端)之 PCR 結果，顯示出不同來源精蟲細胞 gDNA 之 LS PCR pattern 有明顯的差異存在，反應出個體間的不同。將 LS PCR 產物做選殖與定序，共偵測到 7 個染色體轉位。由 Sample B 偵測到的 2 個染色體轉位，與其發生轉位的對象均不同，分別為 4q22.1 及 8q11.23，且位於 *TCF3* 上的斷裂點亦不同，顯示在同一個體內，*TCF3* 發生轉位的對象亦有差異性存在，反應出單一細胞層次的不同(cellular level difference)。

### 3.3.4 推斷之 *TCF3* 轉位(putative translocations)

在 *TCF3* 共偵測到 19 個染色體轉位；使用前置(forward，3'端)引子偵測出 8 個，使用後置(backward，5'端)引子偵測出 11 個。

由前置(forward，3'端)引子偵測出的 8 個染色體轉位

1. t(1;19)(q23.3;p13.3)

來源為 Sample B 的 LS PCR 產物，片段大小為 836 bp。為 *TCF3* 與 1q23.3 的非基因區(intergenic region)發生的轉位(圖八之一)，斷裂點分別位於 *TCF3* 的 intron

13 及 1q23.3 的 *CDCA1* 與 *PBX1* 之間，在斷裂點交界處具有 5 個核苷酸的微同源性(microhomology)序列。經序列還原，推測偵測到的轉位產物為 der(19)。

#### 2. t(1;19)(q32.1;p13.3)

來源為 Sample B 的 LS PCR 產物，片段大小為 647 bp。為 *TCF3* 與 1q32.1 的非基因區發生的轉位(圖八之二)，斷裂點分別位於 *TCF3* 的 intron 12 及 1q32.1 的 *FLJ10901* 與 *CACNA1S* 之間，在斷裂點交界處具有 3 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為二中節染色體(dicentric chromosome)。

#### 3. t(10;19)(q22.1;p13.3)

來源為 Sample A 的 LS PCR 產物，片段大小為 421 bp。為 *TCF3* 與位於 10q22.1 的 *CBARA1* 發生的轉位(圖八之三)，斷裂點分別位於 *TCF3* 的 intron 12 及 *CBARA1* 的 intron 3，在斷裂點交界處具有 3 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為二中節染色體。

#### 4. t(18;19)(q21.1;p13.3)

來源為 Sample A 的 LS PCR 產物，片段大小為 382 bp。為 *TCF3* 與位於 18q21.1 的 *DCC* 發生的轉位(圖八之四)，斷裂點分別位於 *TCF3* 的 intron 12 及 *DCC* 的 intron 2，在斷裂點交界處具有 7 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為二中節染色體。

#### 5. t(3;19)(p22.2;p13.3)

來源為 Sample D 的 LS PCR 產物，因 PCR 時模板 3' 端發生不同的斷裂，形成三種不同大小的片段，分別為 358 bp，947 bp 及 989 bp。為 *TCF3* 與 3p22.2 的非基因區發生的轉位(圖八之五)，斷裂點分別位於 *TCF3* 的 intron 12 及 3p22.2 的 *MOBP* 與 *MYRIP* 之間，在斷裂點交界處具有 7 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(19)。

#### 6. t(17;19)(q11.2;p13.3)

來源為 Sample B 的 LS PCR 產物，片段大小為 547 bp。為 *TCF3* 與 17q11.2 的

非基因區發生的轉位(圖八之六)，斷裂點分別位於 *TCF3* 的 intron 12 及 17q11.2 的 *MGC9564* 與 *SLC13A2* 之間，在斷裂點交界處具有 5 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(19)。

#### 7. t(3;19)(p21.31;p13.3)

來源為 Donor 3 的 LS PCR 產物，片段大小為 212 bp。為 *TCF3* 與 3p21.31 的非基因區發生的轉位(圖八之七)，斷裂點分別位於 *TCF3* 的 intron 12 及 3p21.31 的 *C3orf23* 與 *ZNF445* 之間，在斷裂點交界處具有 6 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(19)。

#### 8. t(5;19)(q35.2;p13.3)

來源為 Donor 5 的 LS PCR 產物，片段大小為 434 bp。為 *TCF3* 與 5q35.2 的非基因區發生的轉位(圖八之八)，斷裂點分別位於 *TCF3* 的 intron 12 及 5q35.2 的 *HMP19* 與 *MSX2* 之間，在斷裂點交界處具有 1 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為二中節染色體。

### 由後置(Backward)引子偵測出的 11 個染色體轉位

#### 1. t(12;19)(q21.31;p13.3)

來源為 Sample B 的 nested PCR 產物，片段大小為 139 bp。為 *TCF3* 與位於 12q21.31 的 *TMTC2* 發生的轉位(圖九之一)，斷裂點分別位於 *TCF3* 的 intron 15 及 *TMTC2* 的 intron 4，在斷裂點交界處具有 10 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為無中節染色體(acentric chromosome)。

#### 2. t(8;19)(q11.23;p13.3)

來源為 Sample C 的 LS PCR 產物，片段大小為 289 bp。為 *TCF3* 與位於 8q11.23 的 *RB1CC1* 發生的轉位(圖九之二)，斷裂點分別位於 *TCF3* 的 intron 15 及 *RB1CC1* 的 intron 1，在斷裂點交界處具有 5 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(8)。

#### 3. t(11;19)(p13;p13.3)

來源為 Sample A 的 LS PCR 產物，片段大小為 601 bp。為 *TCF3* 與 11p13 的非基因區發生的轉位(圖九之三)，斷裂點分別位於 *TCF3* 的 intron 15 及 11p13 的 *MGC34830* 與 *HIPK3* 之間，在斷裂點交界處具有 5 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為無中節染色體。

#### 4. t(2;19)(p24.3;p13.3)

來源為 Sample B 的 LS PCR 產物，片段大小為 1431 bp。為 *TCF3* 與 2p24.3 的非基因區發生的轉位(圖九之四)，斷裂點分別位於 *TCF3* 的 intron 15 及 2p24.3 的 *TRIB1* 與 *NSE1* 之間，在斷裂點交界處具有 8 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(2)。

#### 5. t(11;19)(q13.3;p13.3)

來源為 Donor 5 的 LS PCR 產物，片段大小為 629 bp。為 *TCF3* 與 11q13.3 的非基因區發生的轉位(圖九之五)，斷裂點分別位於 *TCF3* 的 intron 15 及 11q13.3 的 *CPT1A* 與 *MRPL21* 之間，在斷裂點交界處具有 5 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(11)。

#### 6. t(8;19)(q22.2;p13.3)

來源為 Sample C 的 LS PCR 產物，片段大小為 1019 bp。為 *TCF3* 與 8q22.2 的非基因發生的轉位(圖九之六)，斷裂點分別位於 *TCF3* 的 intron 15 及 8q22.2 的 *RNF19* 與 *ANKRD46* 之間，在斷裂點交界處具有 7 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為無中節染色體。

#### 7. t(6;19)(q21;p13.3)

來源為 Sample E 的 LS PCR 產物，片段大小為 529 bp。為 *TCF3* 與 6q21 的非基因區發生的轉位(圖九之七)，斷裂點分別位於 *TCF3* 的 intron 15 及 6q21 的 *PREP* 與 *PRDM1* 之間，在斷裂點交界處具有 2 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為無中節染色體。

#### 8. t(10;19)(p12.1;p13.3)

來源為 Donor 2 的 LS PCR 產物，片段大小為 586 bp。為 *TCF3* 與 10p12.1 的非基因區發生的轉位(圖九之八)，斷裂點分別位於 *TCF3* 的 intron 15 及 10p12.1 的 *WAC* 與 *BAMBI* 之間，在斷裂點交界處具有 11 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為無中節染色體。

#### 9. t(4;19)(q22.1;p13.3)

來源為 Sample C 的 LS PCR 產物，片段大小為 666 bp。為 *TCF3* 與 4q22.1 的非基因區發生的轉位(圖九之九)，斷裂點分別位於 *TCF3* 的 intron 15 及 4q22.1 的 *TIGD2* 與 *GPRIN3* 之間，在斷裂點交界處具有 6 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(4)。

#### 10. t(1;19)(q32.3;p13.3)

來源為 Sample B 的 LS PCR 產物，片段大小為 653 bp。為 *TCF3* 與 1q32.3 的非基因區發生的轉位(圖九之十)，斷裂點分別位於 *TCF3* 的 intron 15 及 1q32.3 的 *NEK2* 與 *SLC30A1* 之間，在斷裂點交界處具有 13 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為無中節染色體。

#### 11. t(2;19)(q33.1;p13.3)

來源為 Sample B 的 LS PCR 產物，片段大小為 1431 bp。為 *TCF3* 與 2q33.1 的非基因區發生的轉位(圖九之十一)，斷裂點分別位於 *TCF3* 的 intron 15 及 2q33.1 的 *PLCL1* 與 *SATB2* 之間，在斷裂點交界處具有 7 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為無中節染色體。

### 3.3.5 染色體轉位斷裂點之定位

經上述分析，發現偵測到的染色體轉位均具有一個明顯的特性，即是在斷裂點交界處均具有一小段無論是在 *TCF3* 或是其轉位對象，序列皆相同的核苷酸，稱為微同源性序列(microhomology sequence)，長度約在 1~13 個核苷酸不等。過去的研究指出，減數分裂時的 DNA 雙股斷裂(double strand breaks, DSBs)是由具有

Topoisomerase II 活性的 Spo11 所產生<sup>(19)</sup>。在以酵母菌及小鼠為模式的實驗，已證實減數分裂 DNA 雙股斷裂後，會先經目前仍未知的核酸內切酶(endonuclease)進行不對稱切割(asymmetric nicking)，再經 5'→3' 核酸酶外切酶(5'→3' exonuclease)切除(resection)後進行修復(圖十之 A)<sup>(20)</sup>。由於我們無法確定 DNA 雙股斷裂的位置及在發生雙股斷裂後，其末端如何處理，才使微同源性序列暴露出來，故在定位斷裂點位置時，先假設 DNA 發生雙股斷裂後，經 5'→3' 切除後，直接使微同源性序列暴露出來。

而在 DNA 雙股斷裂的方式方面，我們假設斷裂方式與限制酶的作用相似，故產生的雙股斷裂末端可能為平滑末端(blunt ends)或黏性末端(cohesive ends)，而黏性末端又有 5' cohesive ends 及 3' cohesive ends 兩種形式。由於在哺乳類體細胞對 DNA 雙股斷裂的修復機制主要有同源性重組(Homologous recombination, HR)及非同源性末端接合(Non-homologous end joining, NHEJ)，而此兩種機制在 DNA 雙股斷裂的末端均會先進行 5'→3' 切除再進行修復，因此我們假設在人類精蟲細胞偵測到的染色體轉位在其雙股斷裂末端，亦會先進行 5'→3' 切除再進行修復(圖十之 A)。此外，考慮 DNA 發生雙股斷裂及 5'→3' 切除後，無論是在 TCF3 或是其轉位對象，微同源性序列均能被保留下來，故可排除雙股斷裂末端為 5' 黏性末端(5' cohesive ends)的可能性。此外，無論是由後置(5' 端)或前置(3' 端)引子所偵測出之染色體轉位均具有微同源性序列，故推測 DNA 雙股斷裂末端應為 3' 黏性末端(3' cohesive ends)，才能確保位於正股(top strand)或是負股(bottom strand)的微同源性序列均得以被保留(圖十之 B)。

因此，我們在定位染色體轉位斷裂點時，遵循以下規則：(1)DNA 發生雙股斷裂後，可能產生的末端為 3' 黏性末端，且微同源性序列位於斷裂端的末端處。(2)不論目標基因的轉錄方向為何，均將其序列以正股的方式呈現，而其轉位對象序列可能為正股或是負股；以圖十之 C 為例，由於已將目標基因的序列以正股表示，故將其斷裂點定在 A(方框標示處)，若轉位對象(TP)序列為正股(圖十之 C-a)，將斷

裂點定在 A(方框標示處)；若轉位對象(TP)序列為負股(圖十之 C-b)，則將斷裂點定在 G(方框標示處)。

依據上述規則，將我們在人類精蟲細胞中偵測到的 *TCF3* 染色體轉位，在 *TCF3* 及其轉位對象上之斷裂點在染色體上的確切位置標定出來(表三)。

### 3.3.6 *TCF3* 染色體轉位斷裂點的分布

將我們在人類精蟲細胞中偵測到的 *TCF3* 染色體轉位斷裂點分布以 GNS 展示出來，並與文獻所發表的斷裂點分布作比較(圖十一)。雖然由 *TCF3*-F(3' 端)及 *TCF3*-B(5' 端)所偵測出的染色體轉位，其斷裂點均呈密集分布；由 *TCF3*-F 偵測出的染色體轉位在 *TCF3* 的斷裂點分布於 intron 15 呈現密集分布的情形，而由 *TCF3*-B 偵測出者，則在 *TCF3* intron 12 處有分布密集的情形，但並無任何斷裂點位於斷裂點叢聚區域。此外，也並未偵測到相對互調(reciprocal translocation)形式的染色體轉位。

其中，以後置(5' 端)引子在 Sample B 偵測到的 t(2;19)(q33.1;p13.3)(圖十一空心三角形標示處)，其位於 *TCF3* 之斷裂點位於二級限制酶切位之外側，經序列分析，發現序列中帶有 HindIII 之切位，推測此結果為 HindIII 作用不完全所致。另外，以前置(3' 端)引子在 Sample B 偵測到的 t(1;19)(q23.3;p13.3)(圖十一空心三角形標示處)，其位於 *TCF3* 之斷裂點亦位於二級限制酶切位之外側，經序列分析，發現 XhoI 之切位發生點突變(CTCGAG→CCCGAG)，而導致 XhoI 無法作用。除上述二者之外，其餘染色體轉位之斷裂點均位於二級限制酶切位之內側。

此外，分別將由不同來源精細胞 gDNA 偵測出的染色體轉位在 *TCF3* 上的斷裂點及其轉位對象展示出來；可觀察到這些染色體轉位除了有個體之間(between individuals)之差異外，亦有出個體本身(within individual)的差異，反應出單一細胞層次之不同(cellular level difference)。

### 3.3.7 TCF3 染色體轉位序列的分析

染色體轉位的發生導因於細胞對 DNA 雙股斷裂的修復，在哺乳類細胞有兩種主要的機制可修復 DNA 的雙股斷裂，分別為同源性重組(HR)及非同源性末端接合(NHEJ)。同源性重組需有同源性序列(homologous sequence)存在才能發生，而同源性序列可源自姐妹染色分體(sister chromatids)、同源染色體(homologous chromosomes)或基因體中的重複序列(repetitive elements)。由於我們偵測到的均為非同源染色體間的轉位，故排除同源性序列源自姐妹染色分體及同源染色體的可能性。因此，我們針對 TCF3 及其轉位對象在斷裂點周邊重複序列的分布作分析，以探討這些染色體轉位是否為重複序列媒介的同源性重組(repeat-mediated HR)所造成的。

重複序列在人類基因體中所佔的比例超過 50%，其中又以來自轉座子的重複序列(transposable elements, TEs)所佔的比例最高(約佔人類基因體 45%)<sup>(21)</sup>。而 TEs 可再區分為四類，分別為 LINEs、SINEs、LTR 反轉座子(LTR retrotransposons)及 DNA 轉座子(DNA transposons)。其中，SINEs 在人類基因體的拷貝數目約有 150 萬個，為四類 TEs 中最多的，故推測假若同源性重組屬於重複序列媒介的，應以 SINE-mediated HR 的出現頻率較高。因此我們將 TCF3 及其轉位對象在斷裂點上下各 2 kb 序列的 SINE 分布以 GNS 展示出來(圖十二)，結果並未呈現出 SINE 在二者上有對應分布的情形，故初步排除偵測到的染色體轉位為重複序列媒介的同源性重組(repeat-mediated HR)之可能性。

而在對 TCF3-B(5'端)所偵測到的染色體轉位序列作 Blat 分析時，我們注意到在 TCF3 上有一段 205 bp 的序列(在此先稱之為 T-element)(圖十三之 A)，與之相似度高的序列幾乎均座落在第十九對染色體短臂 13.3(19p13.3)的區域(圖十三之 B)，數目為 76 個。因此，推測 T-element 可能為人類基因體中的重複序列(repetitive element)。經 Repbase(<http://www.girinst.org/repbase/index.html>)的比對，發現此段序列與 DNA 轉座子中的 MER20(medium reiteration frequency repeat 20)相似(圖十三之 C)。欲更進一步探討 T-element 是否等同於 MER20，我們對 T-element 與 MER20

在 19p13.3 的分布情形做比較，另外也分析此二者在與人類基因體序列相似度高達 98.8%的黑猩猩(chimpanzee)的 19p13.3 之分布情形，作為重複確認(圖十四)。在人類及黑猩猩的 19p13.3，T-element 的分布分別有 3 處及 2 處呈現叢聚分布的情形，且此叢聚分布與 MER20 的分布吻合，故知 T-element 應等同於 MER20。

此外，由於以 T-element 作為 Blast 分析的 Query 時，找到的相似序列，幾乎均座落於 19p13.3，我們好奇此分布狀況在人類基因體中是否僅侷限於此區域，故將 MER20 在人類基因體中的分布以 GNS 展示出來(圖十五)，結果顯示，MER20 在人類基因體中的拷貝數約兩萬個，且其密集分布的情形並非只侷限於 19p13.3。雖然 MER20 在基因體中的拷貝數相較於 SINE 少很多，但為了確認 MER20 是否與 *TCF3*-B(5' 端)偵測到的染色體轉位相關，故將 *TCF3* 及其轉位對象在斷裂點上下各 2 kb 序列的 MER20 分布以 GNS 展示出來(圖十六)。由於在此區域內均無 MER20 分布，因此可知這些染色體轉位並非 MER20-mediated HR 造成的。

### 3.4 染色體轉位確認試驗之實驗設計

我們已利用 RECORD 技術在人類生殖細胞中偵測到染色體轉位，但仍需以轉位特異性引子(translocation-specific primers, TL primers)進行 PCR 反應來加以證實。我們已先後針對 *EWSR1*、*MLL*(from 支若馨學姊)及 *TCF3* 三個 locus 做探討，但僅有 *MLL* 這個 locus 偵測到的染色體轉位斷裂點有明顯叢聚分布的情形，且分布區域與文獻發表的斷裂點位置十分相近，特別是在 *MLL*-F2(fragment 2)(圖十七)，因此我們針對 *MLL*-F2(表四)偵測到的染色體轉位做確認試驗，並先著重於由 Sample B 偵測到的 3 個染色體轉位。

在實驗設計部分，考量若以線性 gDNA 作為模版(linear template)時，不同染色體轉位需設計不同引子，可能因不同引子在進行 PCR 反應的效率不同而影響實驗結果。為排除此顧慮，故採用環化的 gDNA 當做模版(circular template)，優點如下：

(1)可沿用先前設計於目標基因上用於偵測基因轉位的引子，避免因不同引子效率的差異而影響實驗結果。(2)不同的染色體轉位可初步使用相同的引子同時進行大，再以轉位特異性引子(TL primers)進行 PCR 反應來加以證實。(3)PCR 結果或許能反應出生殖細胞的 DNA 重排圖譜(germline rearranged DNA profiles)，除了欲做確認的染色體轉位之外，或許也可反應出先前未偵測到的結構變異。

此外，也針對 *MLL*-F2 設計後置(3'端)引子，以探討生殖細胞的 *MLL* 染色體轉位是否為相對互調(reciprocal translocation)的形式。使用位於 *MLL*-F2 的兩組引子(表五)，並依 *MLL* 的轉錄方向，將此二組引子可分析的區域命名為 *MLL*-F2-F(5'端)及 *MLL*-F2-B(3'端)。模板前處理方式如下：使用 BamHI 及 BglII 將生殖細胞 gDNA 片段化再環化後，以二級限制酶處理，分別為：*MLL*-F2-F(5'端)使用 NsiI，*MLL*-F2-B(3'端)使用 BsmI 及 AfeI。進行確認試驗時，先以位於 *MLL* 上的引子做初步的 iPCR 放大，再以 TL primers 進行 PCR 反應來加以證實。此外，也將 iPCR 產物做 nested PCR 及 LS PCR，並將 LS PCR 產物選殖並定序分析，藉以評估此種方式獲得的生殖細胞 DNA 重排圖譜(germline rearranged DNA profiles)其可信度有多高。

### 3.5 確認試驗之結果

#### 3.5.1 *MLL* 染色體轉位的確認

針對 3 個由 Sample B 偵測到的 *MLL* 染色體轉位進行確認試驗，以 100ng gDNA 為起始模板，做 10 次重覆。取 *MLL*-F2-F iPCR 產物作為模版，以 3 組轉位特異性引子進行 TL PCR，分別為：*MLL*-F2-3F/*NOTCH2NL*-V，*MLL*-F2-3F/8q11.21-V 及 *MLL*-F2-3F/*PIK3R4*-V(表六)。10 次重覆的結果，TL PCR 均無任何一次出現產物(圖十九)。以先前報導人類生殖細胞的復發性基因轉位 t(11;22)(q23;q11)作參考，其發生頻率為  $1.24\text{--}9.46 \times 10^{-5}$ <sup>(16)</sup>，與 RECORD 的偵測極限  $5 \times 10^{-4}$  相比，二者約相差十倍，

故推測起始模板的量至少需增加 10 倍才有出現轉位特異性產物的可能。此外，由於我們偵測到的染色體轉位其轉位對象均不相同，與上述生殖細胞的復發性基因轉位  $t(11;22)(q23;q11)$  明顯不同，並未呈現出與特定轉位對象發生轉位的傾向。故推測其發生頻率可能較  $t(11;22)(q23;q11)$  低，若要對此進行確認試驗，除了需增加重複次數之外，亦需再找尋其他的取代方法以增加偵測靈敏度，如：提升轉位模板的信號(如：Southern blot)或設法再將模版的複雜度降低(如：streptavidin-biotin purification system)等。

### 3.5.2 評估生殖細胞 DNA 重排圖譜之可行性

首先，先針對生殖細胞 DNA 重排圖譜是否可行進行評估，故使用先前偵測出 *MLL* 染色體轉位的 iPCR 產物及帶有染色體轉位序列的質體(positive control，PC) 當作模版，以 LS primer 進行 PCR 反應，比較前者的產物大小是否可與後者的產物大小互相對應(圖二十)。在 Sample B、E 及 D，LS PCR pattern 均能和其 PC 相對應，尤其是在 Sample B，LS PCR pattern 與 PC 產物大小完全符合，因此可知，以此種方式獲得 DNA 重排圖譜的想法確實可行。

### 3.5.3 生殖細胞 DNA 重排圖譜之真實性

使用 100ng 經二級限制酶處理過之環化 gDNA 為模版，以 *MLL-F2-F* 之引子進行 iPCR，nested PCR 及 LS PCR，並將 LS PCR 產物選殖並定序分析，評估此種方式獲得的生殖細胞 DNA 重排圖譜(germline rearranged DNA profiles)之真實性，並比較不同次實驗所得圖譜是否有差異。

圖二十一為不同次實驗所得之生殖細胞 DNA 重排圖譜，由實驗結果發現，不同來源 gDNA 所得的圖譜均不相同，這似乎意味著生殖細胞的 DNA 重排具有個體間的差異；此外，不同次實驗呈現出的圖譜差異頗大。經選殖定序分析後，共偵測到 14 個不同的染色體轉位(表七)，其中並無與先前偵測到的染色體轉位重複者。

推測可能原因如下：(1)我們在精蟲細胞偵測到的染色體轉位，其發生頻率可能小於或與 RECORD 技術偵測極限相近(borderline)，以致相同來源的模板在不同 aliquots 偵測到的染色體轉位均不相同。(2)當精蟲細胞的特定基因發生雙股斷裂後，可能是以隨機的(stochastic)方式與其他發生雙股斷裂的 DNA 發生轉位，且受限於 RECORD 技術偵測極限，而導致並未偵測到具有相同轉位對象的染色體轉位。

而在 *MLL*-F2-B 之偵測部份，共偵測到 6 個染色體轉位(圖二十二)(表七)。其中並未偵測到相對互調(reciprocal translocation)形式的轉位。

此部分實驗結果顯示，部分生殖細胞 DNA 重排圖譜的 pattern 確實能反應出生殖細胞的 DNA 重排，但圖譜中其他未反應出生殖細胞 DNA 重排者，經定序分析，證實大部分為 PCR artifact，且尚有一部份 pattern 未被選殖到。由於使用 PCR based approach，故我們無法完全避免因引子的 mispriming 所產生的 PCR artifact，因此會影響生殖細胞 DNA 重排圖譜反應出發生重排的 DNA 之真實性。

### 3.5.4 確認試驗由 *MLL*-F2 偵測到的染色體轉位

進行確認試驗時，在 *MLL*-F2 共偵測到 20 個染色體轉位；使用前置(forward，5'端)引子偵測出 14 個，使用後置(backward，3'端)引子偵測出 6 個。

由前置(forward，5'端)引子偵測出的 14 個染色體轉位

#### 1. t(10;11)(p12.1;q23.3)

來源為 Sample C 的 LS PCR 產物，片段大小為 153 bp。為 *MLL* 與 10p12.1 的非基因區發生的轉位(圖二十三之一)，斷裂點分別位於 *MLL* 的 intron 9 及 10p12.1 的 *KIAA1217* 與 *ARHGAP21* 之間，在斷裂點交界處具有 5 個核苷酸的微同源性(microhomology)序列。經序列還原，推測偵測到的轉位產物為 der(11)。

#### 2. t(3;11)(p22.1;q23.3)

來源為 Sample D 的 LS PCR 產物，片段大小為 442 bp。為 *MLL* 與 3p22.1 的非基因區發生的轉位(圖二十三之二)，斷裂點分別位於 *MLL* 的 intron 9 及 3p22.1

的 *ZNF621* 與 *CTNNB1* 之間，在斷裂點交界處具有 2 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(11)。

### 3. t(1;11)(p32.1;q23.3)

來源為 Sample E 的 LS PCR 產物，片段大小為 967 bp。為 *MLL* 與 1p32.1 的非基因區發生的轉位(圖二十三之三)，斷裂點分別位於 *MLL* 的 intron 9 及 1p32.1 的 *TACSTD2* 與 *JUN* 之間，在斷裂點交界處具有 4 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(11)。

### 4. t(9;11)(p24.3;q23.3)

來源為 Sample C 的 LS PCR 產物，片段大小為 702 bp。為 *MLL* 與位於 9p24.3 的 *CBWD1* 發生的轉位(圖二十三之四)，斷裂點分別位於 *MLL* 的 exon 10 及 *CBWD1* 的 intron 4，在斷裂點交界處具有 5 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(11)。

### 5. t(4;11)(q35.1;q23.3)

來源為 Sample A 的 LS PCR 產物，片段大小為 527 bp。為 *MLL* 與 4q35.1 的非基因區發生的轉位(圖二十三之五)，斷裂點分別位於 *MLL* 的 exon 10 及 4q35.1 的 *MGC45800* 與 *DCTD* 之間，在斷裂點交界處具有 5 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為二中節染色體(dicentric chromosome)。

### 6. t(4;11)(q25;q23.3)

來源為 Sample C 的 LS PCR 產物，片段大小為 447 bp。為 *MLL* 與 4q25 的非基因區發生的轉位(圖二十三之六)，斷裂點分別位於 *MLL* 的 exon 10 及 4q25 的 *SCYE1* 與 *DKK2* 之間，在斷裂點交界處具有 3 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(11)。

### 7. t(3;11)(q23;q23.3)

來源為 Sample D 的 LS PCR 產物，片段大小為 442 bp。為 *MLL* 與位於 3q23 的 *CLSTN2* 發生的轉位(圖二十三之七)，斷裂點分別位於 *MLL* 的 exon 10 及 *CLSTN2*

的 intron 5，在斷裂點交界處具有 6 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為二中節染色體。

#### 8. t(5;11)(p12;q23.3)

來源為 Sample D 的 LS PCR 產物，片段大小為 679 bp。為 *MLL* 與位於 5p12 的 *GHR* 發生的轉位(圖二十三之八)，斷裂點分別位於 *MLL* 的 intron 10 及 *GHR* 的 intron 1，在斷裂點交界處具有 6 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為二中節染色體。

#### 9. t(11;18)(q23.3;q12.2)

來源為 Donor 4 的 LS PCR 產物，片段大小為 492 bp。為 *MLL* 與位於 18q12.2 的 *BRUNOL4* 發生的轉位(圖二十三之九)，斷裂點分別位於 *MLL* 的 intron 10 及 *BRUNOL4* 的 intron 1，在斷裂點交界處具有 6 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(11)。

#### 10. t(11;14)(q23.3;q31.1)

來源為 Donor 5 的 LS PCR 產物，片段大小為 1065 bp。為 *MLL* 與 14q31.1 的非基因區發生的轉位(圖二十三之十)，斷裂點分別位於 *MLL* 的 intron 10 及 14q31.1 的 *SELL* 與 *FLRT2* 之間，在斷裂點交界處具有 6 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為二中節染色體。

#### 11. t(3;11)(q22.3;q23.3)

來源為 Donor 5 的 LS PCR 產物，片段大小為 769 bp。為 *MLL* 與位於 3q22.3 的 *PIK3CB* 發生的轉位(圖二十三之十一)，斷裂點分別位於 *MLL* 的 intron 10 及 *PIK3CB* 的 intron 12，在斷裂點交界處具有 4 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(11)。

#### 12. t(10;11)(q26.2;q23.3)

來源為 Sample B 的 LS PCR 產物，片段大小為 1007 bp。為 *MLL* 與位於 10q26.2 的 *PTPRE* 發生的轉位(圖二十三之十二)，斷裂點分別位於 *MLL* 的 intron 10 及

*PTPRE* 的 intron 10，在斷裂點交界處具有 8 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為二中節染色體。

#### 13. t(1;11)(q21.1;q23.3)

來源為 Sample B 的 LS PCR 產物，片段大小為 1578 bp。為 *MLL* 與 1q21.1 的非基因區發生的轉位(圖二十三之十三)，斷裂點分別位於 *MLL* 的 intron 10 及 1q21.1 的 *SEC22L1* 與 *NOTCH2NL* 之間，在斷裂點交界處具有 9 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(11)。

#### 14. t(11;15)(q23.3;q22.2)

來源為 Sample D 的 LS PCR 產物，片段大小為 1366 bp。為 *MLL* 與 15q22.2 的非基因區發生的轉位(圖二十三之十四)，斷裂點分別位於 *MLL* 的 intron 10 及 14q31.1 的 *FLJ38723* 與 *TLN2* 之間，在斷裂點交界處具有 10 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為二中節染色體。

#### 由後置(backward, 3'端)引子偵測出的 6 個染色體轉位

##### 1. t(11;18)(q23.3;q12.2)

來源為 Donor 4 的 LS PCR 產物，片段大小為 1450 bp。為 *MLL* 與 18q12.2 的非基因區發生的轉位(圖二十四之一)，斷裂點分別位於 *MLL* 的 intron 10 及 18q12.2 的 *BRUNOL4* 與 *PIK3C3* 之間，在斷裂點交界處具有 4 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為無中節染色體(acentric chromosome)。

##### 2. t(4;11)(q26;q23.3)

來源為 Donor 5 的 LS PCR 產物，片段大小為 1051 bp。為 *MLL* 與位於 4q26 的 *MYOZ2* 發生的轉位(圖二十四之二)，斷裂點分別位於 *MLL* 的 intron 10 及 *MYOZ2* 的 intron 4，在斷裂點交界處具有 2 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(4)。

##### 3. t(4;11)(p12;q23.3)

來源為 Donor 5 的 LS PCR 產物，片段大小為 1039 bp。為 *MLL* 與位於 4p12

的 *GABRA2* 發生的轉位(圖二十四之三)，斷裂點分別位於 *MLL* 的 intron 10 及 *GABRA2* 的 intron 8，在斷裂點交界處具有 5 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為無中節染色體。

#### 4. t(2;11)(p24.3;q23.3)

來源為 Sample C 的 LS PCR 產物，片段大小為 441 bp。為 *MLL* 與 2p24.3 的非基因區發生的轉位(圖二十四之四)，斷裂點分別位於 *MLL* 的 intron 10 及 2p24.3 的 *TRIB2* 與 *NSE1* 之間，在斷裂點交界處具有 5 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為無中節染色體。

#### 5. t(11;14)(q23.3;q31.1)

來源為 Donor 5 的 LS PCR 產物，片段大小為 744 bp。為 *MLL* 與 14q31.1 的非基因區發生的轉位(圖二十四之五)，斷裂點分別位於 *MLL* 的 intron 10 及 14q31.1 的 *SEL1L* 與 *FLRT2* 之間，在斷裂點交界處具有 6 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為無中節染色體。

#### 6. t(11;16)(q23.3;q21)

來源為 Donor 5 的 LS PCR 產物，片段大小為 257 bp。為 *MLL* 與 16q21 的非基因區發生的轉位(圖二十四之六)，斷裂點分別位於 *MLL* 的 intron 10 及 16q21 的 *GOT2* 與 *CDH8* 之間，在斷裂點交界處具有 2 個核苷酸的微同源性序列。經序列還原，推測偵測到的轉位產物為 der(16)。

### 3.5.5 *MLL*-F2 染色體轉位斷裂點的分布

將我們在人類精蟲細胞中偵測到的 *MLL*-F2 染色體轉位斷裂點分布以 GNS 展示出來，並與文獻所發表的斷裂點分布作比較(圖二十三)。由前置(5'端)及後置(3'端)引子所偵測出的染色體轉位，斷裂點均密集分布於 *MLL* intron 9 至 intron 10 之間，且分布狀況與文獻發表之斷裂點分布狀況相似。但並未偵測到使用相同微同源性(microhomology)序列之相對互調(reciprocal translocation)形式的染色體轉位。

此外，分別將由不同來源精細胞 gDNA 偵測出的染色體轉位在 *MLL*-F2 上的斷裂點及其轉位對象展示出來；可觀察到這些染色體轉位除了有個體間(between individuals)之差異外，亦有個體本身(within individual)的差異，反應出單一細胞層次之不同(cellular level difference)。而由 Sample D 偵測出的兩個染色體轉位(表九灰底標示處)，在 *MLL*-F2 的斷裂點位置十分相近，且微同源性(microhomology)序列有部份相同，但卻與不同的對象發生染色體轉位，顯示出在同一個體內，即使 DNA 雙股斷裂的位置幾近相同，暴露出相同的微同源性序列，但亦有可能與不同的對象發生染色體轉位，再一次反應了此種差異為單一細胞層次之不同(cellular level difference)。

### 3.5.6 精蟲細胞染色體轉位之對象在各染色體長短臂上的分布

進一步分析偵測到的精蟲細胞染色體轉位其對象在各染色體長短臂上的分布數量(表八)。若染色體轉位是以隨機的(stochastic)方式發生，理論上應會受染色體長度影響，長度越長發生的機率應該越高。轉位對象分布在染色體短臂及長臂上的比值(p/q ratio)，在 *TCF3* 及 *MLL* 分別為 0.36 及 0.35，而基因體中染色體短臂及長臂長度之比值約為 0.5，雖並未相等，但推測若偵測到的染色體轉位數量更多時，p/q ratio 應會更接近 0.5。

### 3.5.7 精蟲細胞染色體轉位形式為相對互調的可能性

由 Donor 5 在 *MLL*-F2 偵測到的兩個染色體轉位(表七淺灰底標示者) t(11;14)(q23.3;q31.1)，在 *MLL*-F2-F 及 *MLL*-F2-B 均有偵測到，在 *MLL* 上之斷裂點相距約 1400 bp，二者轉位對象均相同，位於 *SEL1L* 及 *FLRT2* 之間的非基因區(intergenic region)，且推斷之染色體轉位產物(putative translocations)，由 *MLL*-F2-F 偵測到的為二中節染色體(dicentric chromosome)，由 *MLL*-F2-B 偵測到的為無中節染色體(acentric chromosome)，恰為發生相對互調後所產生之互補形式之產物。雖然我們

並未直接偵測到具有相同微同源性(microhomology)序列之相對互調形式的染色體轉位，但上述現象似乎暗示著精蟲細胞的染色體轉位可能是屬於相對互調的形式。



## 肆、討論

### 4.1 哺乳類細胞對 DNA 雙股斷裂之修復機制

哺乳類細胞中造成 DNA 雙股斷裂的因素可分為內源性及外源性兩類。內源性因素包含細胞進行正常生理代謝產生之活性氧自由基(reaction oxygen species, ROS)在 DNA 上造成的阻礙性損害(blocking lesion)，當 DNA 複製進行至此處時，即會使複製叉瓦解(fork collapse)而導致雙股斷裂；此外，核酸酶也可造成 DNA 雙股斷裂，例如細胞分裂時的 Topoisomerase II $\alpha$ ，減數分裂時的 Spo11，V(D)J 重組時所需的 Rag1/Rag2，及 class switch recombination(CSR)所需的 AID 等。外源性因素則包含游離輻射(ionizing radiation) 及一些化療藥物。當細胞的 DNA 雙股斷裂未修復或修復不正確時，即可能造成細胞死亡或染色體大規模的刪除、轉位而使基因體不穩定(genome instability)導致癌症發生<sup>(22)</sup>。

哺乳類細胞對 DNA 雙股斷裂的修復機制主要有同源性重組(homologous recombination, HR)及非同源性末端接合(non-homologous end joining, NHEJ)。HR 需以同源性序列(homologous sequence)搜尋同源性模板，再以此為模板進行合成，產物形式為交叉(crossover, CO)或無交叉(non-crossover, NCO)，為 error-free 之修復機制。NHEJ 不需藉助同源性序列配對亦不需進行合成，僅需以位於 DNA 雙股斷裂末端數個核苷酸序列的微同源性序列(microhomology sequence)配對後再接合，或不需 microhomology 配對，直接進行接合；在修復交界處會有數個核苷酸的插入或刪除，為 error-prone 之修復機制。根據過去的研究，細胞種類、細胞週期及 DNA 雙股斷裂的形式均會影響細胞選擇何種機制來進行修復<sup>(22)</sup>。

進行 NHEJ 需仰賴 4 個核心因子(core factors)：Ku70/Ku80、DNA-PKcs、DNA ligase IV 及 XRCC4。早先以 plasmid rejoicing assay 在缺乏 Ku80 的 *S. cerevisiae*<sup>(23)</sup> 及小鼠細胞<sup>(24)</sup>，發現仍可進行末端接合(end joining)形式的修復，顯示有異於 NHEJ

的機制存在。近來在其他 NHEJ 核心因子缺乏的哺乳類細胞中，證實確實存在此修復機制<sup>(25, 26)</sup>。其特色為在修復交界處普遍存在較長的微同源性(microhomology)序列(extensive microhomology usage)，且造成長度更長的核苷酸插入或刪除，為error-prone 之修復，此機制稱為 microhomology mediated end joining (MMEJ)<sup>(27)</sup>。

#### 4.2 生殖細胞染色體轉位發生機制之探討

依據對偵測到之染色體轉位做序列分析之結果，我們排除其發生機制為同源性重組(HR)的可能性。而在染色體轉位斷裂點交界處均具有微同源性序列(microhomology sequence)，此特性與先前在一些癌細胞觀察到之由 NHEJ 所造成之轉位相似，但由於初級精母細胞(primary spermatocyte)在減數分裂前期 I 期(prophase I)的細線期(leptotene) – 粗線期(pachytene)早期發生減數分裂重組，而此時期的精母細胞並未表現 NHEJ 核心因子 Ku70/Ku80<sup>(28, 29)</sup>，故推測 NHEJ 並非造成生殖細胞染色體轉位之機制。

由於在偵測到的染色體轉位斷裂點交界處有 microhomology usage 之特性，因此懷疑是否有為 MMEJ 所造成之可能性。過去的研究大多是在 NHEJ 核心因子缺乏的基因背景下偵測到 MMEJ 修復之產物<sup>(25, 26, 30, 31)</sup>，並認為 MMEJ 是當其他修復機制均無法進行時之後援路徑(backup pathway)<sup>(32)</sup>。但到目前為止，並不清楚正常細胞透過 MMEJ 修復 DNA 雙股斷裂所佔的比例為何。此外，雖有研究指出某些基因的產物可能參與 MMEJ<sup>(27)</sup>，但並未對這些基因在減數分裂重組時的表現做過探討。我們偵測到的生殖細胞染色體轉位雖然在 microhomology usage 之特性與 MMEJ 相似，但在染色體轉位交界處序列較單純(clear junction sequence)，並未出現類似 MMEJ 產物普遍具有的核苷酸插入及刪除之特性。故我們推測，在生殖細胞偵測到之染色體轉位發生機制可能類似 MMEJ，但其發生雙股斷裂後的處理(process)可能有所不同。

為了與發生在體細胞的 MMEJ 有所區隔，我們將此機制命名為 MMIT(micro-homology-mediated interchromosomal translocation)(圖二十三)。其特色如下：(1)需同時存在兩個 DNA 雙股斷裂。(2)DNA 雙股斷裂產生 3' 黏性末端，以確保目標基因及其轉位對象位於正股(top strand)及負股(bottom strand)之微同源性序列(micro-homology sequence)均被保留。(3)遵循微同源性序列之間的 base pairing rule，為微同源性序列媒介的機制；以微同源性序列配對後，直接進行黏合，不需以轉位對象作為模版進行合成。(4)為非同源性染色體間的轉位。(5)為相對互調(reciprocal translocation)形式之染色體轉位。

#### 4.3 相對互調存在生殖細胞基因轉位的可能性

在 *MLL*-F2，我們分別以前置(5' 端)及後置(3' 端)引子進行偵測，但並未偵測到轉位對象及微同源性(microhomology)序列均相同之染色體轉位。但在 Donor 5，由 *MLL*-F2-F(5' 端)及 *MLL*-F2-B(3' 端)均偵測到 t(11;14)(q23.3;q31.1)(表七 淺灰底標示者)，二者在 *MLL* 之斷裂點相距約 1400 bp，二者的轉位對象相同，均位於 *SEL1L* 及 *FLRT2* 之間，且推斷之染色體轉位產物(putative translocations)，由 *MLL*-F2-F(5' 端)偵測到的為二中節染色體，由 *MLL*-F2-B(3' 端)偵測到的為無中節染色體，恰屬於平衡互換(balanced exchange)之形式。此外，在 Donor 4，由 *MLL*-F2-F 及 *MLL*-F2-B 均有偵測到 t(11;18)(q23.3;q12.2)(表七 深灰底標示者)，二者在 *MLL* 上之斷裂點相距約 500 bp，轉位對象均為 18q12.2，且推斷之染色體轉位產物，由 *MLL*-F2-F 偵測到的為 der(11)，由 *MLL*-F2-B 偵測到的為無中節染色體(acentric chromosome)，雖二者並非屬於互補形式之產物，但此結果是由於轉位對象之微同源性序列一者位於正股(top strand)而另一者位於負股(bottom strand)所導致。

若將目標基因序列調整為正股時，依轉位對象位於染色體長臂或短臂、微同源性序列(microhomology sequence)位於正股或負股及目標基因以 5' 端或 3' 端序列

發生轉位，產物形式可分為八種，以 5' 或 3' PCR 分別可偵測到四種。我們發現偵測到的染色體轉位均可依上述規則分類，幾乎所有形式的產物均有被偵測到，並無偏好出現某些特定形式之產物由 5' 及 3' PCR 均偵測到預期形式的產物，且屬於平衡互換(balanced exchange) (圖二十四)；此外，亦顯示出這些染色體轉位的發生需遵循微同源性序列間的 base-pairing rule，為微同源性序列媒介的機制。

雖然我們並未直接偵測到具相同微同源性序列(microhomology sequence)之相對互調形式的基因轉位，但上述現象似乎暗示著生殖細胞的染色體轉位可能是屬於相對互調的形式。

#### 4.4 生殖細胞的基因轉位與減數分裂 DNA 雙股斷裂修復之關係

人類精子生成(spermatogenesis)的過程中，一個初級精母細胞 (primary spermatocyte)在減數分裂 I 期(meiosis I)時，會由 Spo11<sup>(19)</sup>產生約 152 個 DNA 雙股斷裂(DSBs)<sup>(33)</sup>，藉由同源性重組(HR)，使父源及母源染色體的遺傳訊息交換，對維持人類基因體的多樣性扮演很重要的角色。而生殖細胞進行減數分裂重組最重要目的，是在修復由 Spo11 產生的 DNA 雙股斷裂，避免細胞走向凋亡(apoptosis)。

減數分裂時以同源性重組修復 DNA 雙股斷裂可產生三種形式的產物：(1)無交叉(non-crossovers，NCOs，又稱為 gene conversion)。(2)同源染色體中非姐妹染色分體間的交叉(crossovers，COs)。(3)non-allelic HR(NAHR)所造成的拷貝數變異(copy number variations，CNVs)。

其中，同源染色體中非姐妹染色分體間的交叉(COs)數量會受嚴格調控(CO obligation)，使染色體的每一條臂上至少有一次 CO，以確保染色體平均分配至子細胞中<sup>(34, 35)</sup>。CO 的數量可由 cytological(Mlh1 免疫螢光染色)的方式得知<sup>(36)</sup>，一次減數分裂約產生 50 個 COs<sup>(33)</sup>。相對於交叉(CO)，無交叉(NCO)的偵測受限於多型性標誌(polymorphic markers)的分布，所以其數量可能會被低估<sup>(33)</sup>。人類基因體的

減數分裂重組熱點(recombination hotspots)估計約有 25000-50000 個左右<sup>(37)</sup>，雖然不同熱點間 CO/NCO 之比例會有差異，但在總數上 NCO 仍較 CO 多。NAHR 利用 low copy repeats(LCRs)間的配對，依二個 LCRs 的方向性，可能造成 CNVs，如：重覆(duplication)或刪除(deletion)。目前已知發生在減數分裂時的 NAHR 造成 *de novo* 基因體重排(genomic rearrangement)可能導致子代的遺傳性疾病(genomic disorders)，例如：Charcot-Marie Tooth disease type 1A(CMT1A)，Potocki-Lupski syndrome(PTLS)，Smith-Magenis syndrome(SMS)，HNPP 等<sup>(15, 38-40)</sup>。LCRs 大小大於 1kb，序列相似度大於 95%，佔人類基因體的 5-10%<sup>(41)</sup>，故理論上會有 5-10% 之 DNA 雙股斷裂透過此種方式被修復。

綜合上述可知，減數分裂時，同源染色體中非姐妹染色分體間的 COs 約可修復 30% DNA 雙股斷裂。而 NAHR 理論上可修復 5-10% DNA 雙股斷裂。而經同源性重組修復 DNA 雙股斷裂產生 NCOs 所佔的比例目前仍不清楚。雖然大部分減數分裂的 DNA 雙股斷裂，已可利用上述所提及之機制獲得修復，但尚不能排除有其它機制存在的可能性<sup>(33)</sup>。例如：發生雙股斷裂的染色體本身自行黏合；姐妹染色分體(sister chromatid)間的重組；末端接合(end joining)形式的修復，如：MMEJ (microhomology mediated end joining)，MMIT(microhomology mediated inter-chromosomal translocation)。

由於我們在不同來源的精細胞基因體 DNA 中，普遍均偵測到染色體轉位的存在，而這些染色體轉位在序列特性上，明顯異於上述三者，故我們推測 MMIT 應屬於人類生殖細胞中修復 DNA 雙股斷裂的機制之一。

#### 4.5 實驗結果整合

由 K562 及 HEK293 gDNA spiked 的實驗，得知 RECORD 技術的偵測極限約為  $5 \times 10^{-4}$ 。

利用 RECORD 技術，分別以 *TCF3* 及 *MLL* 當作目標基因，證實在人類生殖細胞中目標基因的重排(rearrangement)有個體間(between individuals)及個體本身(within individual)的差異存在，且個體本身的差異反應出染色體轉位具單一細胞層次的不同(cellular level difference)。而目標基因傾向以隨機的(stochastic)方式，與其它對象發生重排。在二者交界處 microhomology usage 的特性，與在體細胞由 NHEJ 及 MMEJ(microhomology mediated end-joining)所造成的基因轉位類似。由於 NHEJ 核心因子 Ku70/Ku80 在減數分裂時並未表現，故並非 NHEJ 所造成。故推測這些染色體轉位可能由類似 MMEJ 的機制所產生。為了與在體細胞發生的 MMEJ 有所區隔，我們將此機制命名為 MMIT(microhomology mediated interchromosomal translocation)。

生殖細胞進行減數分裂重組的重要目的，即是在修復由 Spo11 所產生的 DNA 雙股斷裂。雖然同源性重組(HR)是目前熟知的修復機制，但尚不能排除有其他機制參與其中。由於我們在不同來源的精細胞基因體 DNA 中，普遍均偵測到染色體轉位的存在，而這些染色體轉位在序列特性上，明顯異於由同源性重組修復者，故我們推測 MMIT 應屬於人類生殖細胞中修復 DNA 雙股斷裂的機制之一。

## 参考文献

1. The International HapMap Consortium (2007). "A second generation human haplotype map of over 3.1 million SNPs." *Nature* 449(7164): 851-861.
2. Sebat, J., B. Lakshmi, J. Troge, et al. (2004). "Large-Scale Copy Number Polymorphism in the Human Genome." *Science* 305(5683): 525-528.
3. Iafrate, A.J., L. Feuk, M.N. Rivera, et al. (2004). "Detection of large-scale variation in the human genome." *Nat Genet* 36(9): 949-951.
4. Tuzun, E., A.J. Sharp, J.A. Bailey, et al. (2005). "Fine-scale structural variation of the human genome." *Nat Genet* 37(7): 727-732.
5. Feuk, L., A.R. Carson, and S.W. Scherer (2006). "Structural variation in the human genome." *Nat Rev Genet* 7(2): 85-97.
6. Khaja, R., J. Zhang, J.R. MacDonald, et al. (2006). "Genome assembly comparison identifies structural variants in the human genome." *Nat Genet* 38(12): 1413-1418.
7. Korbel, J.O., A.E. Urban, J.P. Affourtit, et al. (2007). "Paired-End Mapping Reveals Extensive Structural Variation in the Human Genome." *Science* 318(5849): 420-426.
8. Hurles, M.E., E.T. Dermitzakis, and C. Tyler-Smith (2008). "The functional impact of structural variation in humans." *Trends in Genetics* 24(5): 238-245.
9. Nambiar, M., V. Kari, and S.C. Raghavan (2008). "Chromosomal translocations in cancer." *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer* 1786(2): 139-152.
10. Mitelman, F., B. Johansson, and F. Mertens (2007). "The impact of translocations and gene fusions on cancer causation." *Nat Rev Cancer* 7(4): 233-245.

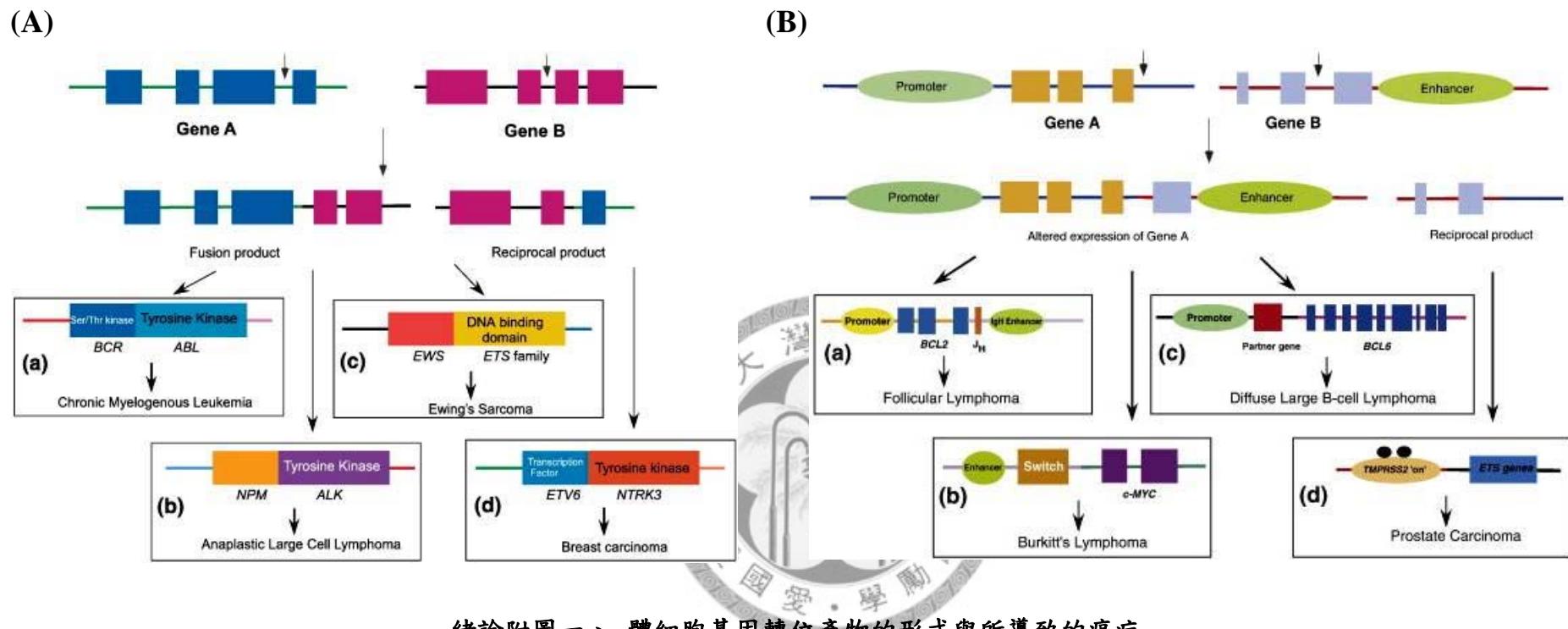
11. Lazorchak, A., M.E. Jones, and Y. Zhuang (2005). "New insights into E-protein function in lymphocyte development." *Trends in Immunology* 26(6): 334-338.
12. Wiemels, J.L., B.C. Leonard, Y. Wang, et al. (2002). "Site-specific translocation and evidence of postnatal origin of the t(1;19) E2A-PBX1 fusion in childhood acute lymphoblastic leukemia." *PNAS* 99(23): 15101-15106.
13. Inaba, T., W. Roberts, L. Shapiro, et al. (1992). "Fusion of the leucine zipper gene HLF to the E2A gene in human acute B-lineage leukemia." *Science* 257(5069): 531-534.
14. Brambillasca, F., G. Mosna, M. Colombo, et al. (1999). "Identification of a novel molecular partner of the E2A gene in childhood leukemia." *Leukemia*. 13(3): 369-375.
15. Lupski, J.R. (2007). "Genomic rearrangements and sporadic disease." *Nat Genet* 39: 543-547.
16. Kurahashi, H. and B.S. Emanuel (2001). "Unexpectedly high rate of de novo constitutional t(11;22) translocations in sperm from normal males." *Nat Genet* 29(2): 139-140.
17. Gribble, S.M., I. Roberts, C. Grace, et al. (2000). "Cytogenetics of the Chronic Myeloid Leukemia-Derived Cell Line K562: Karyotype Clarification by Multicolor Fluorescence In Situ Hybridization, Comparative Genomic Hybridization, and Locus-Specific Fluorescence In Situ Hybridization." *Cancer Genetics and Cytogenetics* 118(1): 1-8.
18. Chissoe, S.L., A. Bodenteich, Y.-F. Wang, et al. (1995). "Sequence and Analysis of the Human ABL Gene, the BCR Gene, and Regions Involved in the Philadelphia Chromosomal Translocation." *Genomics* 27(1): 67-82.
19. Keeney, S., C.N. Giroux, and N. Kleckner (1997). "Meiosis-Specific DNA

- Double-Strand Breaks Are Catalyzed by Spo11, a Member of a Widely Conserved Protein Family." *Cell* 88(3): 375-384.
- 20. Neale, M.J., J. Pan, and S. Keeney (2005). "Endonucleolytic processing of covalent protein-linked DNA double-strand breaks." *Nature* 436(7053): 1053-1057.
  - 21. International Human Genome Sequencing Consortium (2001). "Initial sequencing and analysis of the human genome." *Nature* 409(6822): 860-921.
  - 22. Shrivastav, M., L.P. De Haro, and J.A. Nickoloff "Regulation of DNA double-strand break repair pathway choice." *Cell Res* 18(1): 134-147.
  - 23. Boulton, S. and S. Jackson (1996). "Identification of a *Saccharomyces cerevisiae* Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance." *Nucl. Acids Res.* 24(23): 4639-4648.
  - 24. Liang, F., P.J. Romanienko, D.T. Weaver, et al. (1996). "Chromosomal double-strand break repair in Ku80-deficient cells." *PNAS* 93: 8929-8933.
  - 25. Yan, C.T., C. Boboila, E.K. Souza, et al. (2007). "IgH class switching and translocations use a robust non-classical end-joining pathway." *Nature* 449(7161): 478-482.
  - 26. Corneo, B., R.L. Wendland, L. Deriano, et al. (2007). "Rag mutations reveal robust alternative end joining." *Nature* 449(7161): 483-486.
  - 27. McVey, M. and S.E. Lee (2008). "MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings." *Trends in Genetics* 24(11): 529-538.
  - 28. Hamer, G., H.B. Kal, C.H. Westphal, et al. (2004). "Ataxia Telangiectasia Mutated Expression and Activation in the Testis." *Biol Reprod* 70(4): 1206-1212.

29. Goedecke, W., M. Eijpe, H.H. Offenberg, et al. (1999). "Mre11 and Ku70 interact in somatic cells, but are differentially expressed in early meiosis." *Nat Genet* 23(2): 194-198.
30. Bogue, M.A., C. Wang, C. Zhu, and D.B. Roth (1997). "V(D)J Recombination in Ku86-Deficient Mice: Distinct Effects on Coding, Signal, and Hybrid Joint Formation." *Immunity* 7(1): 37-47.
31. Jankovic, M., A. Nussenzweig, and M.C. Nussenzweig (2007). "Antigen receptor diversification and chromosome translocations." *Nat Immunol* 8(8): 801-808.
32. Nussenzweig, A. and M.C. Nussenzweig (2007). "A Backup DNA Repair Pathway Moves to the Forefront." *Cell* 131(2): 223-225.
33. Baudat, F. and B. de Massy (2007). "Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis." *Chromosome Research* 15(5): 565-577.
34. Jones, G.H. and F.C.H. Franklin (2006). "Meiotic Crossing-over: Obligation and Interference." *Cell* 126(2): 246-248.
35. Martini, E., R.L. Diaz, N. Hunter, and S. Keeney (2006). "Crossover Homeostasis in Yeast Meiosis." *Cell* 126(2): 285-295.
36. Baker, S.M., W. PlugAnnemieke, T.A. Prolla, et al. (1996). "Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over." *Nat Genet* 13(3): 336-342.
37. Myers, S., L. Bottolo, C. Freeman, et al. (2005). "A Fine-Scale Map of Recombination Rates and Hotspots Across the Human Genome." *Science* 310(5746): 321-324.
38. Lupski, J.R. and P. Stankiewicz (2005). "Genomic Disorders: Molecular

- Mechanisms for Rearrangements and Conveyed Phenotypes." *PLoS Genet* 1(6): e49.
39. Turner, D.J., M. Miretti, D. Rajan, et al. (2008). "Germline rates of de novo meiotic deletions and duplications causing several genomic disorders." *Nat Genet* 40(1): 90-95.
40. Gu, W., F. Zhang, and J. Lupski (2008). "Mechanisms for human genomic rearrangements." *PathoGenetics* 1(1): 4.
41. Jeffrey, A.B., M.Y. Amy, F.M. Hillary, et al. (2001). "Segmental Duplications: Organization and Impact Within the Current Human Genome Project Assembly." *Genome Res.* 11: 1005-1017.

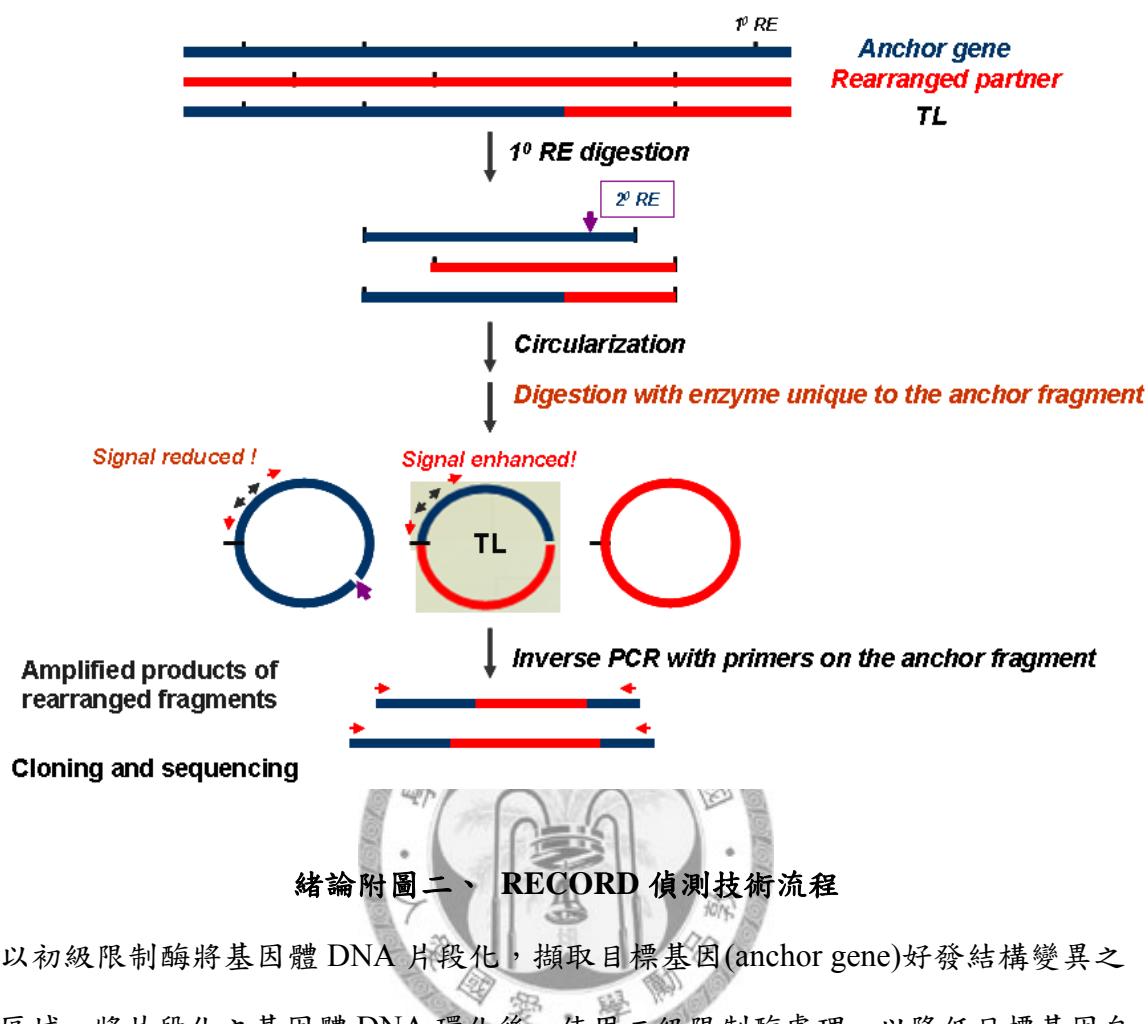




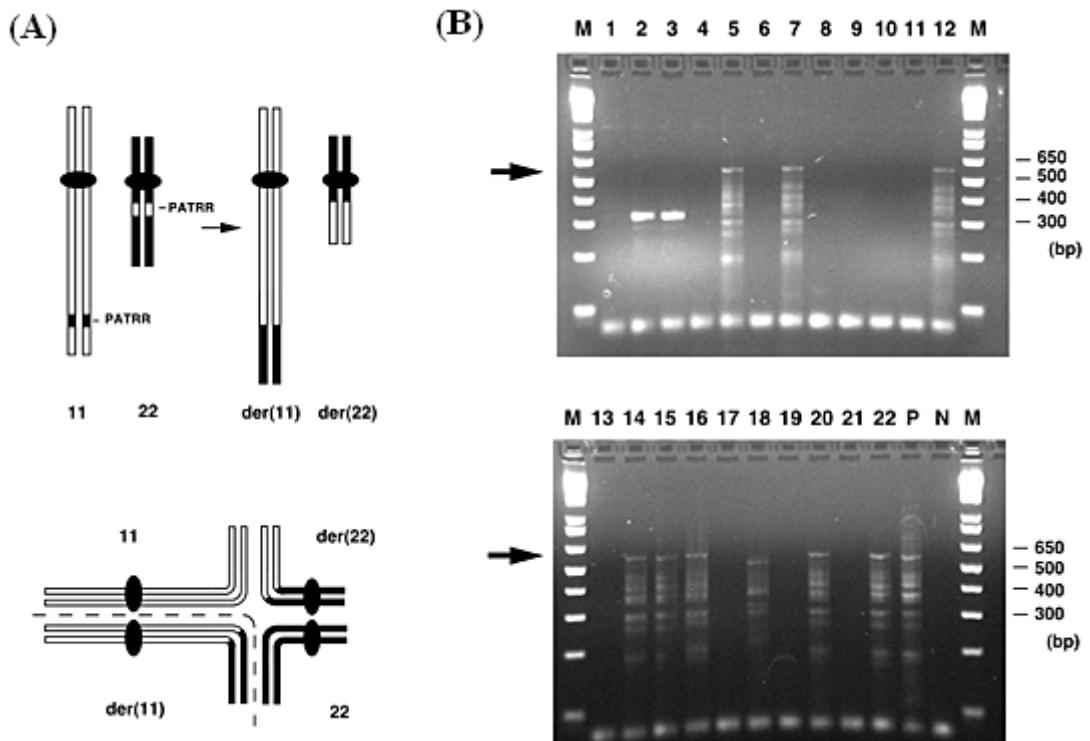
緒論附圖一、體細胞基因轉位產物的形式與所導致的癌症

(A)基因轉位產生致癌性的融合蛋白質產物。(B)基因轉位將致癌基因帶到其他基因的啟動子(promoter)或增強子(enhancer)附近，影響該基因的表現。

(選自參考文獻 9)



以初級限制酶將基因體DNA片段化，擷取目標基因(anchor gene)好發結構變異之區域。將片段化之基因體DNA環化後，使用二級限制酶處理，以降低目標基因自體環化片段(unrearranged form)被擴製的機率，並提升目標基因發生結構變異模板(rearranged form, TL)之檢出率，再使用設計於目標基因上的引子進行反向聚合酶連鎖反應(inverse PCR)，再將產物進行選殖與定序分析，即可得知與目標基因發生結構變異的對象為何。



緒論附圖三、偵測源自同一個體的  $t(11;22)(q23;q11)$  轉位產物 der(11)之電泳結果

(A)在人類精蟲細胞中，由 11q23 及 22q11 的 palindromic AT-rich region(PATRR)所造成的復發性基因轉位  $t(11;22)(q23;q11)$ 。(B)箭頭標示處為經定序證實的 der(11)片段。電泳結果顯示，相同檢體的不同 aliquots 之間，未必均存在 der(11)片段，且存在的 der(11)片段大小仍有些微差異。lane 2&3 為 Alu repeat 造成的 PCR artifact；呈現階梯狀(ladder)的部份，為放大 AT-rich 片段所造成的 PCR artifact。

(選自參考文獻 16)

	Primers	Sequence	Unrearranged product	Rearranged product	Annealing temp.	Elongation time
<b>Inverse PCR</b>	<i>BCR-i-1F</i>	TCCAGCCACTCATTCTCAGACC	4558 bp	7086 bp	65°C ◇ 61°C	7 min
	<i>BCR-i-1R</i>	GGAAGTGCCCAGAGAAGGAAGG				
<b>Nested PCR</b>	<i>BCR-i-2F</i>	ATGCTCTGTGCCTTGGATCTGG	3297 bp	5825 bp	65°C ◇ 61°C	6 min
	<i>BCR-i-2R</i>	TCAGGACAGAGGATACTCACACTTC				
<b>TL PCR</b>	<i>BCR-i-2F</i>	ATGCTCTGTGCCTTGGATCTGG	-	974 bp	56°C	1 min
	<i>BCR/ABL-R</i>	ACTATCAAGGGACAAGTACTGAGTG				
<b>Control</b>	<i>BCR-i-1F</i>	TCCAGCCACTCATTCTCAGACC	2996 bp	2996 bp	65°C ◇ 61°C	3 min
	<i>BCR-i-C</i>	TGTACACGCCATTGTTACACAGTG				

表一、評估 RECORD 偵測敏感度之引子及 PCR 反應條件

Annealing temp.: (X°C ◇ Y°C) 。Elongation time: Z min。

Fragment	Primers		Sequence	Unrearranged product	Annealing temp.	Elongation time
<b>TCF3-B (Backward) (5')</b>	<b>Inverse PCR</b>	<i>TCF3-B-1F</i>	ACTGTGCCCATCCGAAGTGTCC	4537 bp	65°C ◇ 61°C	5 min
		<i>TCF3-B-2R</i>	GCACTCTGGTGTGACAGGGTGAGAC			
	<b>Nested PCR</b>	<i>TCF3-B-3F</i>	TGTGGTCCTTAATCTGTCTGGCATTGG	4215 bp	67°C ◇ 63°C	4 min
		<i>TCF3-B-4R</i>	ACACCCACCACAGATGTAGCCAACC			
	<b>LS PCR</b>	<i>TCF3-B-lig.</i>	ACACATGGCGTGAGGTAGGGATCT	4209 bp	59°C ◇ 55°C	4 min
		<i>TCF3-B-4R</i>	ACACCCACCACAGATGTAGCCAACC			
	<b>Inverse PCR</b>	<i>TCF3-F-1F</i>	CGTCCAGGTTAGGTACAGCACAGG	5066 bp	62°C ◇ 58°C	5 min
		<i>TCF3-F-2R</i>	GCCCTCCCTCACTGCCTGGTTATG			
<b>TCF3-F (Forward) (3')</b>	<b>Nested PCR</b>	<i>TCF3-F-3F</i>	ATGTGTGCGTGGCCATTGCTC	4795 bp	65°C ◇ 61°C	5 min
		<i>TCF3-F-4R</i>	CGGTTCCCTGACCAGCGATGCC			
	<b>LS PCR</b>	<i>TCF3-F-3F</i>	ATGTGTGCGTGGCCATTGCTC	4525 bp	60°C ◇ 56°C	4.5 min
		<i>TCF3-F-lig.</i>	TGAGTGATCCGGGGAGTAGATCC			
<b>TCF3-Control</b>	<b>Control</b>	<i>TCF3-F-3F</i>	ATGTGTGCGTGGCCATTGCTC	3601 bp	59°C ◇ 55°C	4 min
		<i>TCF3-B-4R</i>	ACACCCACCACAGATGTAGCCAACC			

表二、用於偵測人類精蟲細胞 *TCF3* 結構變異之引子及 PCR 反應條件

Annealing temp.: (X°C ◇ Y°C) 。Elongation time: Z min。

TCF3	Breakpoint locus		Breakpoint location (hg17)		Microhomology sequence	Putative translocation	TL chromosome	Sample
	TCF3	TPG/locus	TCF3	TPG/locus				
<b>Backward (5')</b>	Intron 15	<i>TMT2C</i> (Intron 4)	1567934	chr12: 81853824	<u>AGAATCAATG</u>	t(12;19)(q21.31;p13.3)	Acentric	B
	Intron 15	<i>RB1CC1</i> (Intron 1)	1567971	chr 8: 53778555	AAATA	t(8;19)(q11.23;p13.3)	der(8)	C
	Intron 15	<i>HIPK3-MGC34830</i>	1567989	chr11: 33471443	<u>AGAAG</u>	t(11;19)(p13;p13.3)	Acentric	A
	Intron 15	<i>TRIB2-NSE1</i>	1568009	chr 2: 14413880	<u>CAAGGAAA</u>	t(2;19)(p24.3;p13.3)	der(2)	A
	Intron 15	<i>CPT1A-MRPL21</i>	1568123	chr11: 68396167	TTGAG	t(11;19)(q13.3;p13.3)	Acentric	5
	Intron 15	<i>RNF19-ANKRD46</i>	1568167	chr 8: 101564458	<u>CTTGTGC</u>	t(8;19)(q22.2;p13.3)	Acentric	C
	Intron 15	<i>PREP-PRDM1</i>	1568186	chr 6: 106335314	CA	t(6;19)(q21;p13.3)	Acentric	E
	Intron 15	<i>WAC-BAMBI</i>	1568266	chr10: 28998582	<u>TGTAGGAATGA</u>	t(10;19)(p12.1;p13.3)	Acentric	2
	Intron 15	<i>TIGD2-GPRIN3</i>	1568387	chr 4: 90516874	ACACGC	t(4;19)(q22.1;p13.3)	der(4)	C
	Intron 15	<i>SLC30A1-NEK2</i>	1568504	chr 1: 208182310	<u>GGCCTGGGCCACC</u>	t(1;19)(q32.3;p13.3)	Acentric	B
	Intron 15	<i>PLCL1-SATB2</i>	1568826	chr 2: 199343922	CTGAGCA	t(2;19)(q33.1;p13.3)	Acentric	B
<b>Forward (3')</b>	Intron 13	<i>CDCA1-PBX1</i>	1570755	chr 1: 160653590	<u>GGTGG</u>	t(1;19)(q23.3;p13.3)	der(19)	B
	Intron 12	<i>FLJ10901-CACNA1S</i>	1571247	chr 1: 197731372	GAG	t(1;19)(q32.1;p13.3)	Dicentric	B
	Intron 12	<i>CBARA1</i> (Intron 3)	1571248	chr10: 73989787	AGG	t(10;19)(q22.1;p13.3)	Dicentric	A
	Intron 12	<i>DCC</i> (Intron 2)	1571307	chr18: 48546216	<u>ACACCCA</u>	t(18;19)(q21.1;p13.3)	Dicentric	A
	Intron 12	<i>MOBP-MYRIP</i>	1571404	chr 3: 39598820	AGGAGGG	t(3;19)(p22.2;p13.3)	der(19)	D
	Intron 12	<i>MGC9564-SLC13A2</i>	1571425	chr17: 23796351	<u>GGATC</u>	t(17;19)(q11.2;p13.3)	der(19)	B
	Intron 12	<i>C3orf23-ZNF445</i>	1571440	chr 3: 44513341	<u>GGTGAA</u>	t(3;19)(p21.31;p13.3)	der(19)	3
	Intron 12	<i>HMP19-MSX2</i>	1571594	chr 5: 174050309	C	t(5;19)(q35.2;p13.3)	Dicentric	5

表三、由人類精蟲細胞基因體DNA偵測到的TCF3染色體轉位總表

在微同源性(microhomology)序列部分，有底線標示之核苷酸為偵測到的染色體轉位序列之 microhomology 在 TCF3 與其轉位對象相異之處；粗底線者為源自 TCF3 之序列，細底線標示者為源自轉位對象之序列，虛線底線標示者為刪除(deletion)之核苷酸序列。

Fragment	Breakpoint locus		Breakpoint location (hg17)		Microhomology sequence	Putative translocation	TL chromosome	Sample
	MLL	TPG/locus	MLL	TPG/locus				
<b>MLL-F2</b>	Intron 9	<i>NOTCH2NL</i> (Intron 4)	117860618	chr1: 142770835	TAG	t(1;11)(q21.1;q23.3)	der(11)	B
	Intron 9	<i>PDE4IP-SEC22B2</i>	117860645	chr1: 142575076	ACTTTTAA	t(1;11)(q21.1;q23.3)	der(11)	C
	Intron 9	<i>C8orf22-SNTG1</i>	117860663	chr8: 50692908	TGG	t(8;11)(q11.21;q23.3)	Dicentric	B
	Intron 9	<i>PIK3R4</i> (Exon 4)	117860770	chr3: 131946707	TGAC	t(3;11)(q21.3;q23.3)	Dicentric	B
	Exon 10	<i>CHD8-CHD11</i>	117860886	chr16: 61860379	TCTGTGCCAGTA	t(11;16)(q23.3;q21)	Dicentric	D
	Intron10	<i>RAD23B-ZNF462</i>	117860919	chr9: 106877015	CCTGCTTCTT	t(9;11)(q31.22;q23.3)	Dicentric	E

表四、由人類精蟲細胞基因體 DNA 在 MLL-F2 偵測到的染色體轉位總表

(實驗結果來自支若馨學姊)

灰底標示者為由 Sample B 偵測到的染色體轉位，確認試驗針對此部分進行。



Fragment		Primers	Sequence	Unrearranged product	Annealing temp	Elongation time
<b>MLL-F2-F (5')</b>	Inverse PCR	MLL-F2-1F	AACATTCCATCCATCCTGAGC	4537 bp	59°C ◇ 55°C	4.5 min
		MLL-F2-2R	ATTCCTCTTCAACACACCTTGG			
	Nested PCR	MLL-F2-2F	CCATGATGATTCCCTGAGTCAG	4384 bp	60°C ◇ 56°C	4 min
		MLL-F2-4R	TTGCTTAGAACTATTGCCATTGG			
	LS PCR	MLL-F2-3F	TATGAATTGAACAACTAGGTGAGC	4185 bp	60°C ◇ 56°C	4 min
		MLL-F2-Lig.	TGCAAATGACAAATTTATTAATAGA			
	Inverse PCR	MLL-F2-B-1F	TGGAGTGTATAAGTGCGAACAG	3566 bp	62°C ◇ 58°C	4 min
		MLL-F2-B-6R	ACCCTGTGAGATTGCTATGTACAAC			
	Nested PCR	MLL-F2-B-3F	TGCCTCATTACTAGGAAATCATCTCAG	2533 bp	60°C ◇ 56°C	3 min
		MLL-F2-B-8R	TGTGGTAAGTAATAAGCCAAAGAGG			
	LS PCR	MLL-F2-B-Lig.	CTATAAATGGATGCATTAAAGA	2500 bp	58°C ◇ 54°C	2.5 min
		MLL-F2-B-10R	AACTGTAAAGGATGTAGGATATG			
<b>MLL-Control</b>	Control	MLL-F2-1F	AACATTCCATCCATCCTGAGC	2814 bp	60°C ◇ 56°C	3 min
		MLL-F2-B-6R	ACCCTGTGAGATTGCTATGTACAAC			

表五、用於偵測人類精蟲細胞 MLL-F2 結構變異之引子及 PCR 反應條件

Annealing temp.: (X°C ◇ Y°C) 。 Elogation time: Z min 。

Translocation	Primers	Sequence	Product	Annealing temp	Elongation time
t(1;11)(q21.1;q23.3) <i>(NOTCH2NL)</i>	<i>MLL</i> -F2-3F	TATGAATTGAACAACTAGGTGAGC	347 bp	53°C	30 sec
	<i>NOTCH2NL</i> -V	ATCACCTACTGTTAATGTC			
t(8;11)(q11.21;q23.3)	<i>MLL</i> -F2-3F	TATGAATTGAACAACTAGGTGAGC	212 bp	60°C	30 sec
	8q11.21-V	AGATGTTGAACCAGGGAATG			
t(3;11)(q21.3;q23.3) <i>(PIK3R4)</i>	<i>MLL</i> -F2-3F	TATGAATTGAACAACTAGGTGAGC	272 bp	60°C	30 sec
	<i>PIK3R4</i> -V	ACCTTCACAAACGACCAG			

表六、 *MLL* 染色體轉位特異性 PCR 反應之引子及 PCR 反應條件



MLL-F2	Breakpoint locus		Breakpoint location (hg17)		Microhomology sequence	Putative translocation	TL chromosome	Sample
	MLL	TPG/locus	MLL	TPG/locus				
<b>F (5')</b>	Intron 9	<i>KIAA1217-ARHGAP21</i>	117860617	chr10:24902887	TAATA	t(10;11)(p12.1;q23.3)	der(11)	C
	Intron 9	<i>ZNF621-CTNNB1</i>	117860650	chr 3:40772009	AG	t(3;11)(p22.1;q23.3)	der(11)	D
	Intron 9	<i>TACSTD2-JUN</i>	117860764	chr 1:58770610	CCTA	t(1;11)(p32.1;q23.3)	der(11)	E
	Exon 10	<i>CBWD1</i> (Intron 4)	117860788	chr 9:159331	CAGGA	t(9;11)(p24.3;q23.3)	der(11)	C
	Exon 10	<i>MGC45800-DCTD</i>	117860846	chr 4:183903028	CTTTT	t(4;11)(q35.1;q23.3)	Dicentric	A
	Exon 10	<i>SCYE1-DKK2</i>	117860863	chr 4:108051238	GGT	t(4;11)(q25;q23.3)	der(11)	C
	Exon 10	<i>CLSTN2</i> (Intron 5)	117860881	chr 3:141634177	CTGT <u>GC</u>	t(3;11)(q23;q23.3)	Dicentric	D
	Intron 10	<i>GHR</i> (Intron 1)	117860907	chr 5:42528402	TAAGGC	t(5;11)(p12;q23.3)	Dicentric	D
	Intron 10	<i>BRUNOL4</i> (Intron 1)	117860980	chr18:33353510	TTTTGT	t(11;18)(q23.3;q12.2)	der(11)	4
	Intron 10	<i>SEL1L-FLRT2</i>	117860992	chr14:81086587	TT <u>AG</u> CA	t(11;14)(q23.3;q31.1)	Dicentric	5
	Intron 10	<i>PIK3CB</i> (Intron 12)	117861252	chr 3:139904381	ATAA	t(3;11)(q22.3;q23.3)	der(11)	5
	Intron 10	<i>PTPRE</i> (Intron 10)	117861365	chr10:129753807	<u>T</u> AGCAGGG	t(10;11)(q26.2;q23.3)	Dicentric	B
	Intron 10	<i>SEC22L1-NOTCH2NL</i>	117861545	chr 1:142618300	<u>T</u> ATTAAAAT	t(1;11)(q21.1;q23.3)	der(11)	B
	Intron 10	<i>FLJ38723-TLN2</i>	117861793	chr15:60390851	<u>G</u> CAAACAAGG	t(11;15)(q23.3;q22.2)	Dicentric	D
<b>B (3')</b>	Intron 10	<i>BRUNOL4-PIK3C3</i>	117861446	chr18:34331338	T <u>CTT</u>	t(11;18)(q23.3;q12.2)	Acentric	4
	Intron 10	<i>MYOZ2</i> (Intron 4)	117861689	chr 4:120441554	CT	t(4;11)(q26;q23.3)	der(4)	5
	Intron 10	<i>GABRA2</i> (Intron 8)	117861926	chr 4:46124732	<u>GG</u> CAT	t(4;11)(p12;q23.3)	Acentric	5
	Intron 10	<i>TRIB2-NSE1</i>	117862155	chr 2:13765524	GCAGG	t(2;11)(p24.3;q23.3)	Acentric	C
	Intron 10	<i>SEL1L-FLRT2</i>	117862406	chr14:83480857	CATT <u>TA</u>	t(11;14)(q23.3;q31.1)	Acentric	5
	Intron 10	<i>GOT2-CDH8</i>	117862413	chr16:59161617	GA	t(11;16)(q23.3;q21)	der(16)	5

表七、確認試驗時由 MLL-F2 偵測到的人類精蟲細胞染色體轉位總表

在 microhomology sequence 部分，有底線標示之核苷酸為偵測到的染色體轉位序列之 microhomology 在 MLL 與其轉位對象相異之處；粗底線者為源自 MLL 之序列，細底線標示者為源自轉位對象之序列。

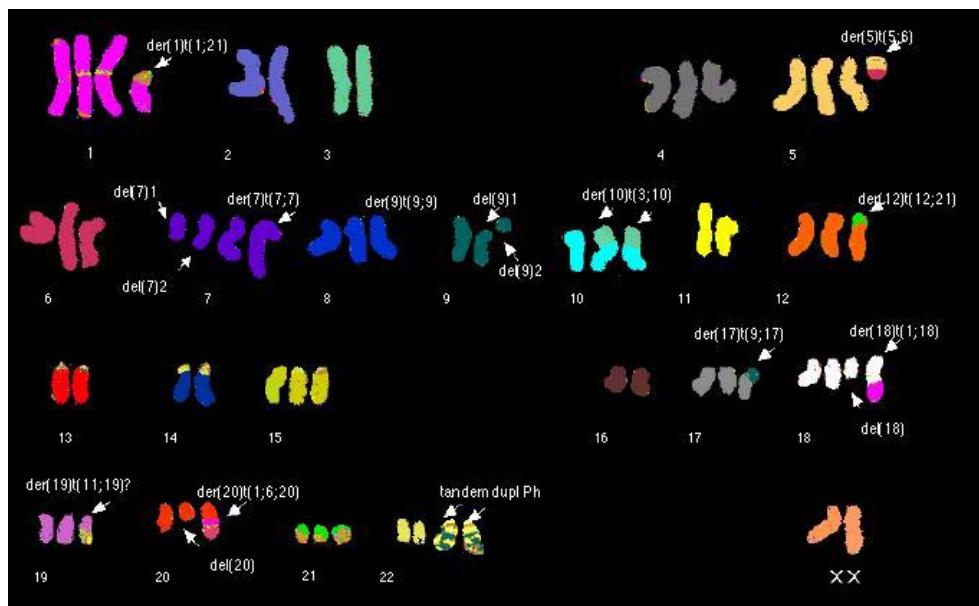
Sample	TCF3	Translocation partner distribution		MLL-F2	Translocation partner distribution	
		chromosome p arm	chromosome q arm		chromosome p arm	chromosome q arm
A	5'	2p24.3, 11p13	-	5'	-	4q35.1, 5q11.2
	3'	-	10q22.1, 18q21.1	3'	-	-
B	5'	-	1q32.3, 2q33.1, 12q21.31	5'	-	1q21.1, 1q21.1, 3q21.3, 8q11.21, 10q26.2
	3'	-	1q23.3, 1q32.1, 17q11.2	3'	-	-
C	5'	-	4q22.1, 8q11.23, 8q22.2	5'	9p24.3, 10p12.1	1q21.1, 4q25
	3'	-	-	3'	2p24.3	-
D	5'	-	-	5'	3p22.1, 5p12	3q23, 15q22.2, 16q21
	3'	3p22.2	-	3'	-	-
E	5'	-	6q21	5'	1p32.1	9q31.22
	3'	-	-	3'	-	-
d1	5'	-	-	5'	-	-
	3'	-	-	3'	-	-
d2	5'	10p12.1	-	5'	-	-
	3'	-	-	3'	-	-
d3	5'	-	-	5'	-	-
	3'	3p21.31	-	3'	-	-
d4	5'	-	-	5'	-	18q12.2
	3'	-	-	3'	-	18q12.2
d5	5'	-	11q13.3	5'	-	3q22.3, 14q31.1
	3'	-	5q35.2	3'	4p12	4q26, 14q31.1, 16q21
d6	5'	-	-	5'	-	-
	3'	-	-	3'	-	-
d7	5'	-	-	5'	-	-
	3'	-	-	3'	-	-
Total		5	14	Total	7	20
p/q ratio		0.36		p/q ratio		0.35

表八、以 TCF3 及 MLL 為目標基因所偵測出之人類精蟲細胞染色體轉位其轉位對象在染色體長短臂上的分布

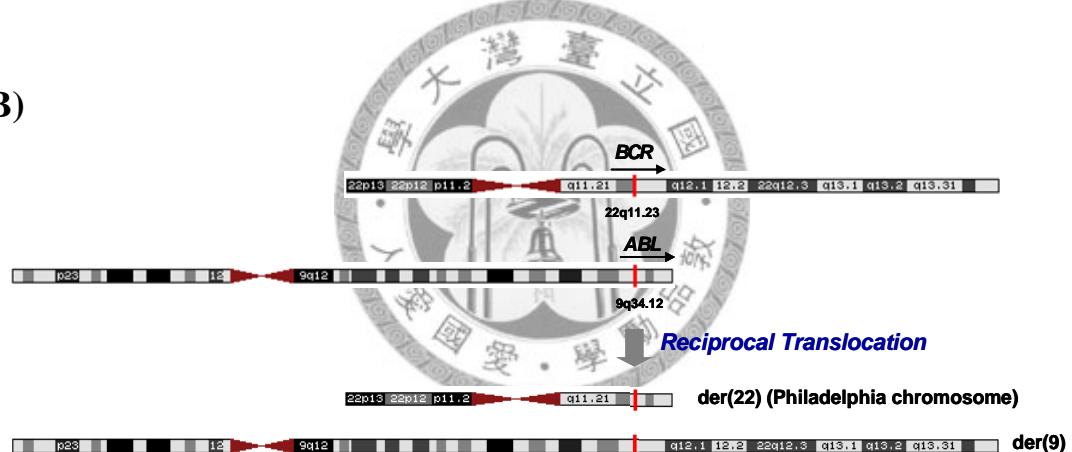
MLL-F2-F (11q23.3)	Breakpoint locus			Microhomology sequence	Putative translocation	TL chromosome	Sperm gDNA
	MLL	Partner band	Partner gene				
Discovery	Intron 9	1q21.1	<i>NOTCH2NL</i> (Intron 4)	TAG	t(1;11)(q21.1;q23.3)	der(11)	B
	Intron 9	1q21.1	-	ACTTTTAA	t(1;11)(q21.1;q23.3)	der(11)	C
	Intron 9	8q11.21	-	TGC	t(8;11)(q11.21;q23.3)	Dicentric	B
	Intron 9	3q21.3	<i>PIK3R4</i> (Exon 2)	TGAC	t(3;11)(q21.3;q23.3)	Dicentric	B
	Exon 10	16q21	-	TCTGTGCCAGTA	t(11;16)(q23.3;q21)	Dicentric	D
	Intron 10	9q31.22	-	CCTGCTCTT	t(9;11)(q31.22;q23.3)	Dicentric	E
	Intron 10	5q11.2	-	TAGCAA	t(5;11)(q11.2;q23.3)	Dicentric	A
Validation	Intron 9	10p12.1	-	TAATA	t(10;11)(p12.1;q23.3)	der(11)	C
	Intron 9	3p22.1	-	AG	t(3;11)(p22.1;q23.3)	der(11)	D
	Intron 9	1p32.1	-	CCTA	t(1;11)(p32.1;q23.3)	der(11)	E
	Exon 10	9p24.3	<i>CBWD1</i> (Intron 4)	CAGGA	t(9;11)(p24.3;q23.3)	der(11)	C
	Exon 10	4q35.1	-	CTTTT	t(4;11)(q35.1;q23.3)	Dicentric	A
	Exon 10	4q25	-	GGT	t(4;11)(q25;q23.3)	der(11)	C
	Exon 10	3q23	<i>CLSTN2</i> (Intron 5)	CTGT <u>GC</u>	t(3;11)(q23;q23.3)	Dicentric	D
	Intron 10	5p12	<i>GHR</i> (Intron 1)	TAAGGC	t(5;11)(p12;q23.3)	Dicentric	D
	Intron 10	18q12.2	<i>BRUNOL4</i> (Intron 1)	TTTTGT	t(11;18)(q23.3;q12.2)	der(11)	4
	Intron 10	14q31.1	-	TT <u>AGCA</u>	t(11;14)(q23.3;q31.1)	Dicentric	5
	Intron 10	3q22.3	<i>PIK3CB</i> (Intron 12)	ATAA	t(3;11)(q22.3;q23.3)	der(11)	5
	Intron 10	10q26.2	<i>PTPRE</i> (Intron 10)	T <u>AGCAGGG</u>	t(10;11)(q26.2;q23.3)	Dicentric	B
	Intron 10	1q21.1	-	T <u>GTTAAAAT</u>	t(1;11)(q21.1;q23.3)	der(11)	B
	Intron 10	15q22.2	-	<u>GCAAACAA</u> AGG	t(11;15)(q23.3;q22.2)	Dicentric	D

表九、由人類精蟲細胞偵測出之 MLL-F2-F 染色體轉位總表

(A)



(B)

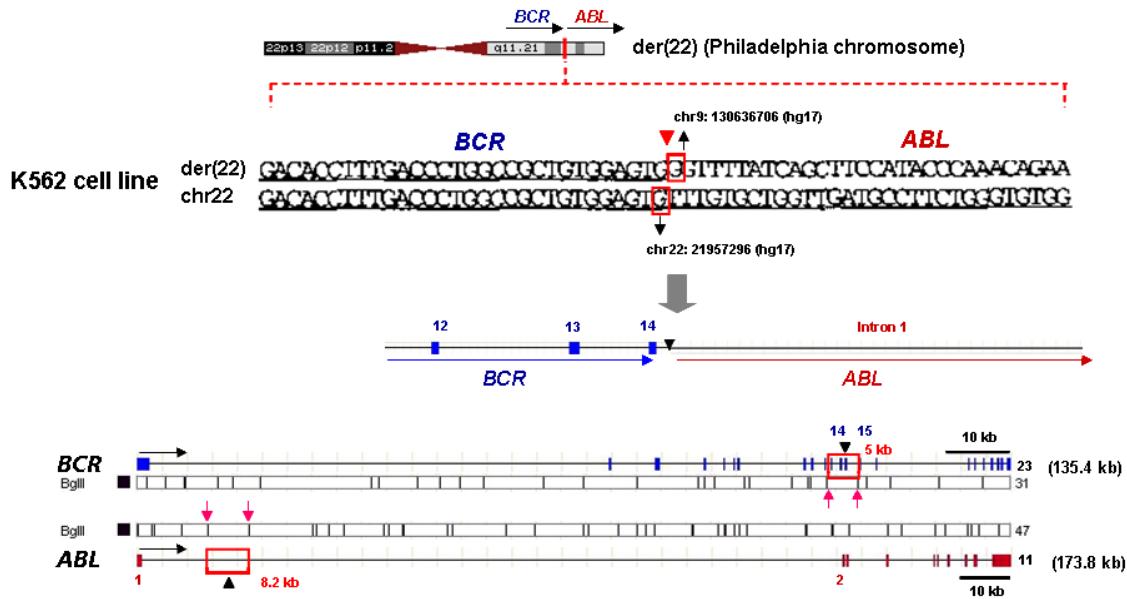


圖一、K562細胞株之染色體核型分析

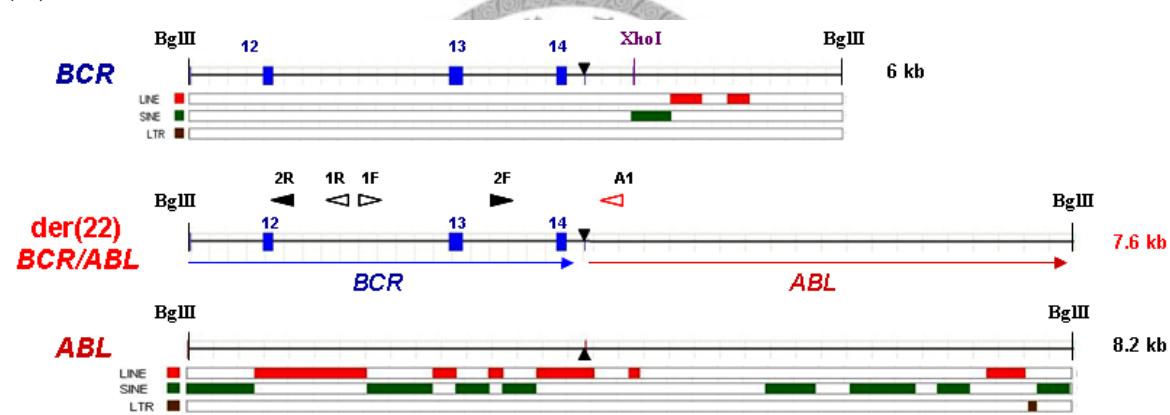
(A)K562細胞株之染色體核型分析。(B)K562細胞株第九及第二十二對染色體發生相互對調產生的衍生染色體示意圖。

(圖一之A選自參考文獻 17)

(A)



(B)

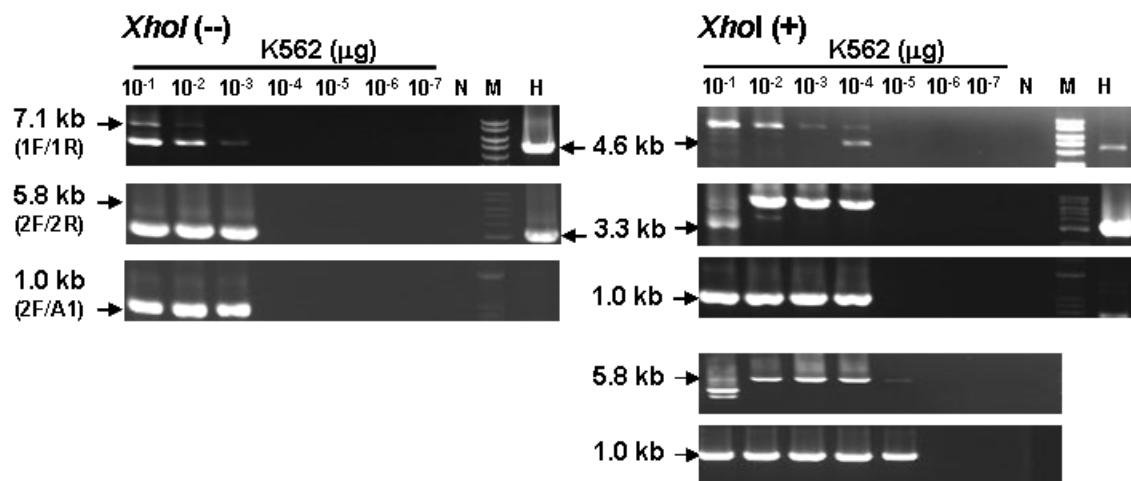


圖二、評估 RECORD 技術偵測敏感度之實驗設計

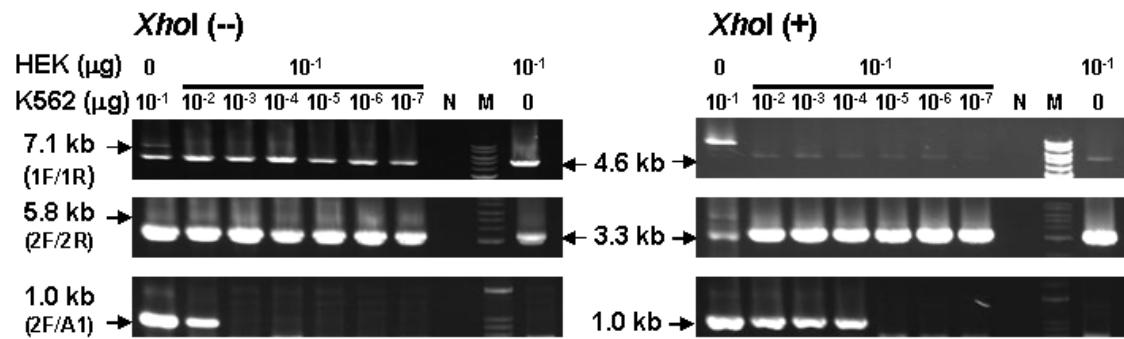
(A)K562 細胞株 der(22)轉位斷裂點之定位。分別位於 BCR 的 intron 14 及 ABL 的 intron 1，以 BglIII 作為初級限制酶，擷取此區域。黑色箭號為基因的轉錄方向，三角形標示處為轉位之斷裂點。(B)以 BglIII 擷取未發生轉位的 BCR 及 ABL 和 BCR/ABL。由 repeat element 分布可知，位於轉位斷裂點 5' 端的 ABL 序列為 repetitive elements 分布密集的區域(repeat masking region)。橫向三角形為引子設計的位置。使用 XhoI 作為二級限制酶，在未發生轉位的 BCR 做切割，以相對提升 BCR/ABL 的訊號。

(圖二之 A 定序結果來自參考文獻 18)

(A)

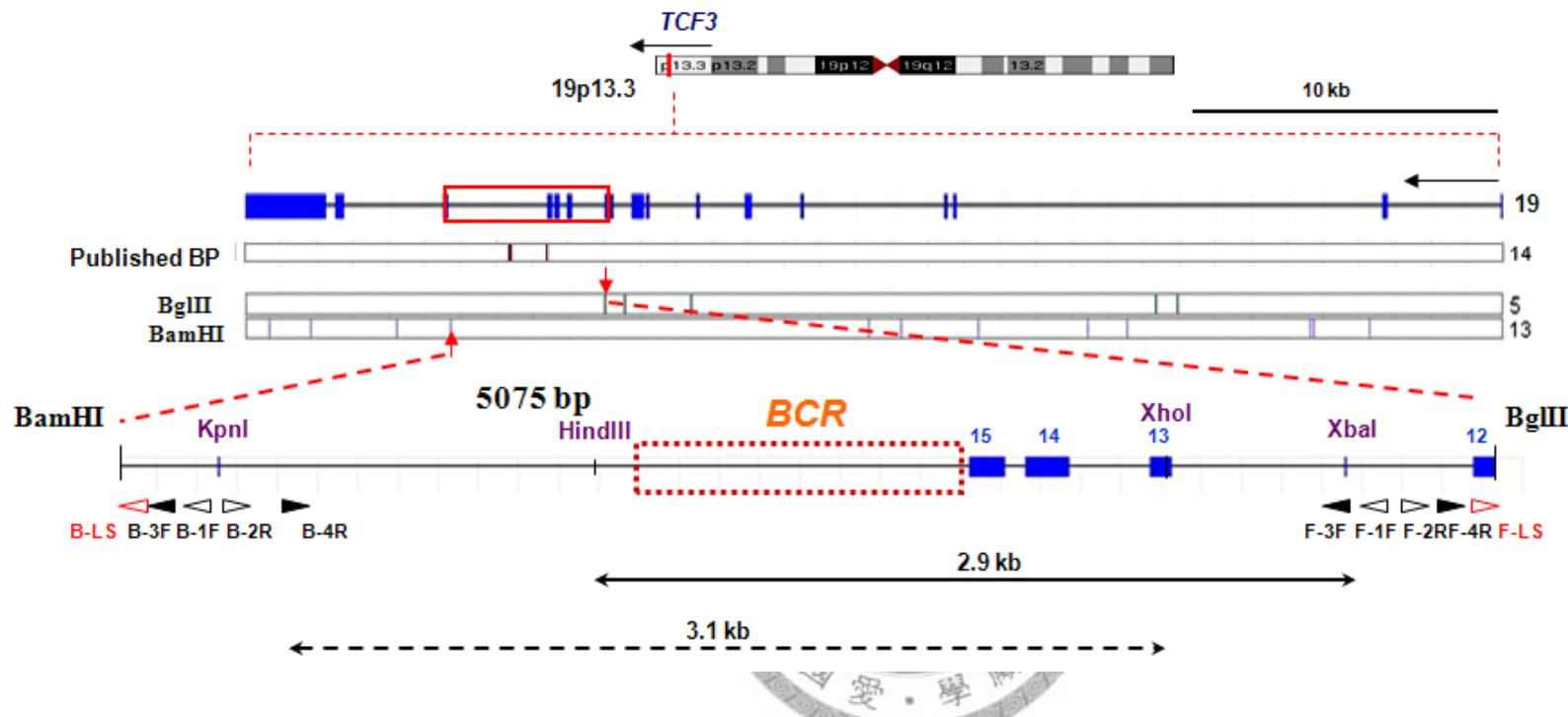


(B)



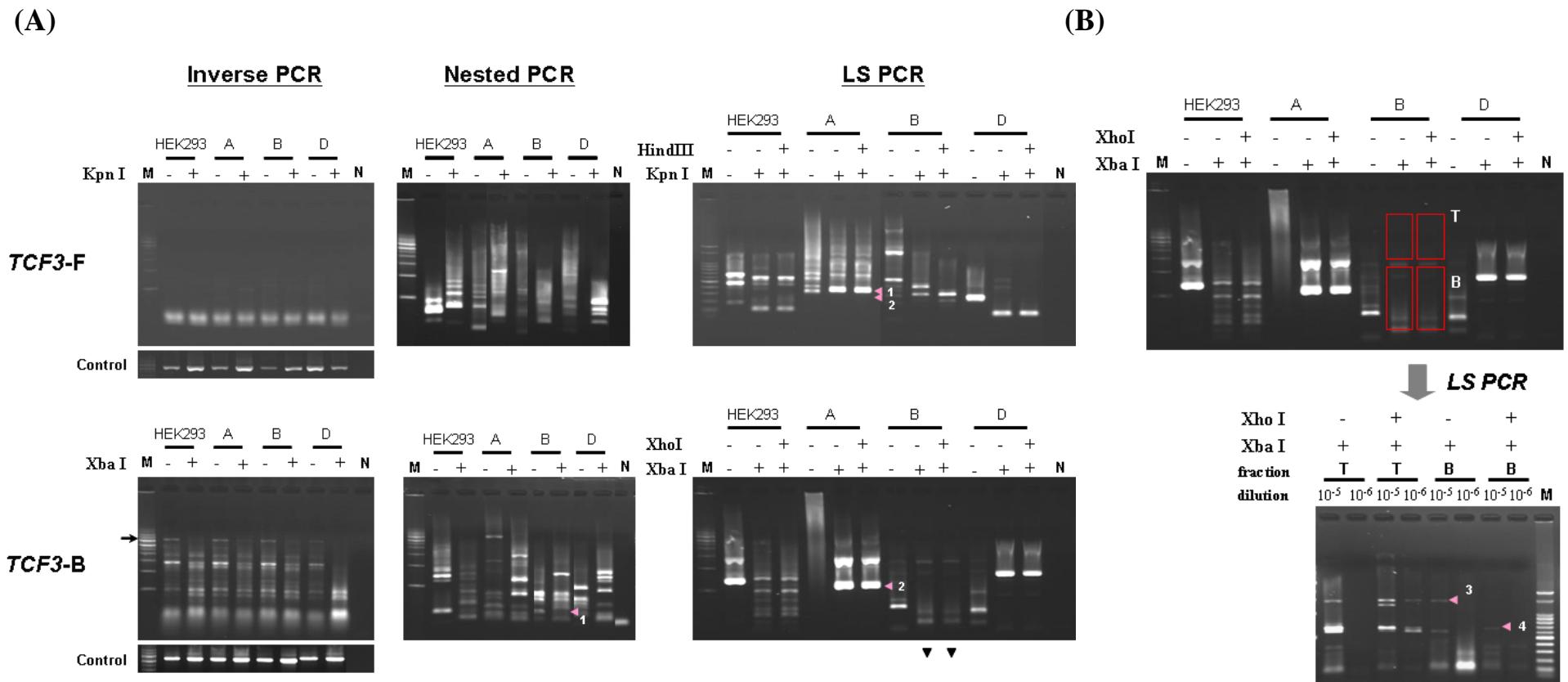
圖三、評估 RECORD 技術偵測敏感度之實驗結果

(A) 使用 K562 gDNA 連續十倍序列稀釋當作模版，作為引子敏感度的測試。(B) 使用 HEK293 gDNA 混合十倍序列稀釋的 K562 gDNA 做當作模版，模擬 RECORD 技術在精細胞中對目標基因轉位的偵測。左方為未經 *Xho*I 處理的模版之結果，右方為經過 *Xho*I 處理的模板之結果。由上而下依序為 inverse PCR，nested PCR 及 TL PCR 之結果。



圖四、*TCF3* 之實驗設計

以 *Bam*HI 及 *Bgl*III 擷取大小約 5.1 kb 涵蓋 *TCF3* 斷裂點叢聚區域(BCR)的片段。於此片段的 5' 端及 3' 端分別設計後置(backward)及前置(forward)兩組引子，並選用不同的二級限制酶；後置引子配合 *Xho*I 及 *Xba*I，可偵測的範圍約 3.1 kb(虛線雙向箭號)，前置引子配合 *Kpn*I 及 *Hind*III，可偵測的範圍約 2.9 kb(實線雙向箭號)。

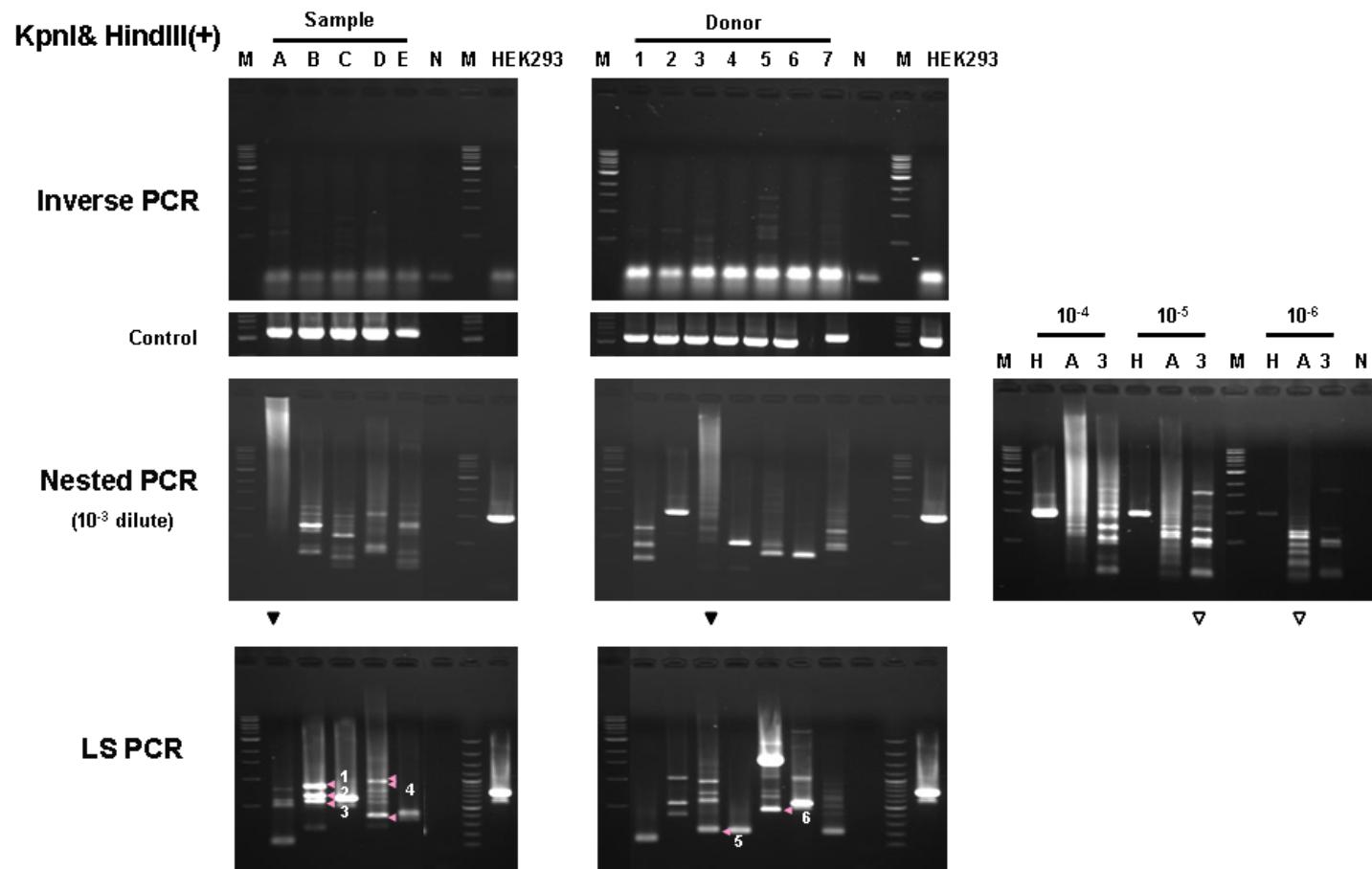


圖五、 *TCF3-F*(3'端)及*TCF3-B*(5'端)模版有無二級限制酶處理之比較

(A)箭號標示處為 unarranged form 模版所放大出來的產物。有二級限制酶處理過的模版之電泳結果與未經二級限制酶處理者，明顯乾淨許多，此外，在 pattern 相似處，有些 band 的亮度有提升的效果。(B)將圖(A)中 Sample B 的 *TCF3-B* 之 LS PCR 產物(倒三角

形標示處)分 fraction(top, T & bottom, B)收集，稀釋後再做一次 LS PCR。電泳結果出現兩條之前未出現的 band(三角形標示處)。三角形標示處為經過選殖與定序後，偵測到的染色體轉位；在 *TCF3*-F 有 2 個，分別如下：1 為 t(10;19)(q22.1;p13.3)，PCR 產物大小為 421 bp；2 為 t(18;19)(q21.1;p13.3)，PCR 產物大小為 383 bp。在 *TCF3*-B 有 4 個，分別如下：1 為 t(12;19) (q21.31;p13.3)，PCR 產物大小為 139 bp；2 為 t(11;19)(p13;p13.3)，PCR 產物大小為 601 bp；3 為 t(1;19)(q32.3;p13.3)，PCR 產物大小為 1407 bp；4 為 t(2;19)(q33.1;p13.3)，PCR 產物大小為 653 bp。



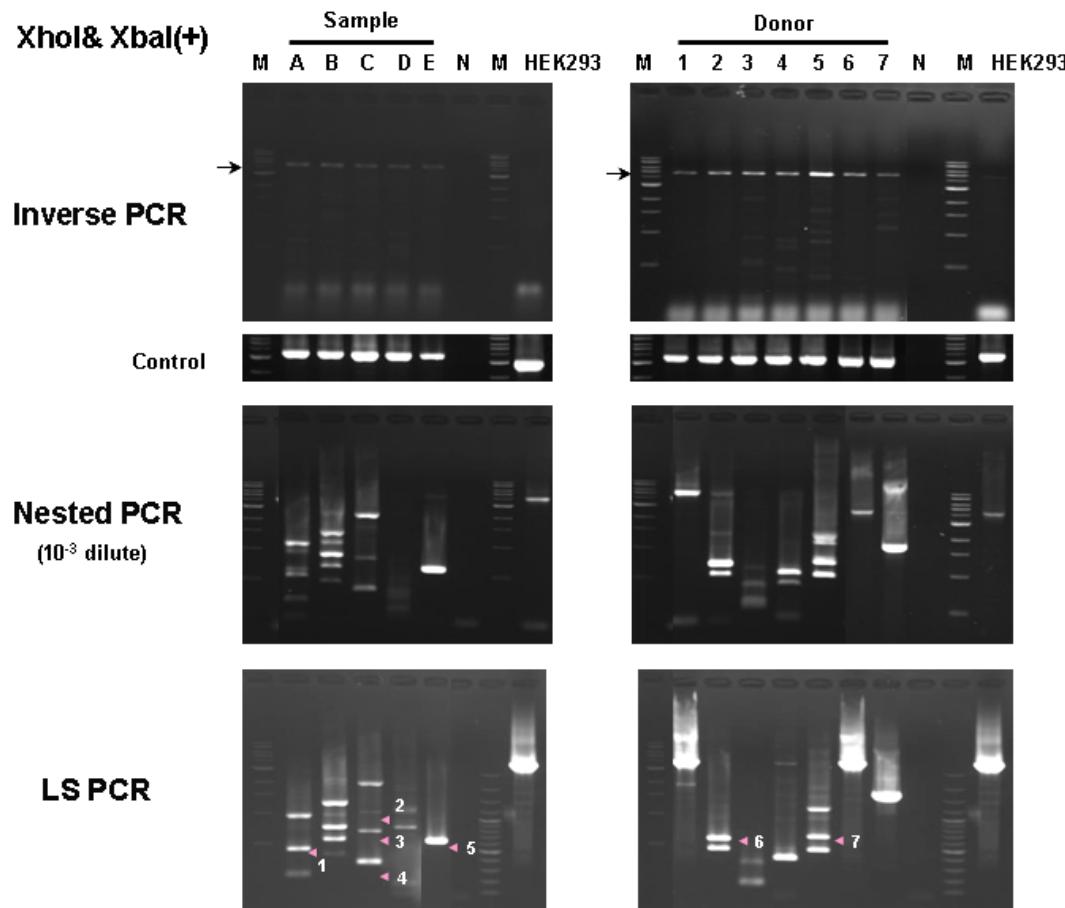


圖六、 TCF3-F(3'端)之 PCR 電泳結果

Sample A 與 Donor 3 在以  $10^{-3}$  倍稀釋的 Inverse PCR 產物當模板時，nested PCR 未有明顯的 band 出現(實心倒三角形標示處)，將 Inverse PCR 產物往下稀釋，Nested PCR 即有明顯的 band 出現(空心倒三角形標示處)。不同來源的精細胞 gDNA 模板所得的 LS PCR

pattern 有明顯的差異性存在。三角形標示處為經過選殖與定序後，偵測到的染色體轉位，共有 6 個，分別如下：1 為  $t(1;19)(q23.3;p13.3)$ ，PCR 產物大小為 836 bp；2 為  $t(1;19)(q32.1;p13.3)$ ，PCR 產物大小為 647 bp；3 為  $t(17;19)(q11.2;p13.3)$ ，PCR 產物大小為 547 bp；4 為  $t(3;19)(p22.2;p13.3)$ ，同一條 lane 的 3 個箭號標示均為此轉位所造成，大小分別為 989 bp，947 bp 及 358 bp；5 為  $t(3;19)(p21.32;p13.3)$ ，PCR 產物大小為 212 bp；6 為  $t(5;19)(q35.2;p13.3)$ ，PCR 產物大小為 434 bp。





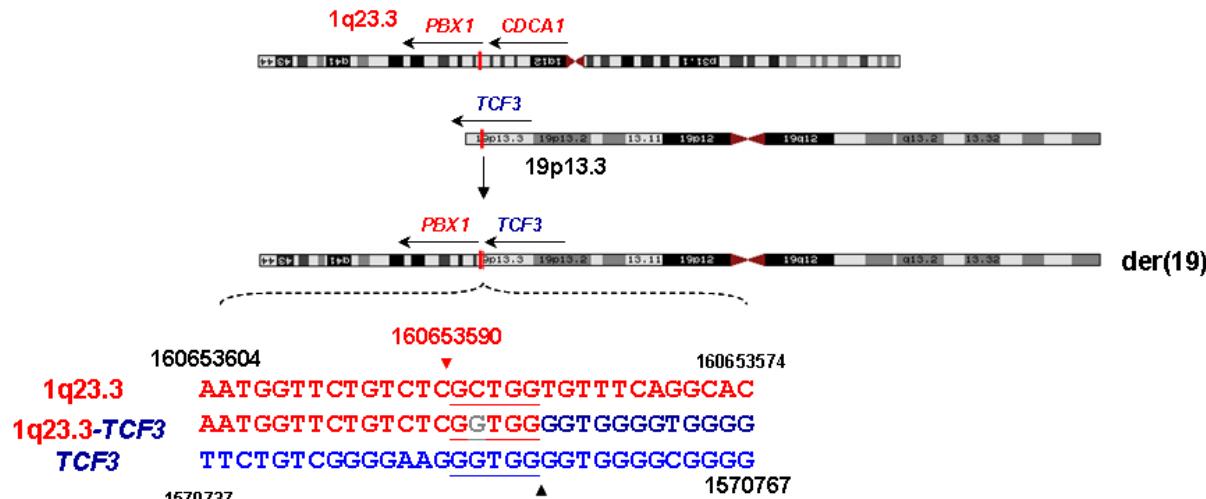
圖七、*TCF3-B(5'端)*之 PCR 電泳結果

箭頭標示處為未被二級限制酶處理乾淨的 unarranged form 模版所放大出來的產物。不同來源的精細胞 gDNA 模板所得的 LS PCR pattern 有明顯的差異存在。三角形標示處為經過選殖與定序後，偵測到 7 個染色體轉位，分別如下：1 為 t(2;19)(p24.3;p13.3)，PCR

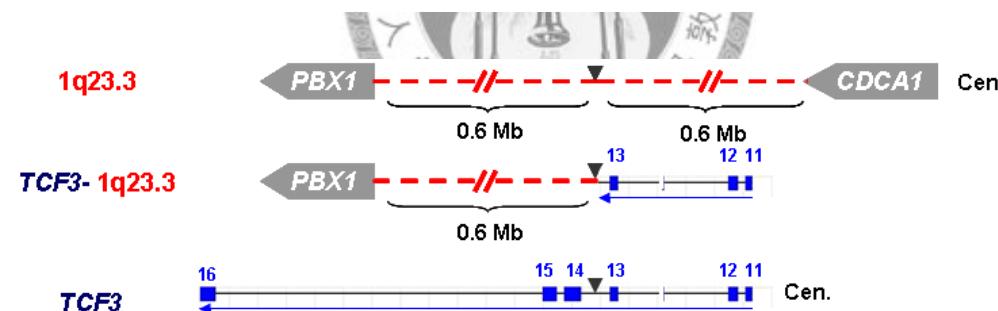
產物大小為 422 bp；2 為 t(8;19)(q22.2;p13.3)，PCR 產物大小為 1019 bp；3 為 t(4;19)(q22.1;p13.3)，PCR 產物大小為 666 bp；4 為 t(8;19)(q11.23;p13.3)，PCR 產物大小為 289 bp；5 為 t(6;19)(q21;p13.3)，PCR 產物大小為 529 bp；6 為 t(10;19)(p12.1;p13.3)，PCR 產物大小為 586 bp；7 為 t(11;19)(q13.3;p13.3)，PCR 產物大小為 629 bp。



A.

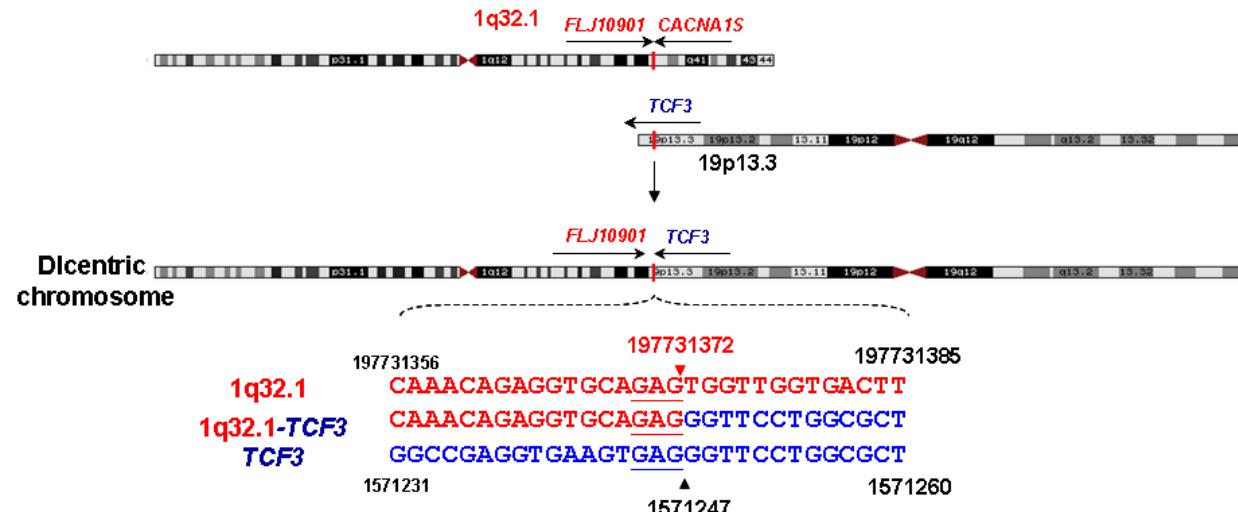


B.

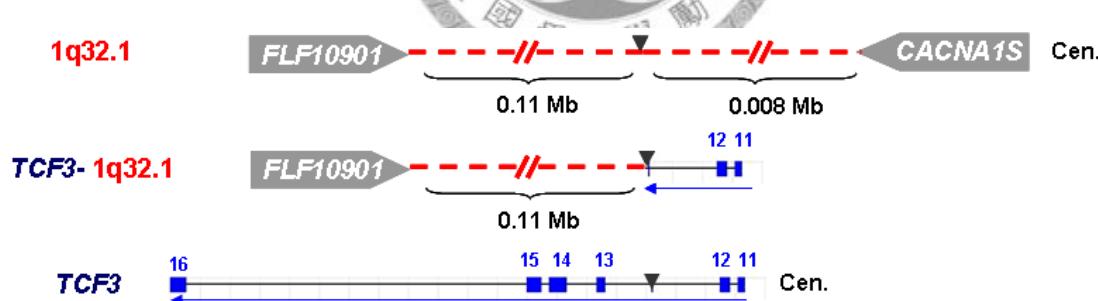


圖八之一、*TCF3-F(3'端)*所偵測出的染色體轉位 t(1;19)(q23.3;p13.3)

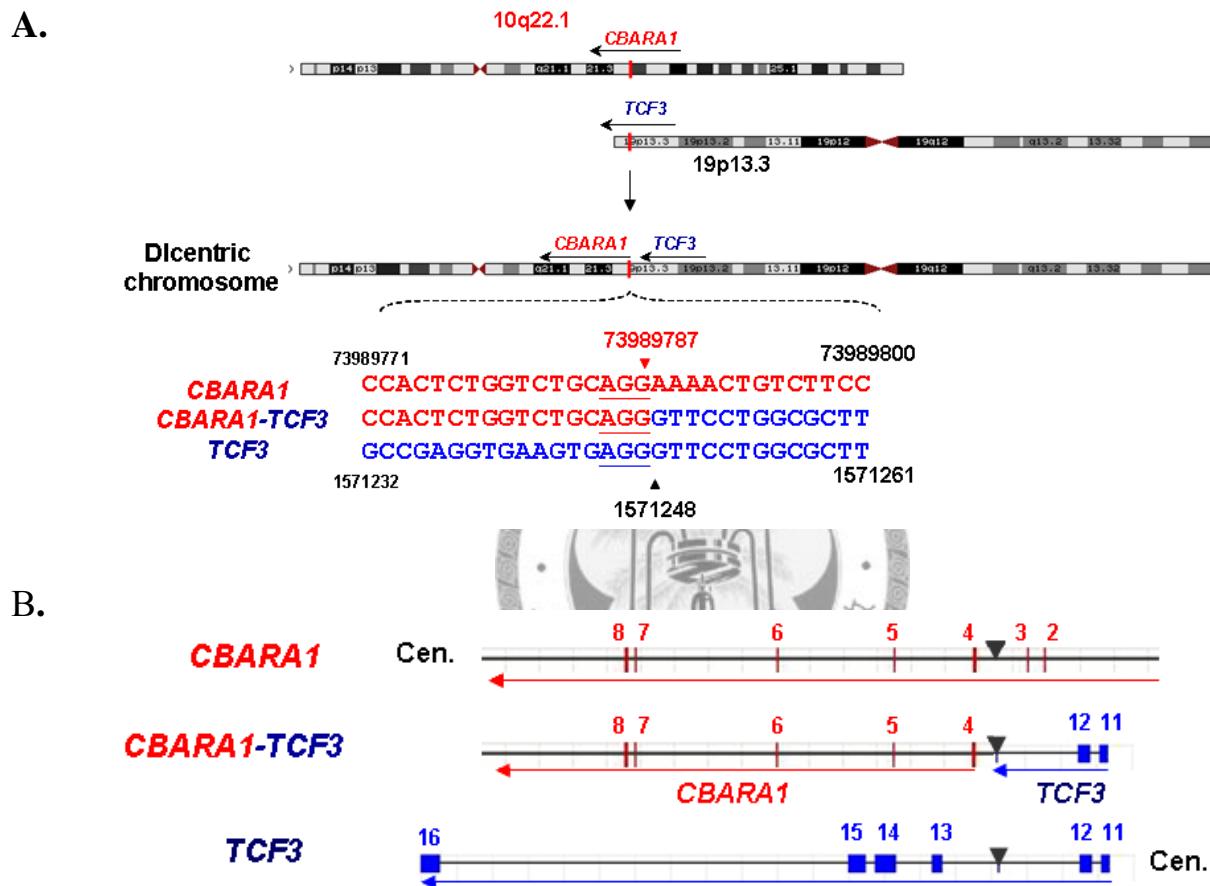
A.



B.

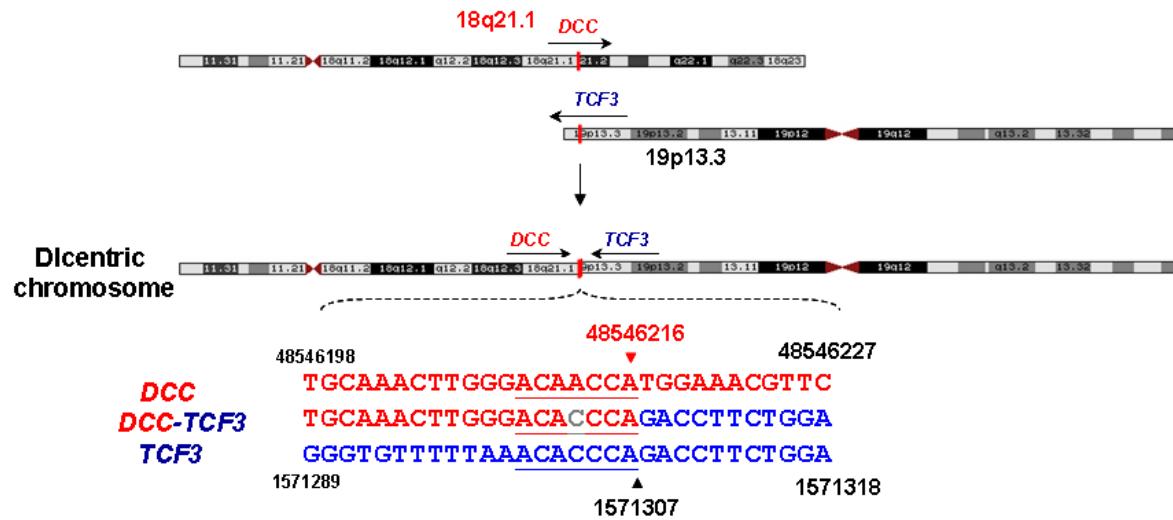


圖八之二、*TCF3-F(3'端)*所偵測出的染色體轉位 t(1;19)(q32.1;p13.3)

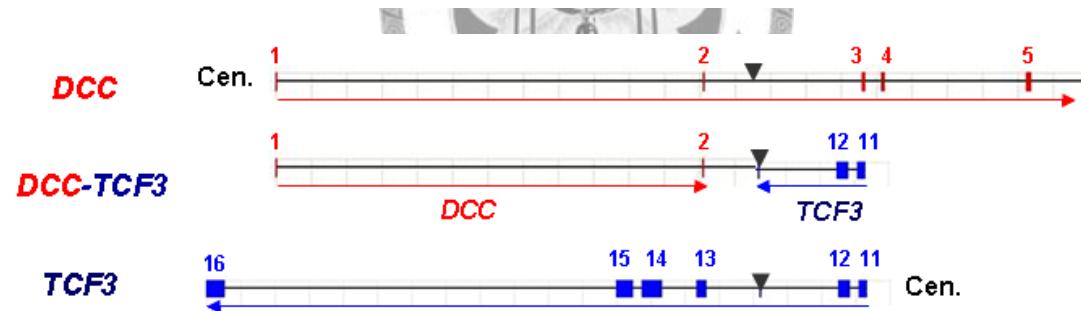


圖八之三、*TCF3-F(3'端)*所偵測出的染色體轉位  $t(10;19)(q22.1;p13.3)$

A.

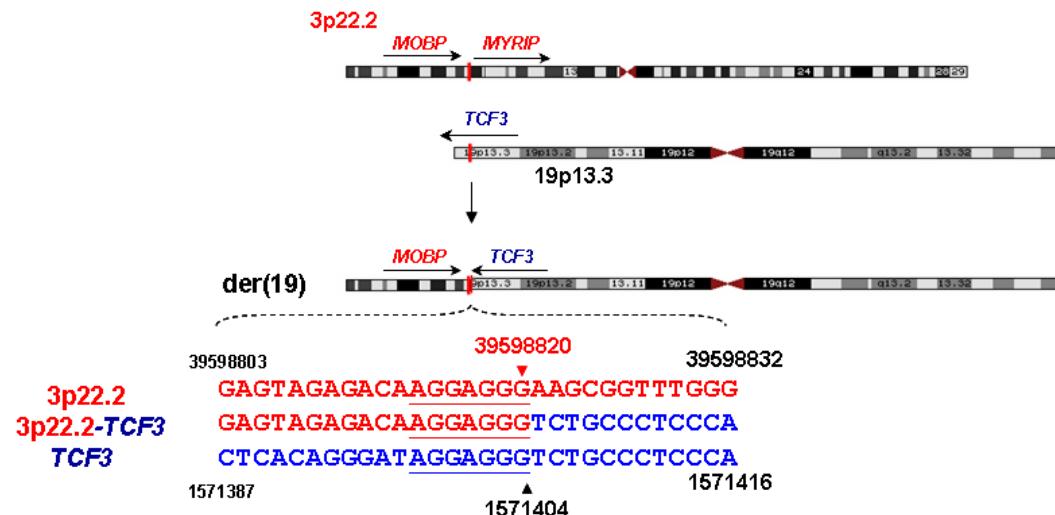


B.

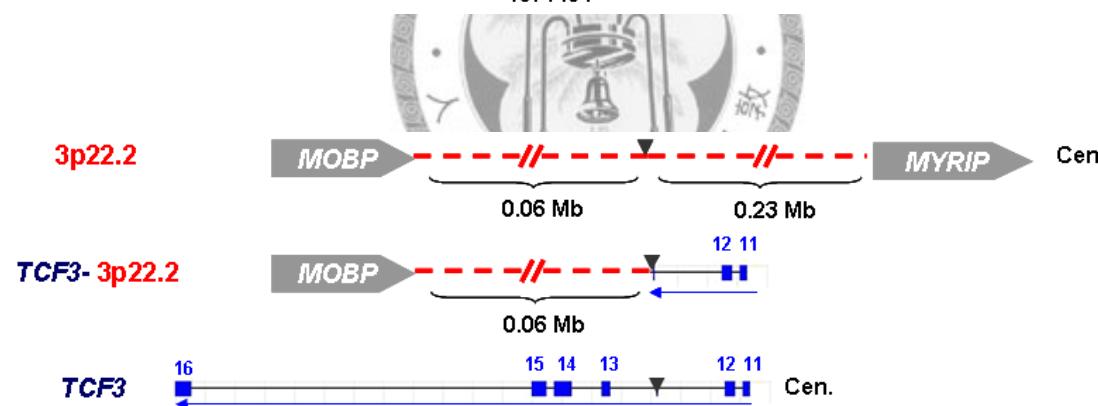


圖八之四、*TCF3-F(3'端)*所偵測出的染色體轉位 t(18;19)(q21.1;p13.3)

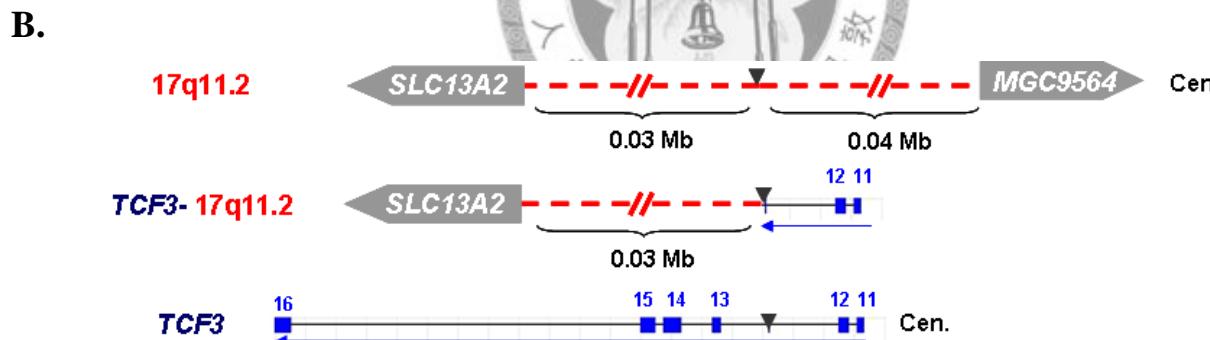
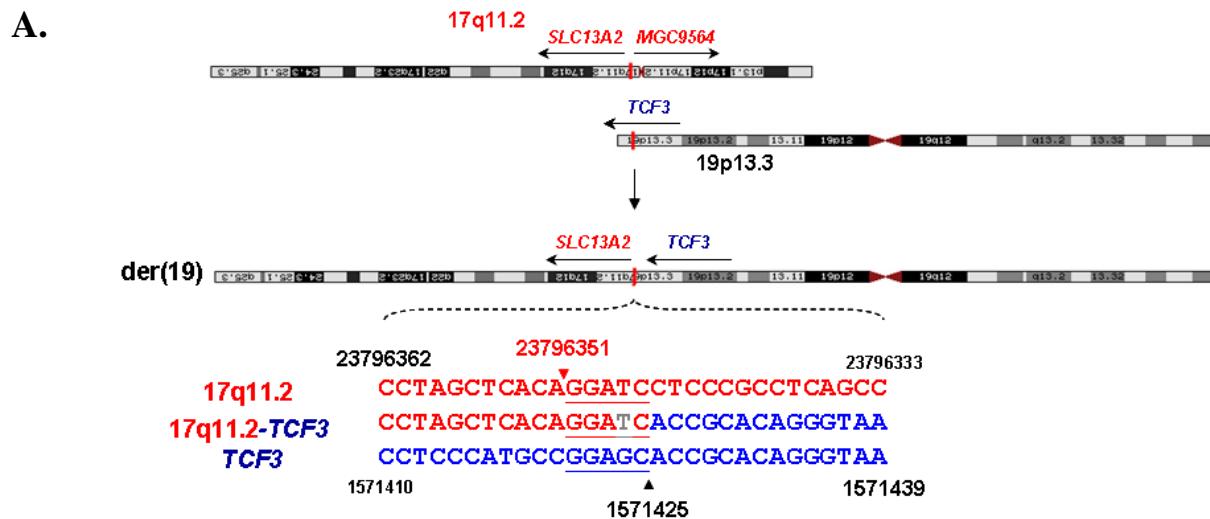
A.



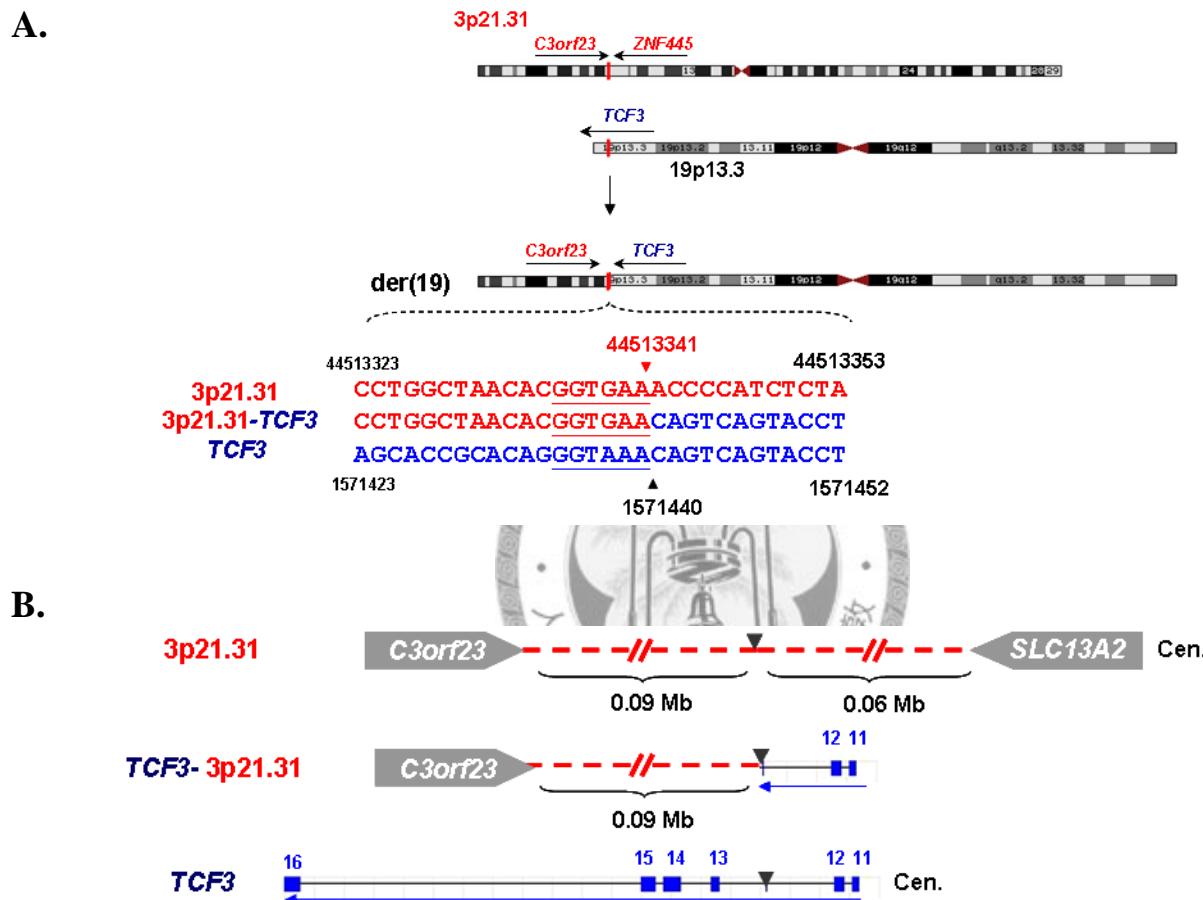
B.



圖八之五、*TCF3*-F(3'端)所偵測出的染色體轉位 *t*(3;19)(p22.2;p13.3)

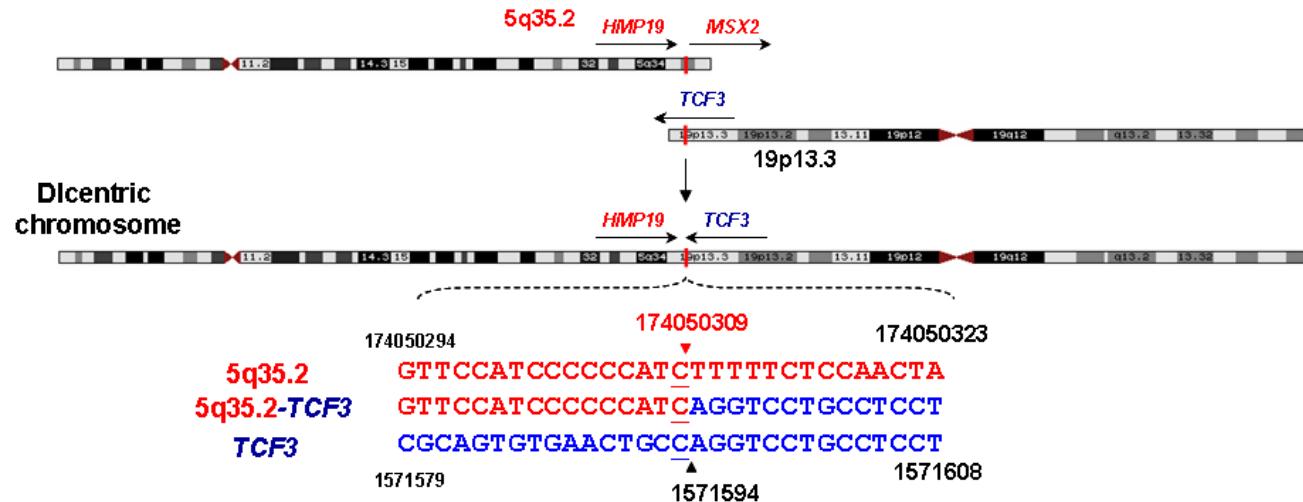


圖八之六、 *TCF3-F(3'端)*所偵測出的染色體轉位 t(17;19)(q11.2;p13.3)

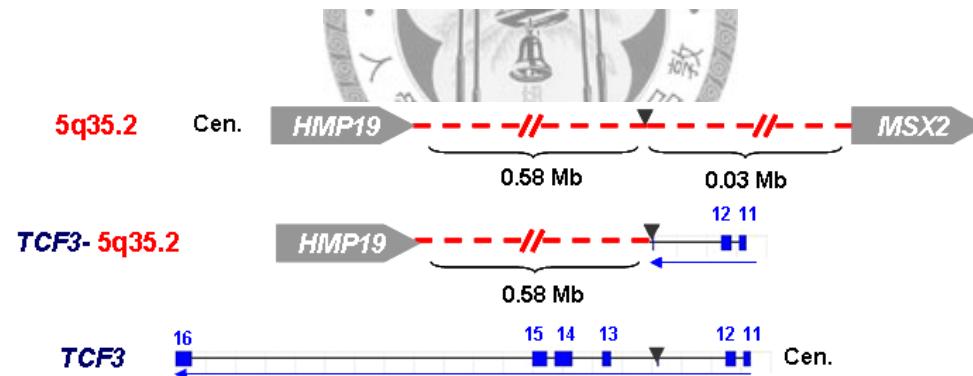


圖八之七、*TCF3-F(3'端)*所偵測出的染色體轉位 t(3;19)(p21.31;p13.3)

A.

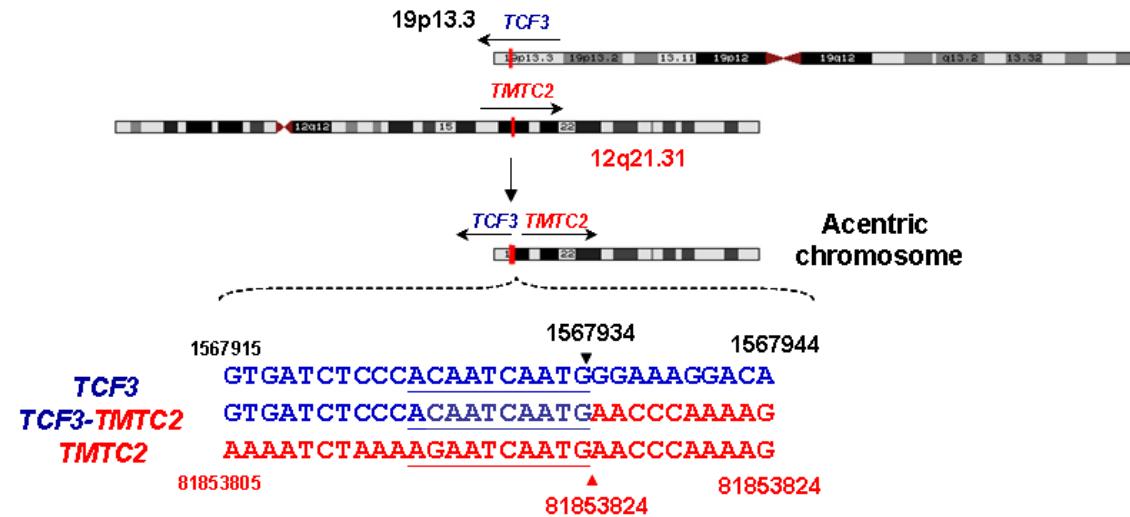


B.

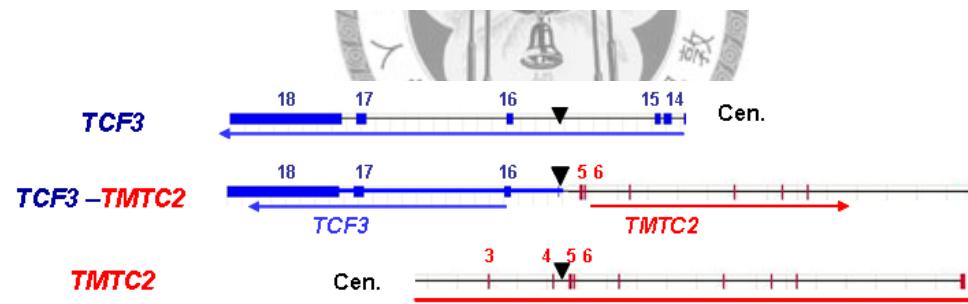


圖八之八、*TCF3-F(3'端)*所偵測出的染色體轉位  $t(5;19)(q35.2;p13.3)$

A.

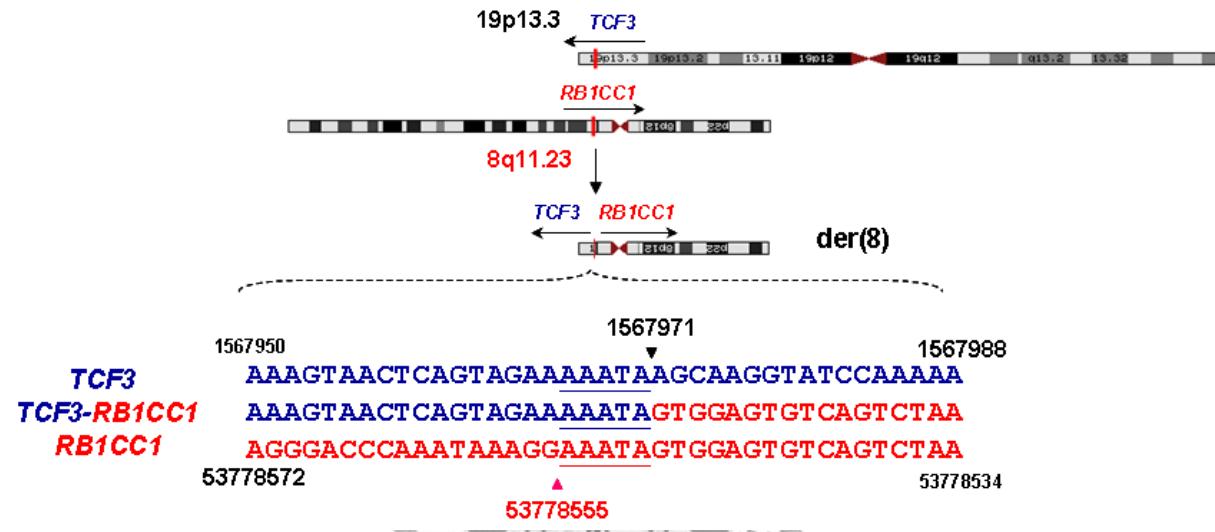


B.

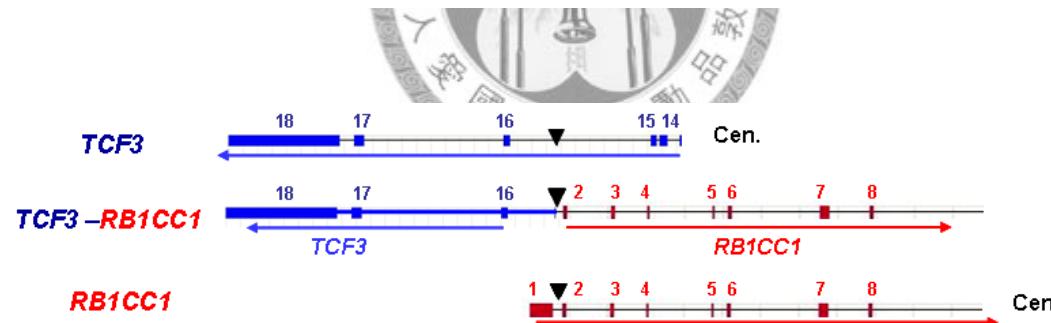


圖九之一、*TCF3*-B(5'端)所偵測出的染色體轉位 *t*(12;19)(q21.31;p13.3)

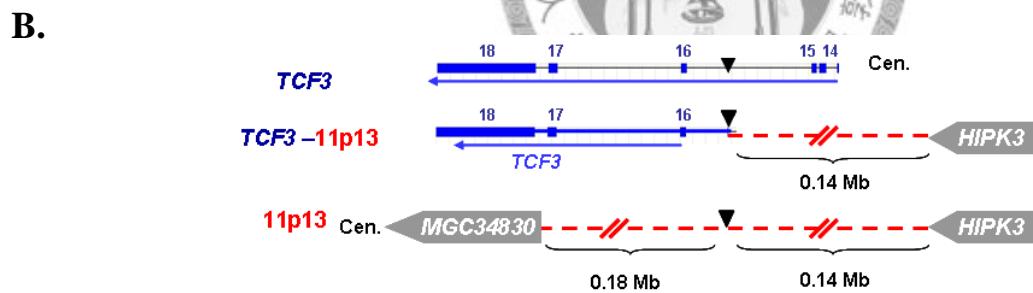
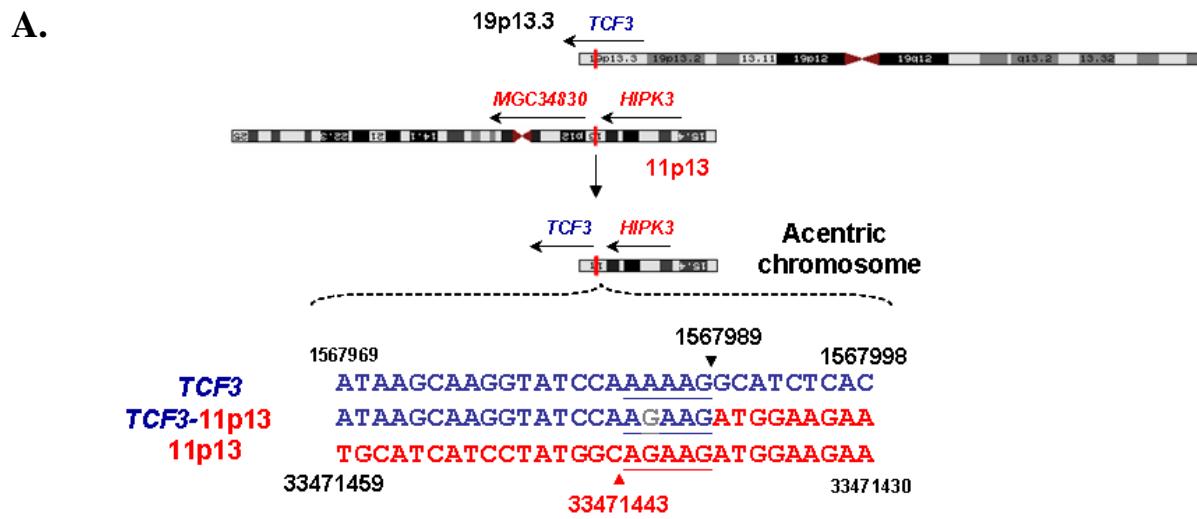
A.



B.

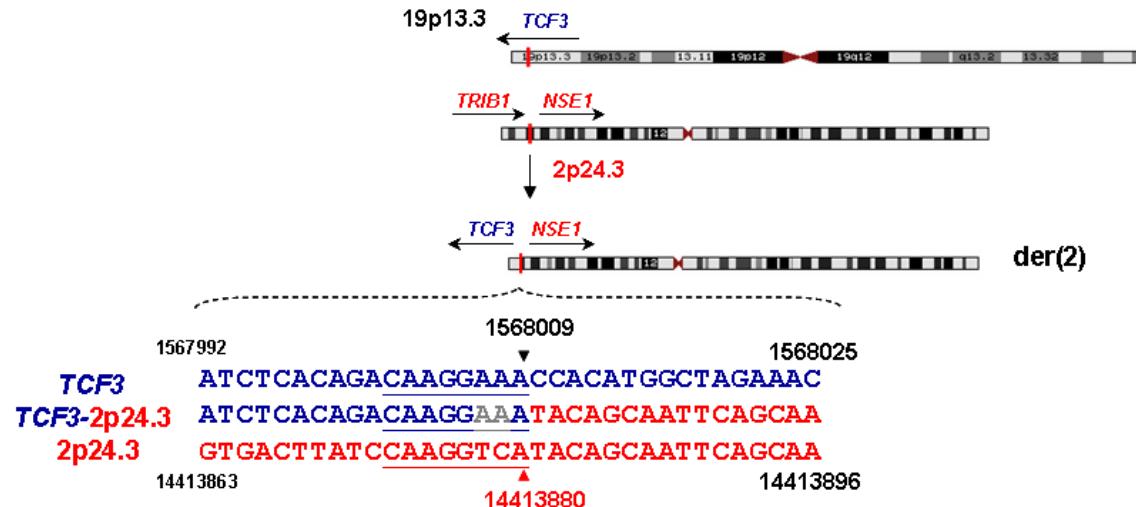


圖九之二、 *TCF3-B*(5'端)所偵測出的染色體轉位 *t*(8;19)(q11.23;p13.3)

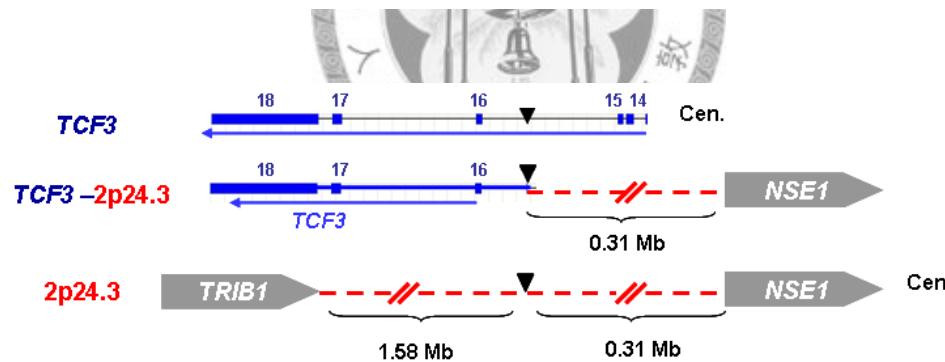


圖九之三、*TCF3-B*(5'端)所偵測出的染色體轉位  $t(11;19)(p13;p13.3)$

A.

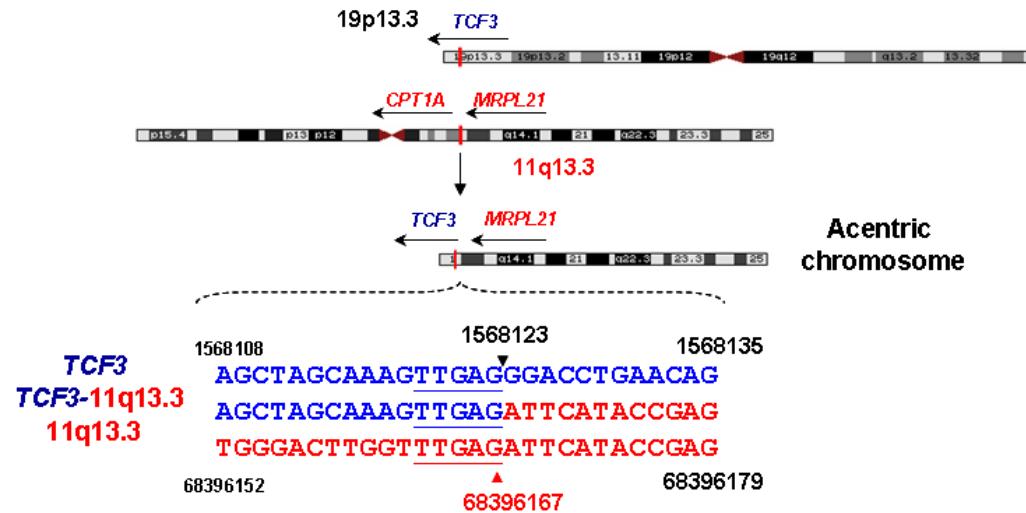


B.

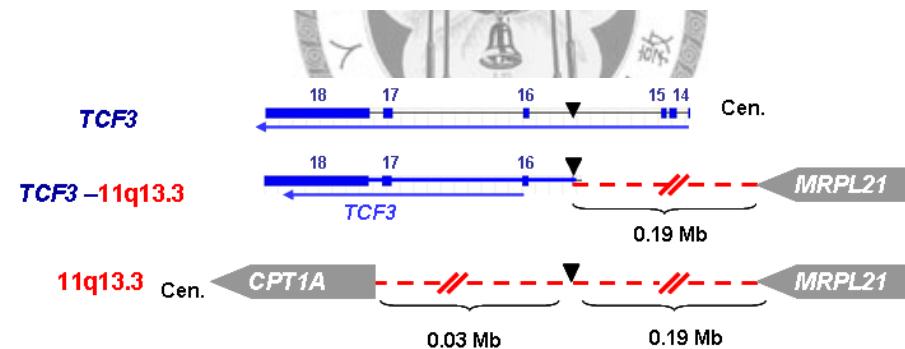


圖九之四、*TCF3*-B(5'端)所偵測出的染色體轉位 *t*(2;19)(p24.3;p13.3)

A.

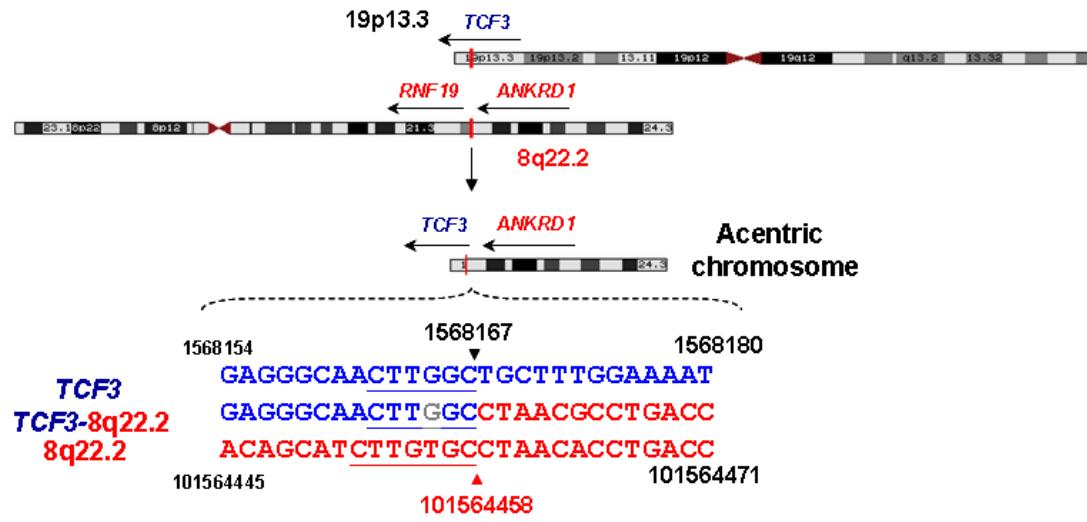


B.

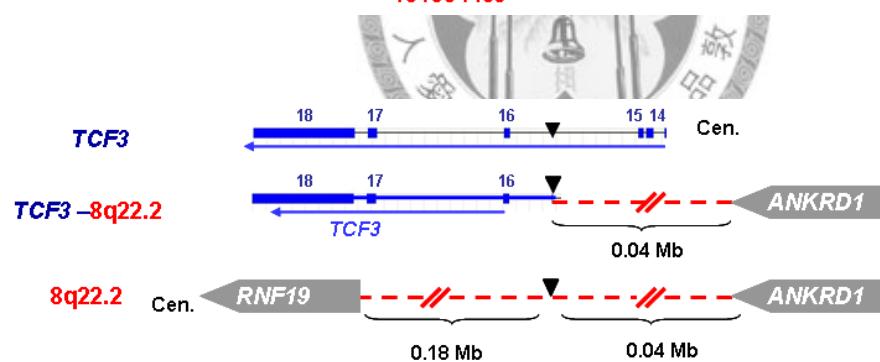


圖九之五、*TCF3-B(5'端)*所偵測出的染色體轉位 *t*(11;19)(q13.3;p13.3)

A.

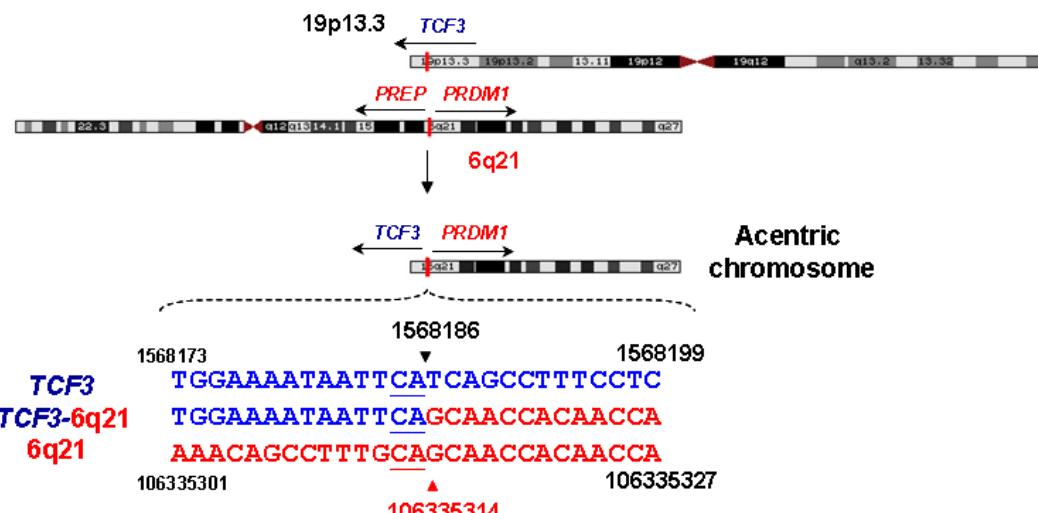


B.

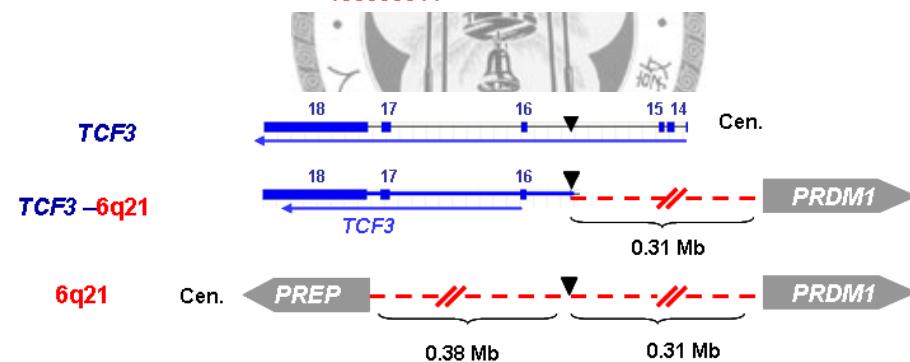


圖九之六、*TCF3-B*(5'端)所偵測出的染色體轉位 *t*(8;19)(q22.2;p13.3)

A.

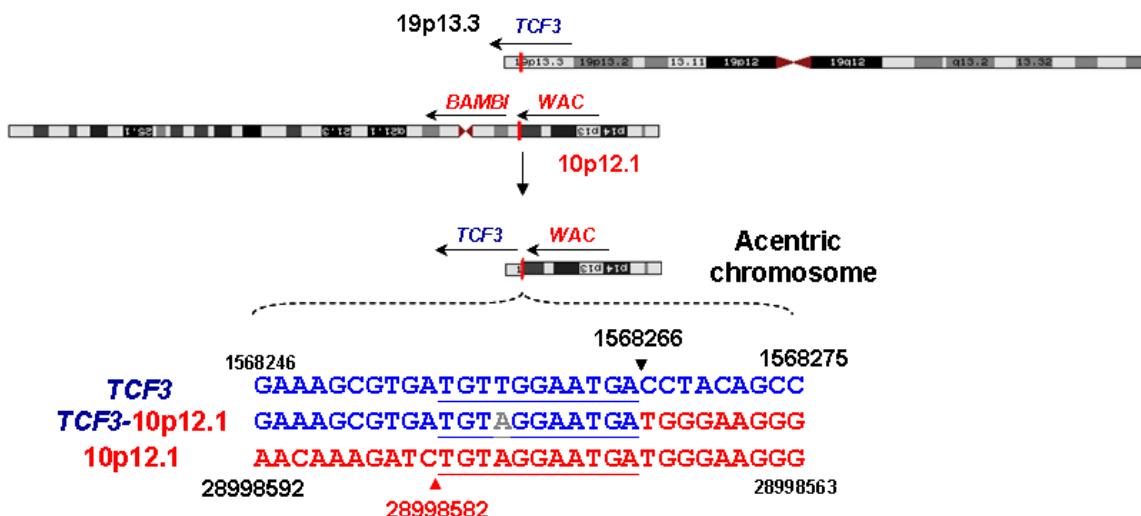


B.

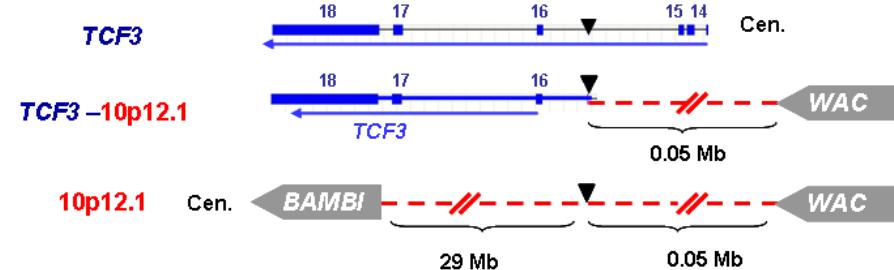


圖九之七、*TCF3-B(5'端)*所偵測出的染色體轉位  $t(6;19)(q21;p13.3)$

A.

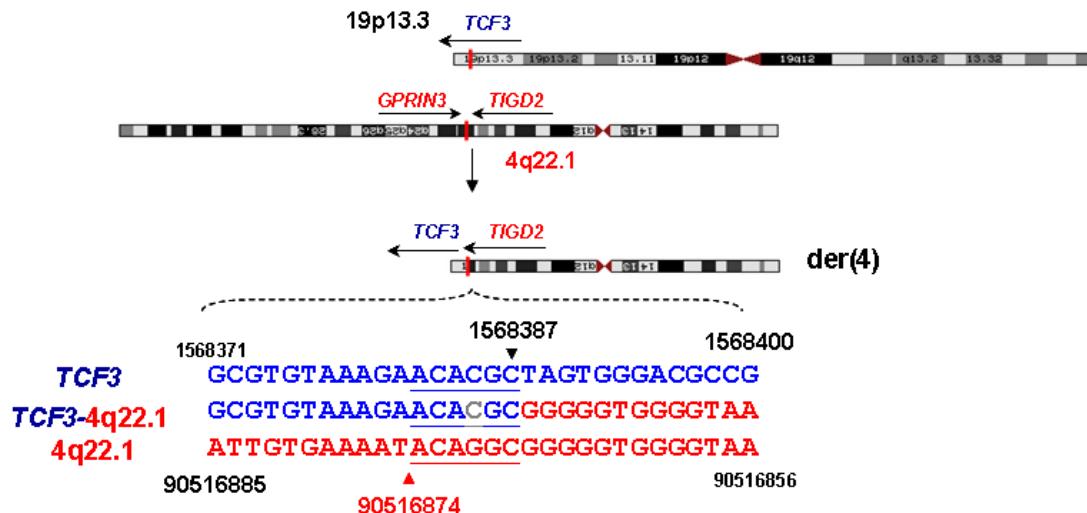


B.

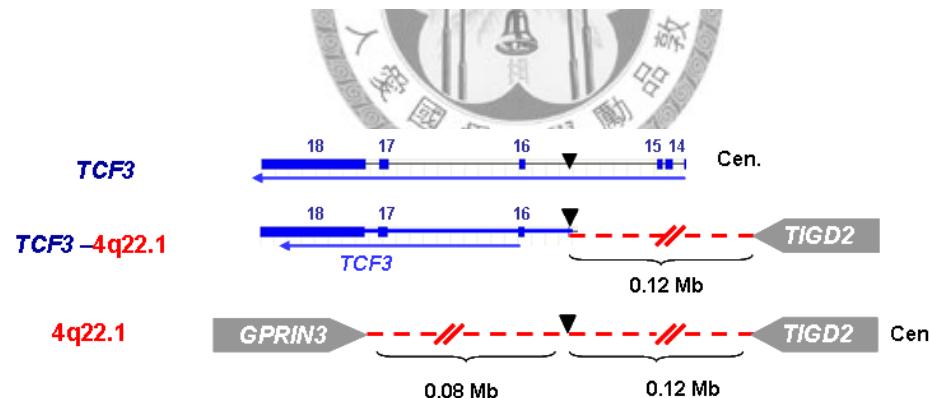


圖九之八、*TCF3-B(5'端)*所偵測出的染色體轉位  $t(10;19)(p12.1;p13.3)$

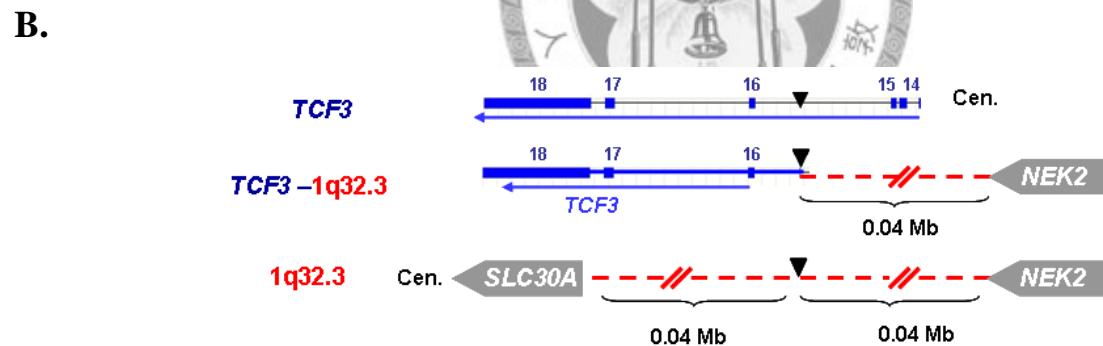
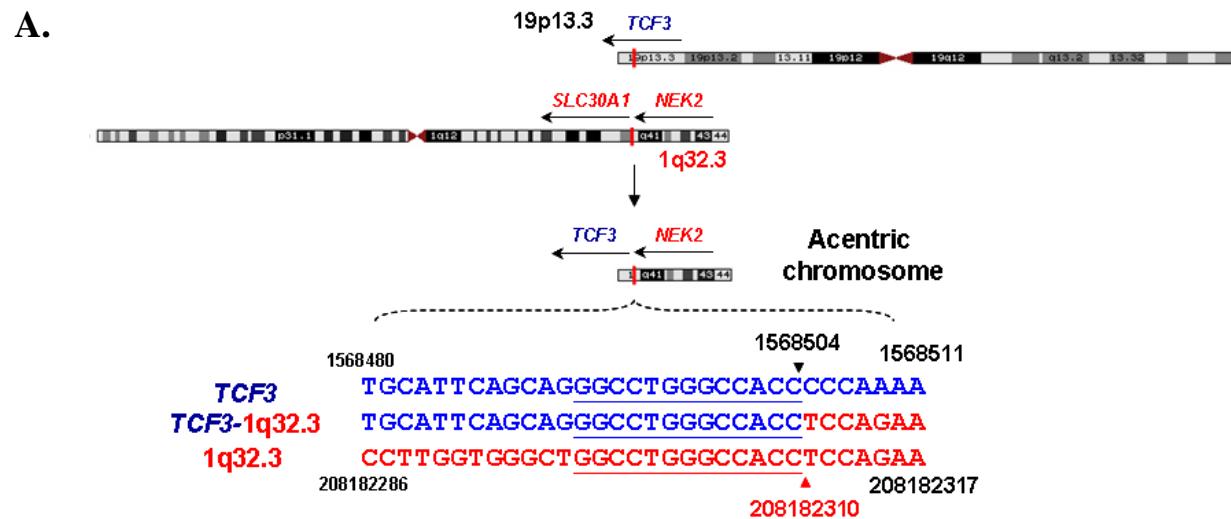
A.



B.

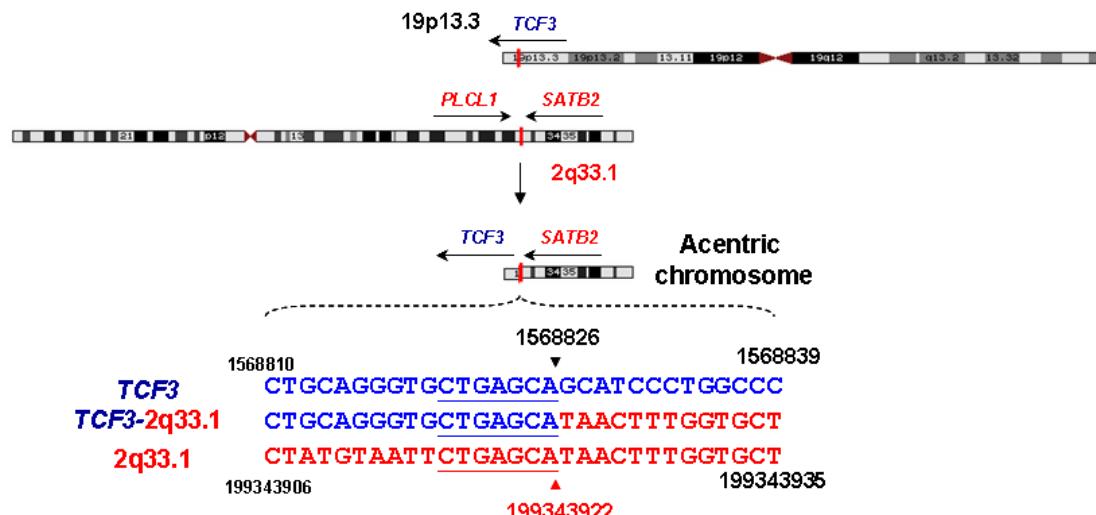


圖九之九、 *TCF3*-B(5'端)所偵測出的染色體轉位 *t*(4;19)(q22.1;p13.3)

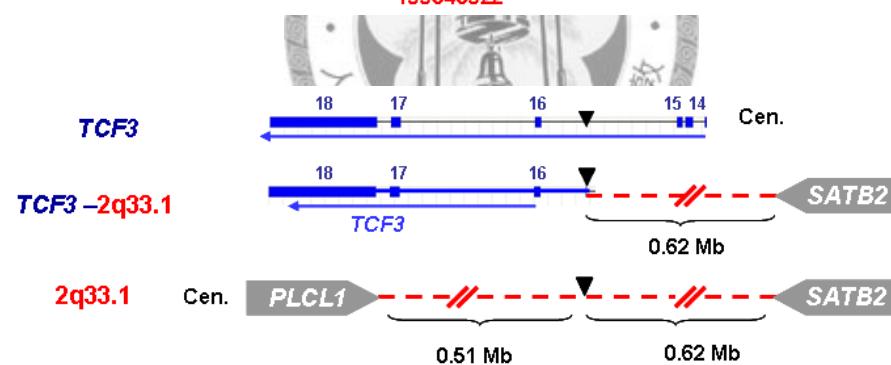


圖九之十、*TCF3-B(5'端)*所偵測出的染色體轉位  $t(1;19)(q32.3;p13.3)$

A.

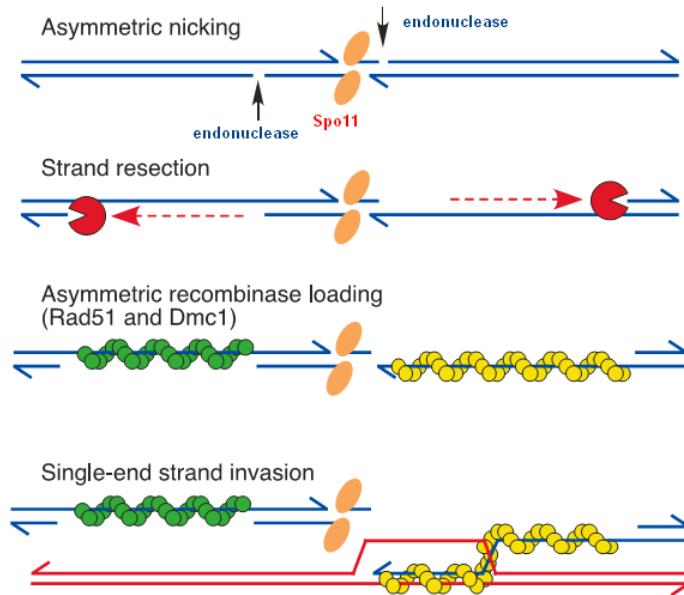


B.

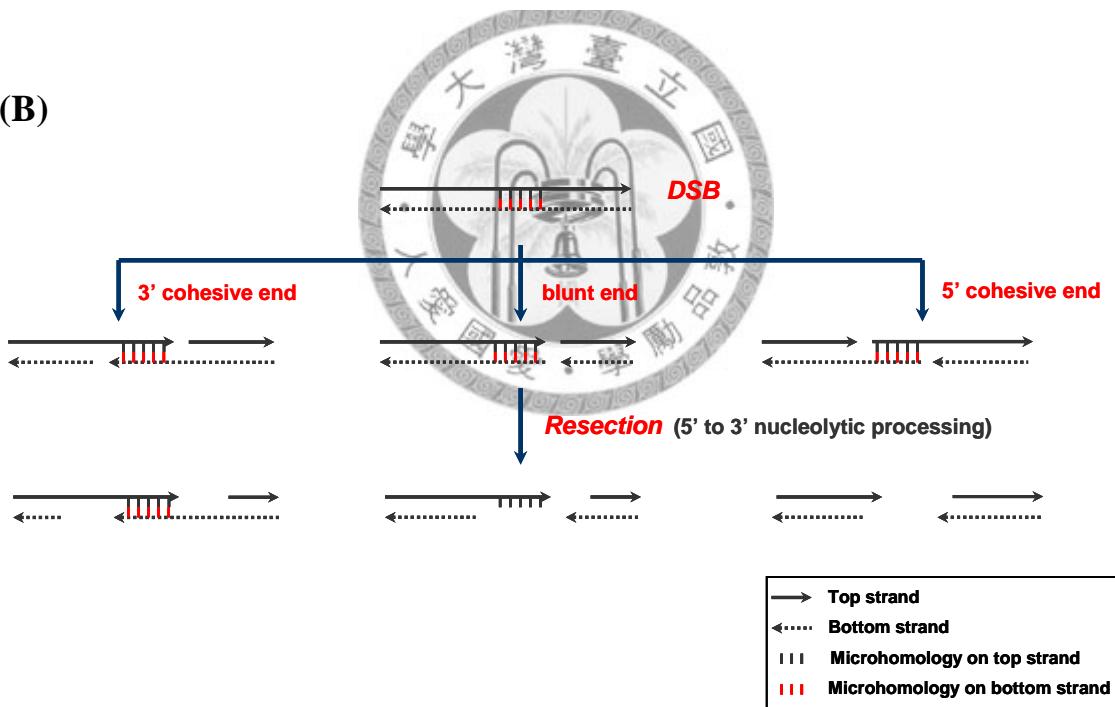


圖九之十一、 *TCF3-B(5'端)*所偵測出的染色體轉位  $t(2;19)(q33.1;p13.3)$

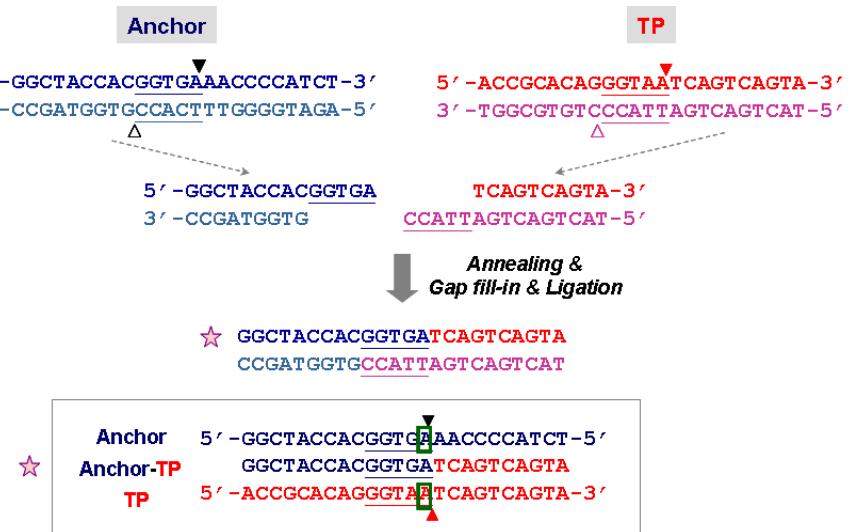
(A)



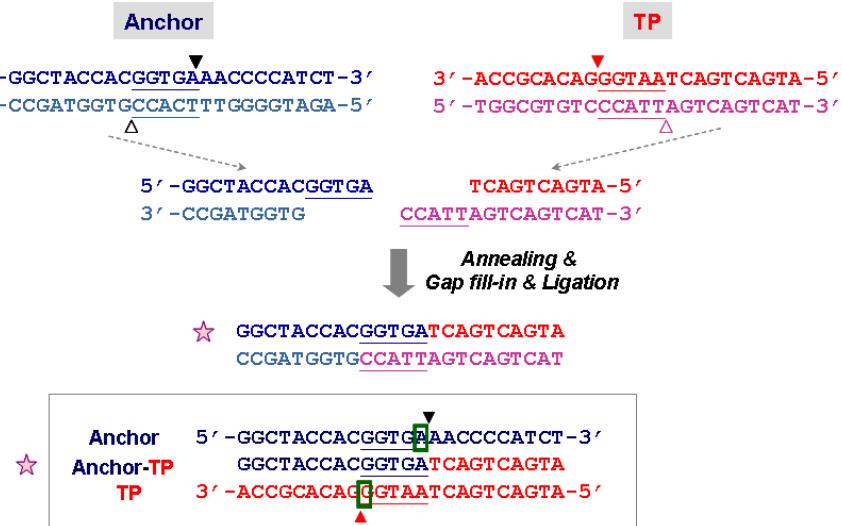
(B)



(C) a.



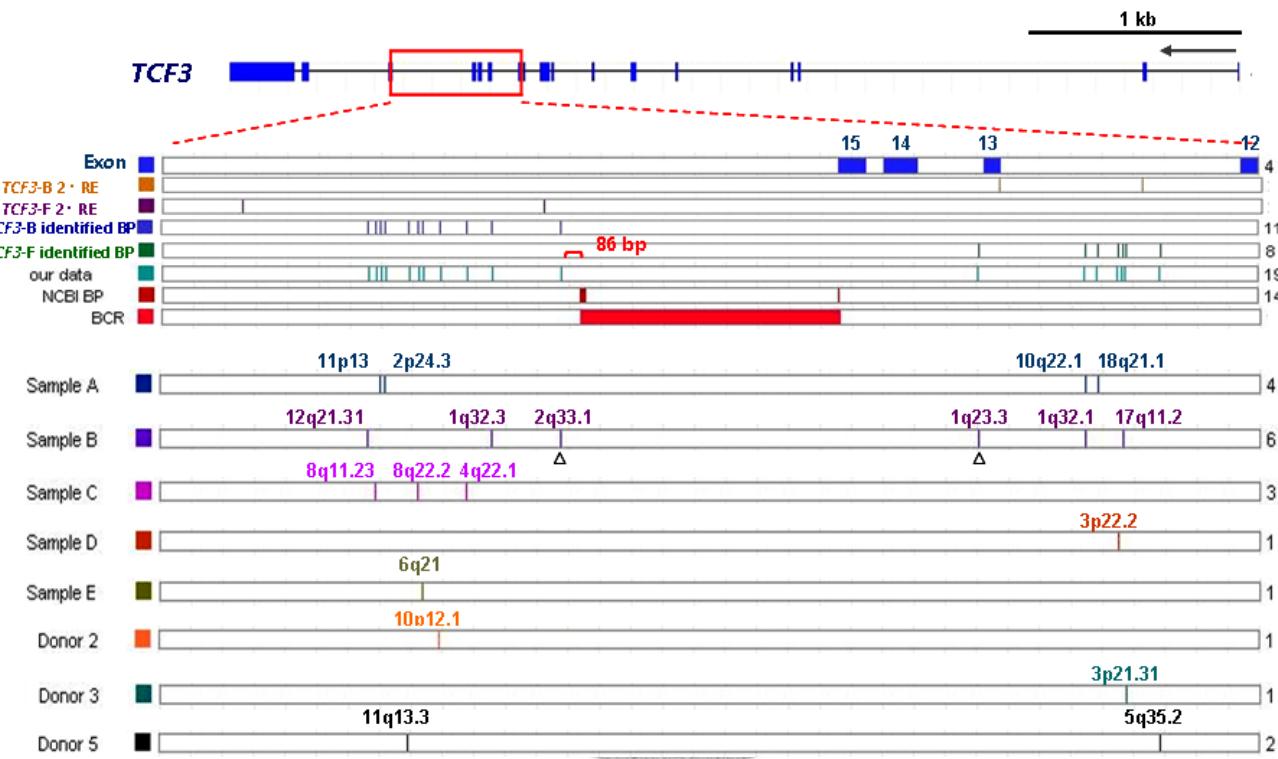
b.



#### 圖十、推測生殖細胞 DNA 雙股斷裂的可能形式及斷裂點之定位方式

- (A)減數分裂時由 Spo11 產生 DNA 雙股斷裂，經 asymmetric nicking 及  $5' \rightarrow 3'$  resection 後，以 HR 進行修復。(B)推測 DNA 雙股斷裂的方式及經過  $5' \rightarrow 3'$  resection 後的結果。若 DNA 雙股斷裂方式為產生 3'黏性末端時，無論是位於正股(top strand)或是負股(bottom strand)的微同源性序列(microhomology sequence)均得以被保留。
- (C)生殖細胞染色體轉位斷裂點定位方式。下方有橫線標示之核苷酸序列为 microhomology；三角形標示處為推測的 DNA 雙股斷裂處；星號標示處為偵測到之染色體轉位序列；方框標示之核苷酸為斷裂點定位處。

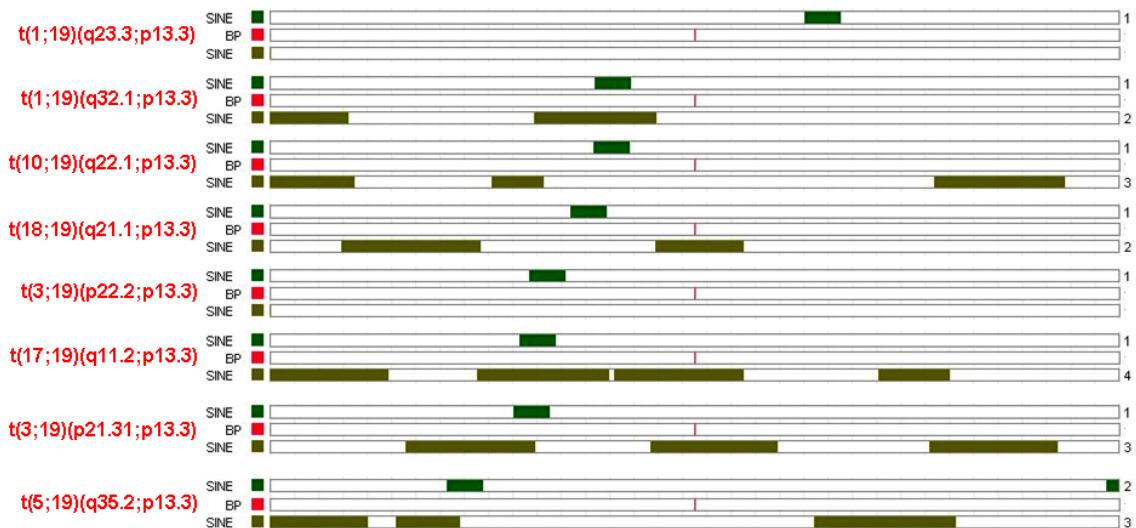
(圖十之 A 選自參考文獻 20)



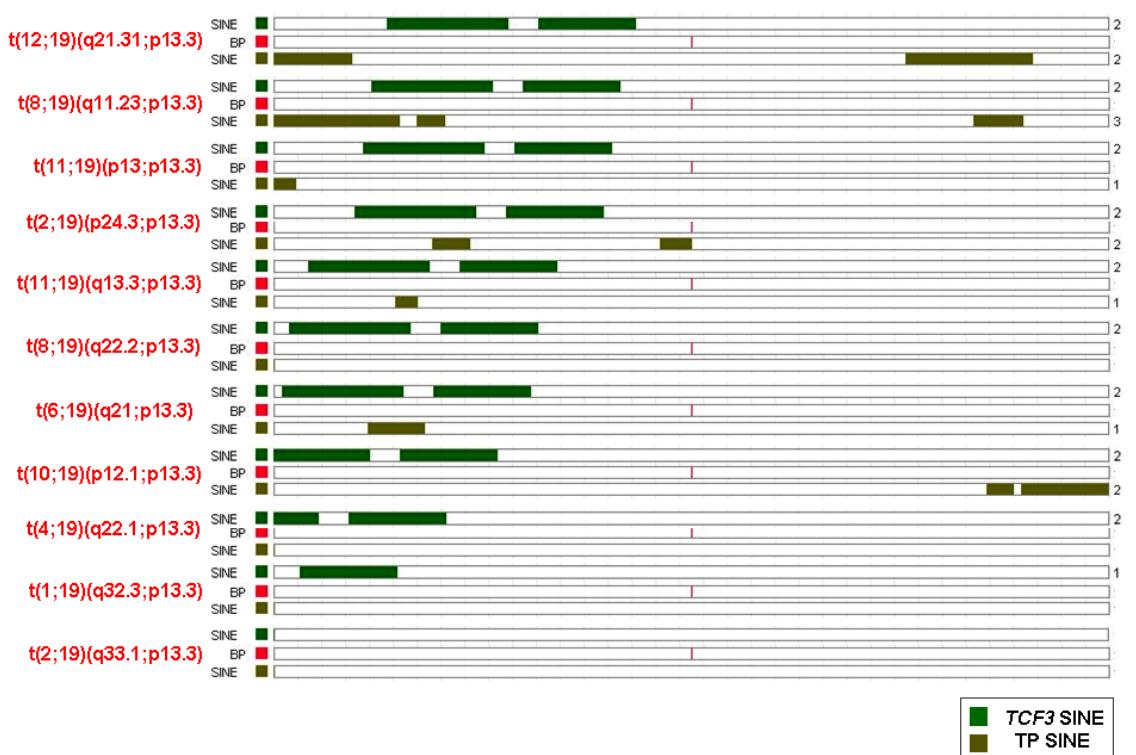
圖十一、*TCF3* 之染色體轉位斷裂點分布

由前置引子(3'端)及後置引子(5'端)偵測出的染色體轉位斷裂點，分別在 *TCF3* intron 12 及 intron 15 有密集分佈的情形。偵測到的斷裂點均未分布於文獻所發表的斷裂點叢聚區域(BCR)，與之相距最近的距離為 86 bp。箭頭方向表示 *TCF3* 的轉錄方向。由不同來源精細胞 gDNA(between individuals)偵測出的染色體轉位，在 *TCF3* 上的斷裂點不同；而斷裂位置在個體本身(within individual)也有差異，且轉位對象亦不同。

(A)



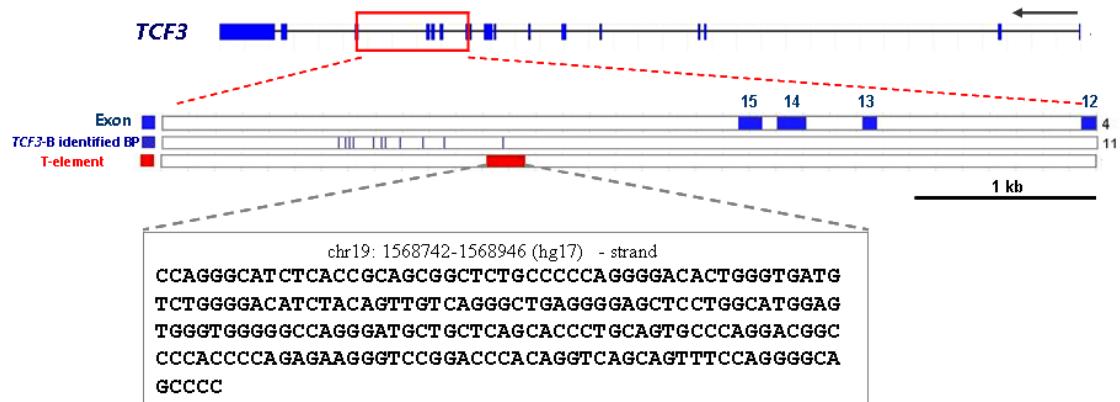
(B)



圖十二、*TCF3* 之染色體轉位斷裂點周邊序列 SINE 分布

(A)*TCF3-F(3'端)*偵測出之染色體轉位斷裂點周邊 2 kb 序列 SINE 分布。(B)*TCF3-B(5'端)*偵測出之染色體轉位斷裂點周邊 2 kb 序列 SINE 分布。

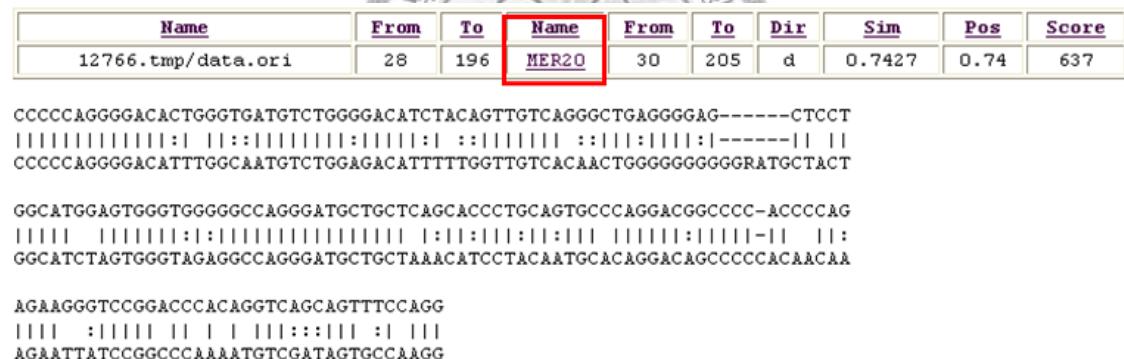
(A)



(B)

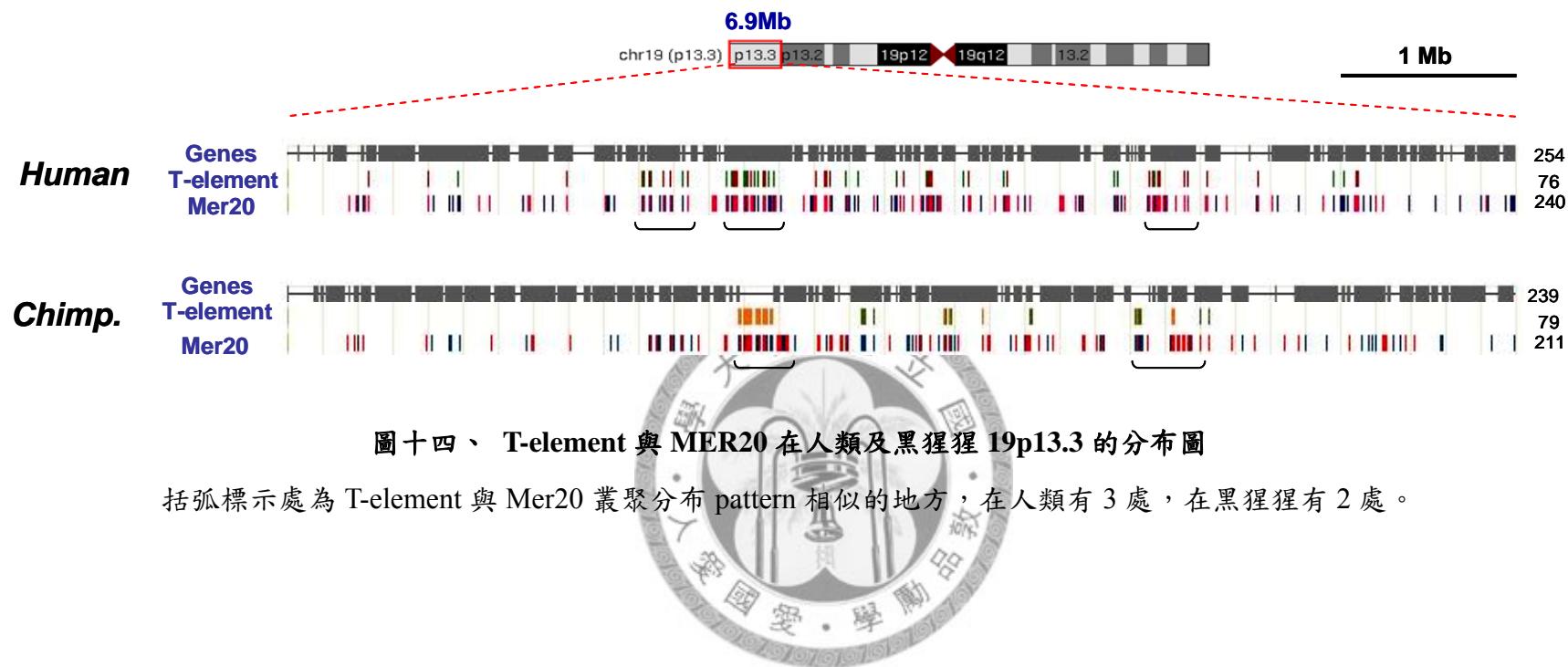


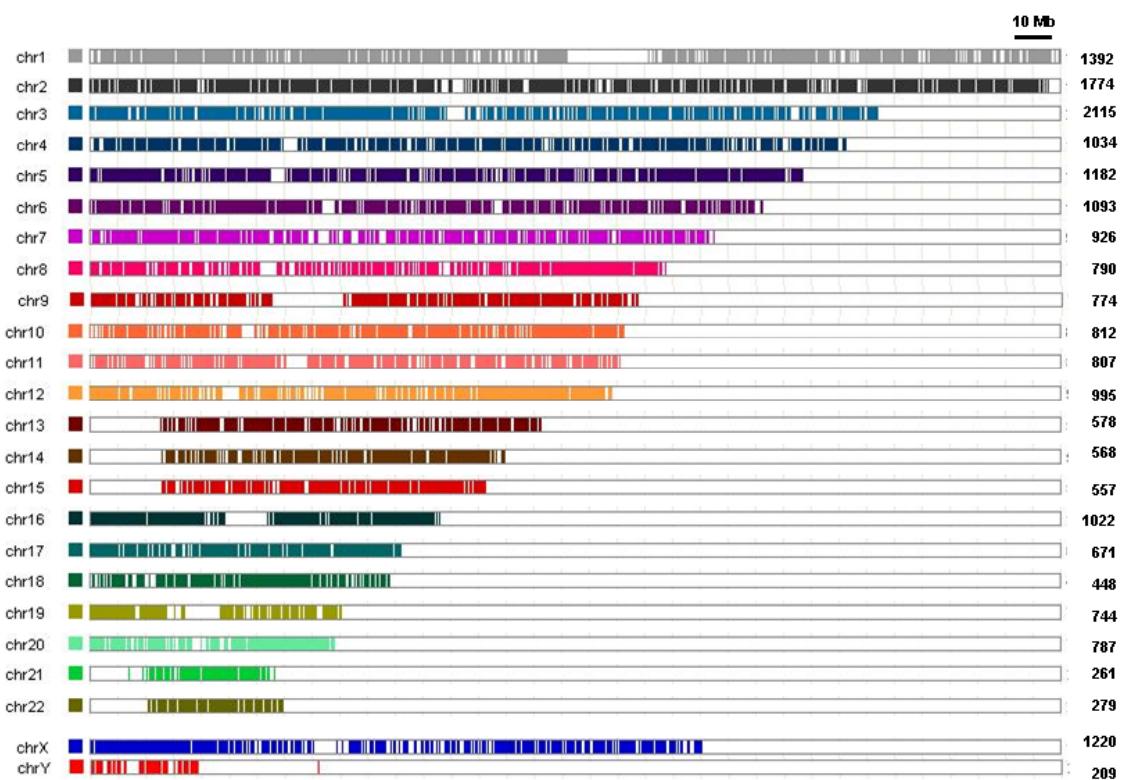
(C)



### 圖十三、T-element 之分析

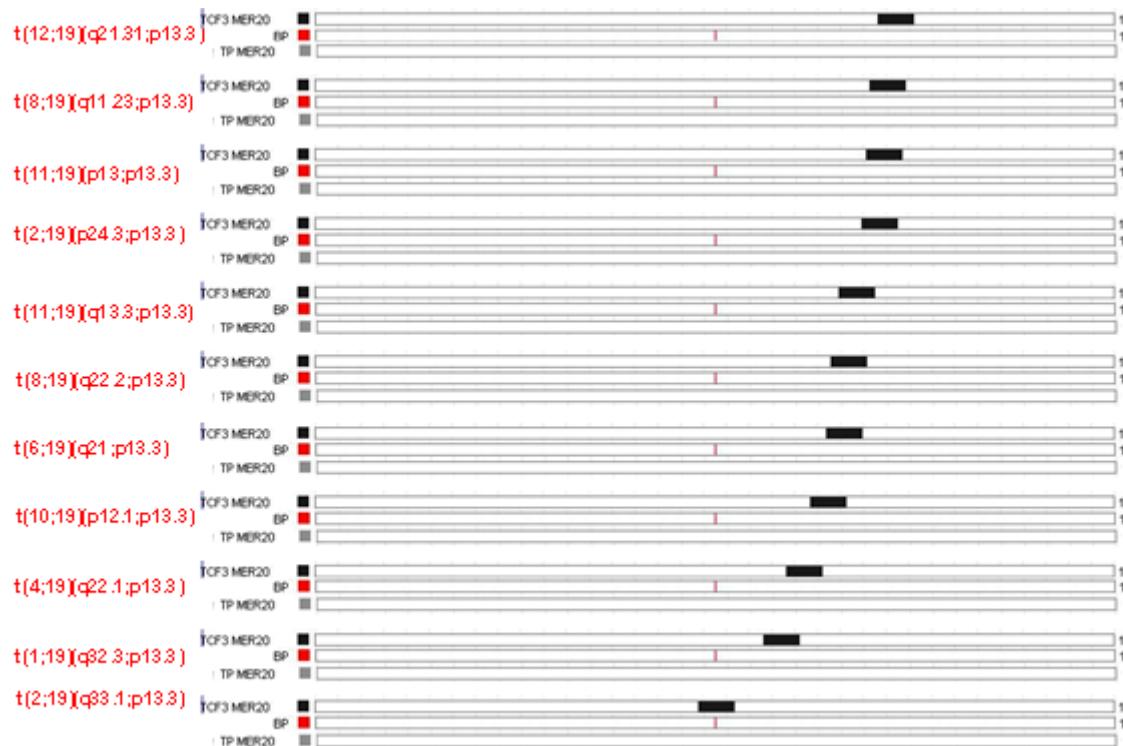
(A)T-element 在 *TCF3* 的座落處。(B)與 T-element 相似的序列(hits)在人類 19p13.3 的分布。(C)將 T-element 利用 Repbase 比對之結果。





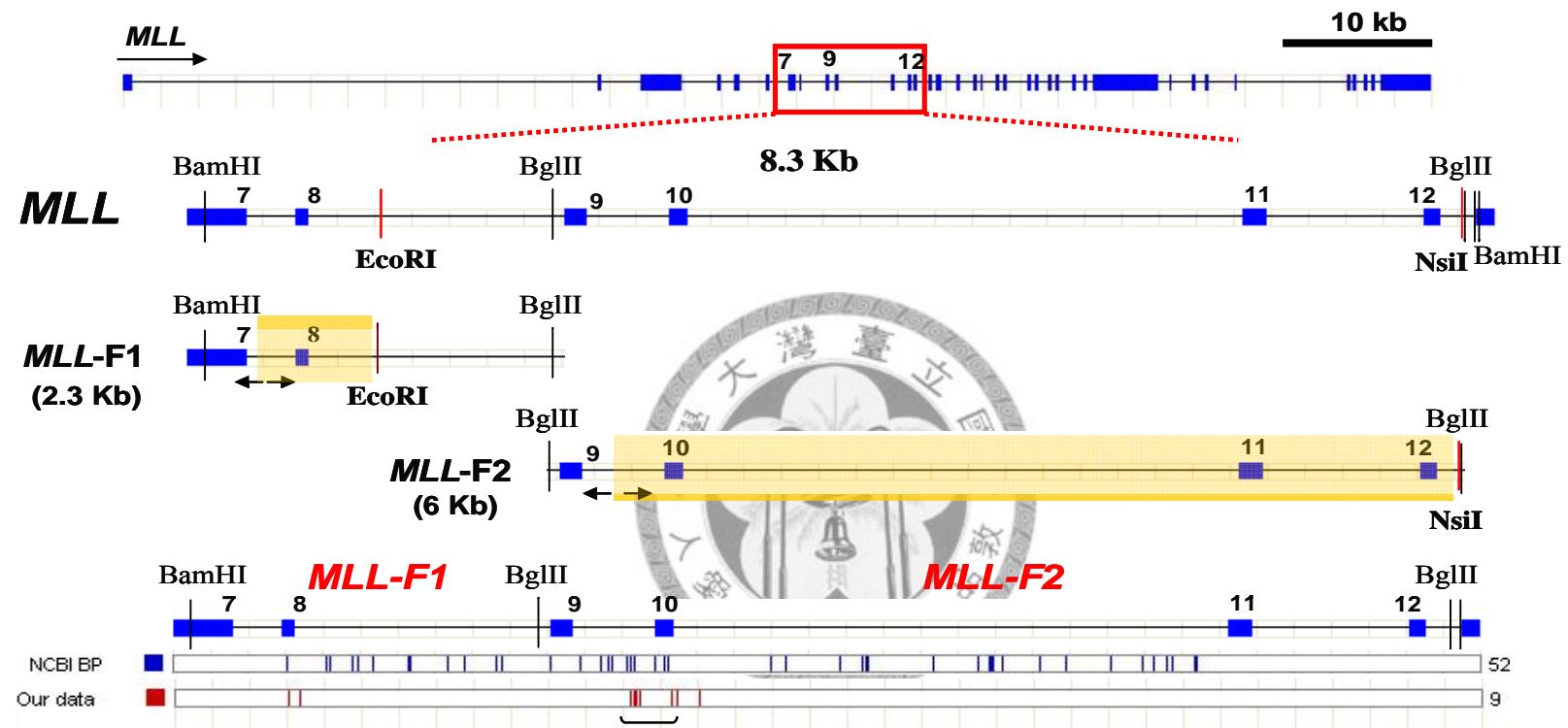
圖十五、MER20 在人類基因體中之分布圖





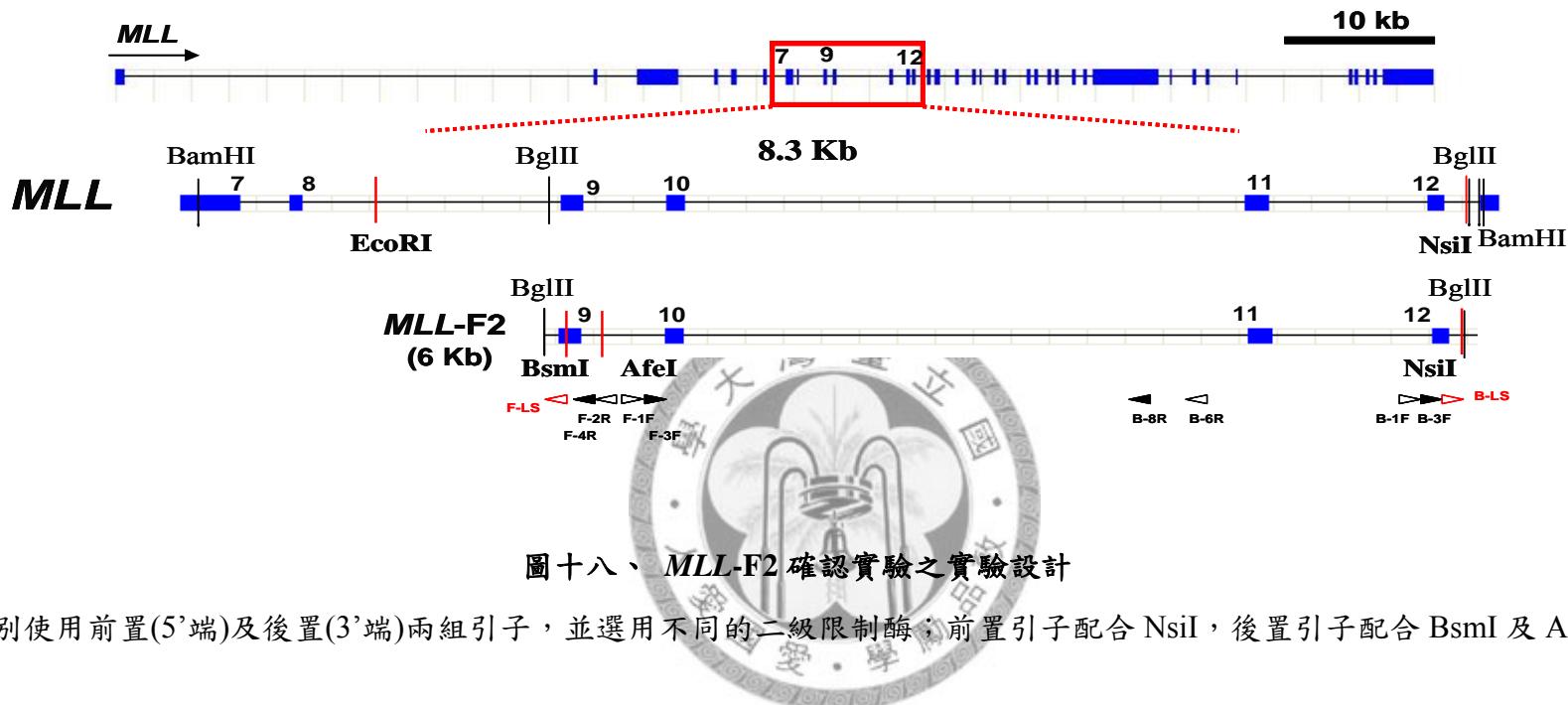
圖十六、 *TCF3-B* 之染色體轉位斷裂點周圍序列 MER20 分布

*TCF3-B*(5' 端)偵測出之染色體轉位斷裂點周邊 2 kb 序列 MER20 分佈。

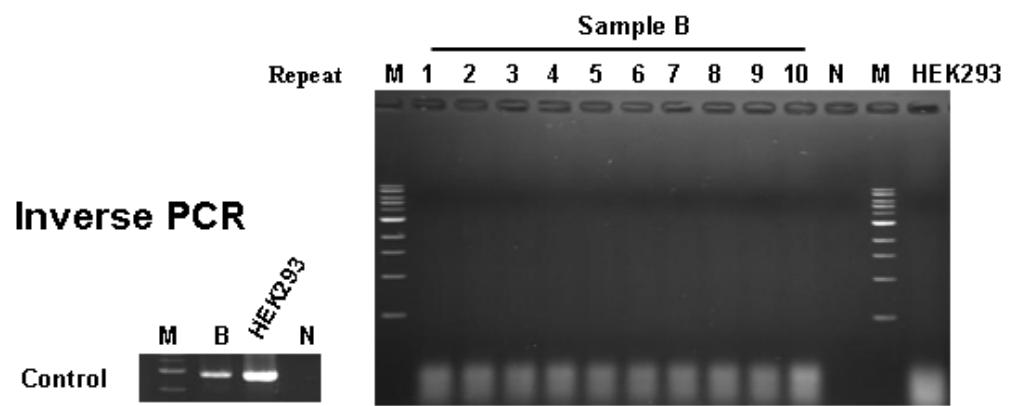


圖十七、在人類生殖細胞偵測到的 *MLL* 染色體轉位斷裂點分布位置

(實驗結果來自支若馨學姊)

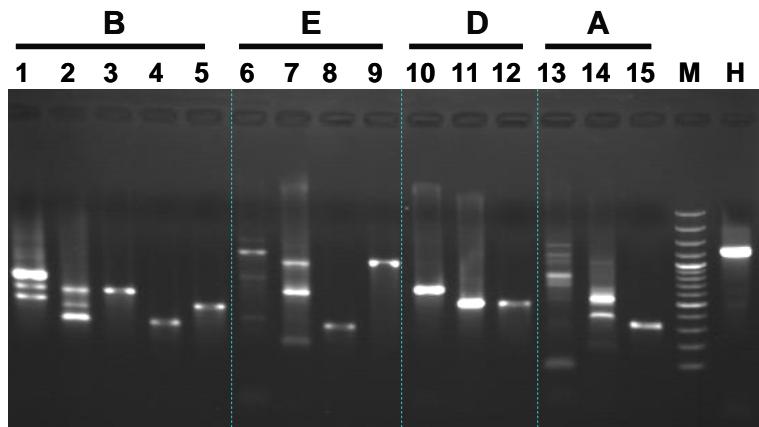


分別使用前置(5'端)及後置(3'端)兩組引子，並選用不同的二級限制酶；前置引子配合 NsiI，後置引子配合 BsmI 及 AfeI。



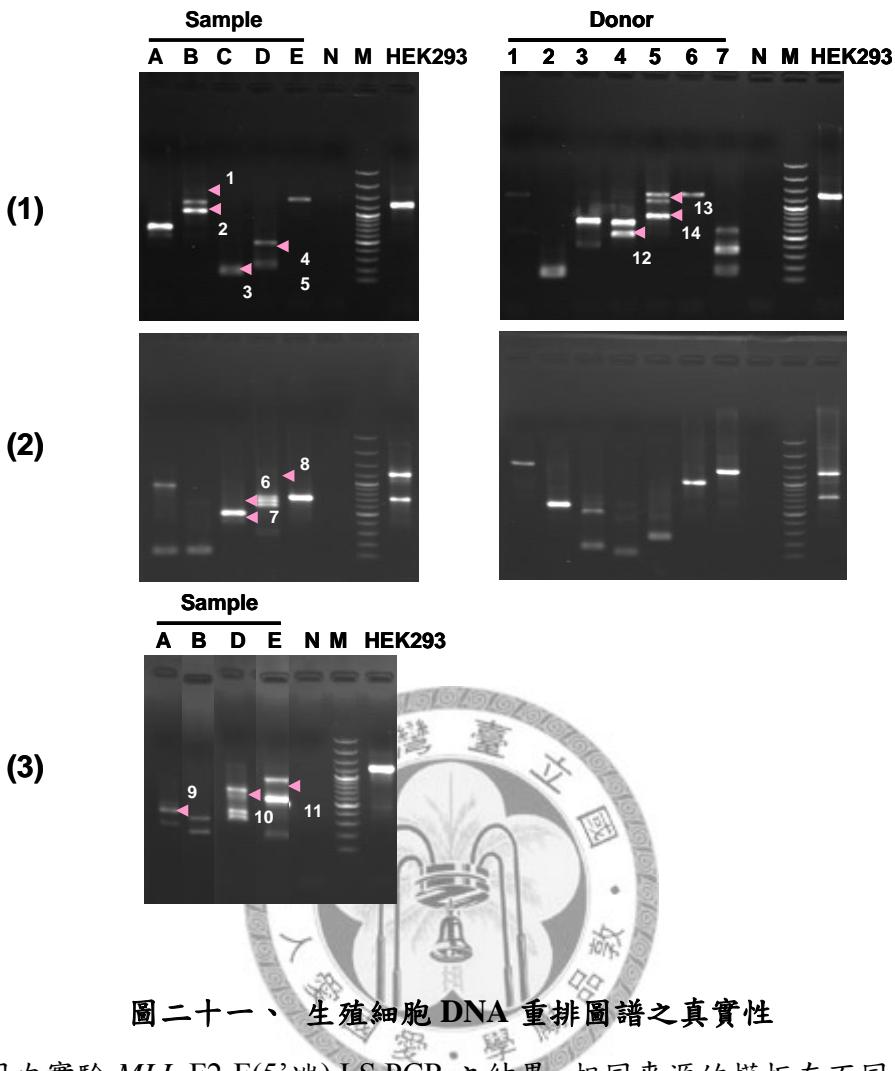
圖十九、*MLL* 染色體轉位的確認

以 Sample B 做十次重複，取 iPCR 產物當模板，分別針對 3 個先前由 *MLL*-F2 偵測到的染色體轉位進行確認。



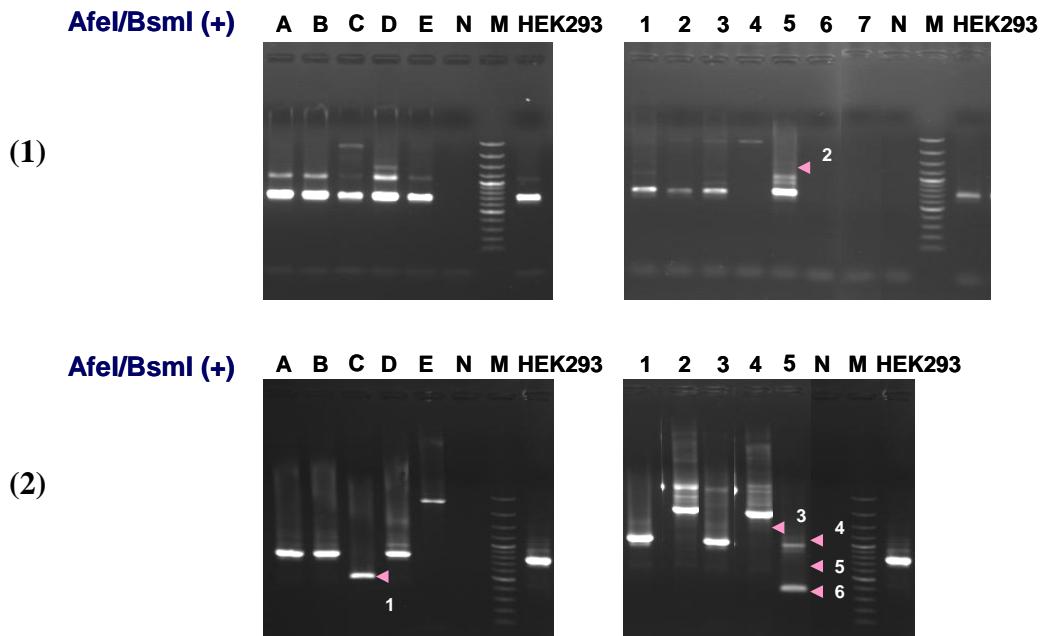
**圖二十、 生殖細胞 DNA 重排圖譜可信度之評估**

用先前偵測到 *MLL* 染色體轉位之 iPCR 產物當模板，lane 1, 6, 10, 13 為 nested PCR 產物，lane 2, 7, 11, 14 為 LS PCR 產物。lane 3, 4, 5 是以帶有由 sample B 偵測出的染色體轉位序列之質體為模板之 LS PCR 產物。lane 8, 9 是以帶有由 sample E 偵測出的染色體轉位序列之質體當模板之 LS PCR 產物。lane 12 是以帶有由 sample D 偵測出的染色體轉位序列之質體當模板之 LS PCR 產物。lane 15 是以帶有由 sample A 偵測出的染色體轉位序列之質體當模板之 LS PCR 產物。LS PCR pattern 在 sample B, E, D 均能與其 positive control 相對應，尤其是在 Sample B，其 LS PCR pattern 完全可反應出 DNA 重排的圖譜。



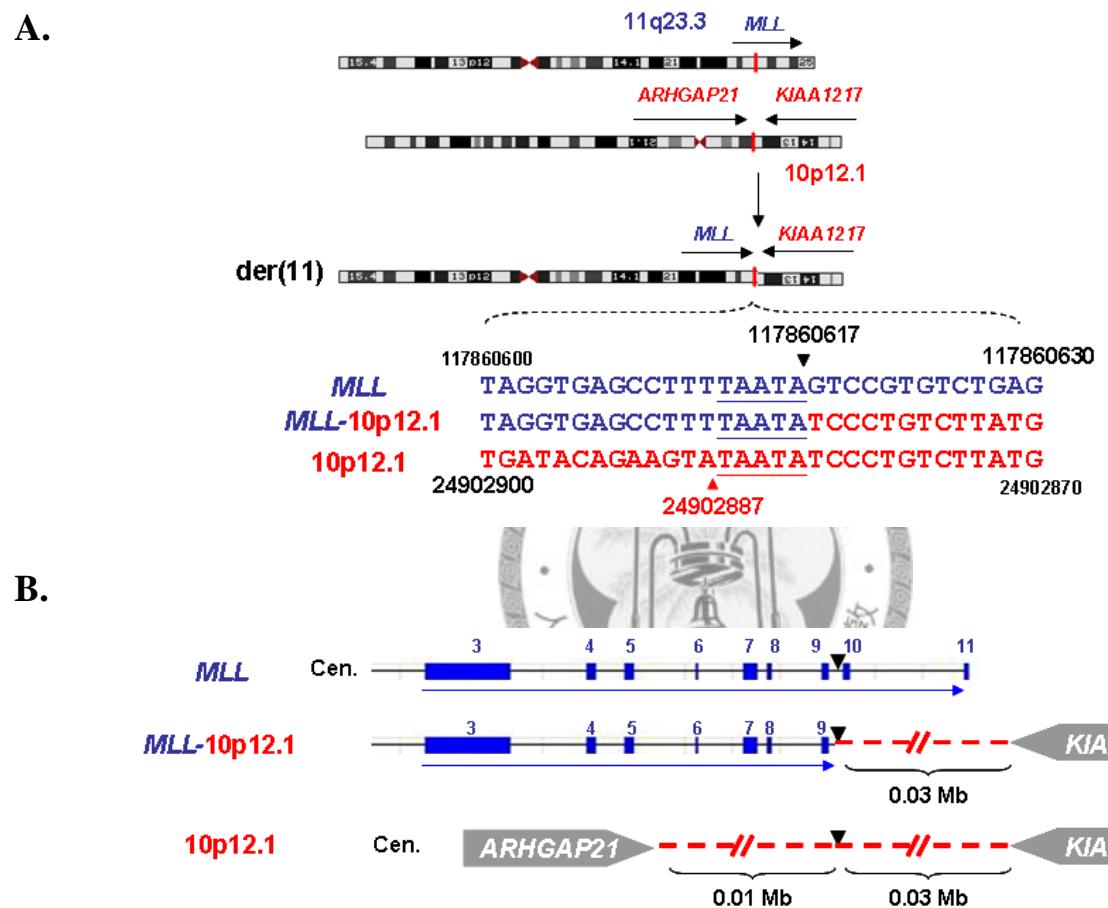
圖二十一、生殖細胞 DNA 重排圖譜之真實性

圖為不同次實驗 *MLL-F2-F(5'端)* LS PCR 之結果，相同來源的模板在不同次實驗顯現出的 pattern 不同。三角形標示處為經選殖定序後偵測到的染色體轉位，共有 14 個。分別如下：1 為  $t(1;11)(q21.1;q23.3)$ ，PCR 產物大小為 1578 bp；2 為  $t(10;11)(q26.2;q23.3)$ ，PCR 產物大小為 1007 bp；3 為  $t(10;11)(p12.1;q23.3)$ ，PCR 產物大小為 153 bp；4 為  $t(3;11)(p22.1;q23.3)$ ，PCR 產物大小為 442 bp；5 為  $t(3;11)(q23;q23.3)$ ，PCR 產物大小為 442 bp；6 為  $t(9;11)(p24.3;q23.3)$ ，PCR 產物大小為 702 bp；7 為  $t(4;11)(q25;q23.3)$ ，PCR 產物大小為 447 bp；8 為  $t(11;15)(q23.3;q22.2)$ ，PCR 產物大小為 1366 bp；9 為  $t(4;11)(q35.1;q23.3)$ ，PCR 產物大小為 527 bp；10 為  $t(5;11)(p12;q23.3)$ ，PCR 產物大小為 679 bp；11 為  $t(1;11)(p32.1;q23.3)$ ，PCR 產物大小為 967 bp；12 為  $t(11;18)(q23.3;q12.2)$ ，PCR 產物大小為 492 bp；13 為  $t(11;14)(q23.3;q31.1)$ ，PCR 產物大小為 1065 bp；14 為  $t(3;11)(q22.3;q23.3)$ ，PCR 產物大小為 769 bp。

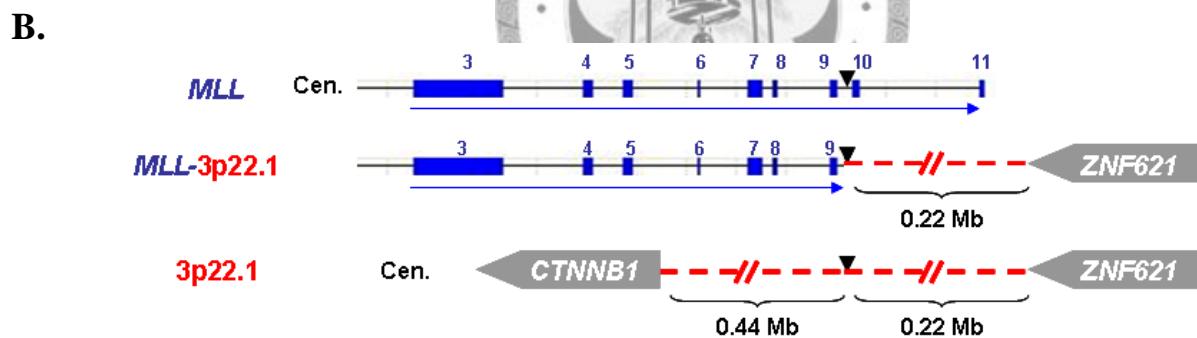
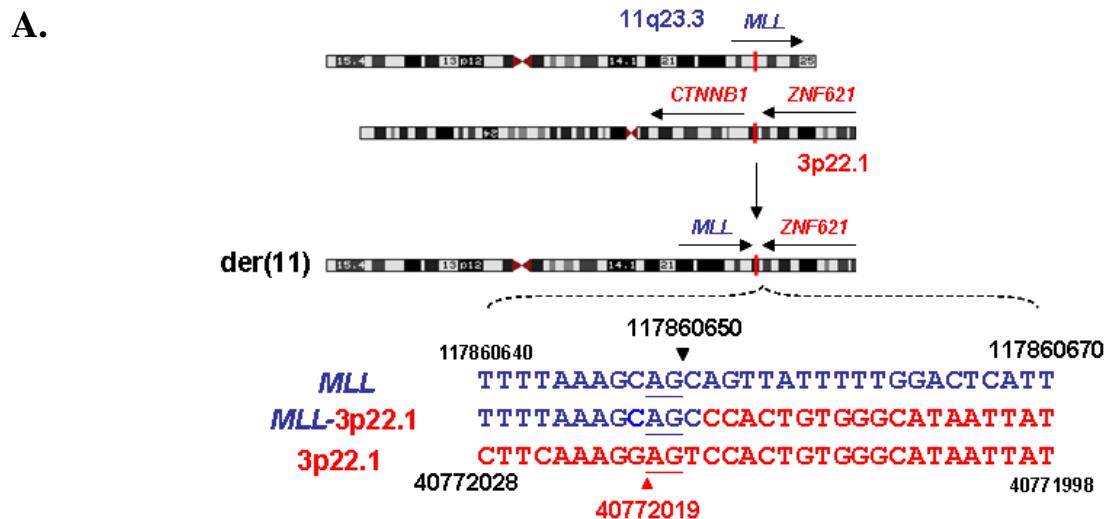


圖二十二、*MLL-F2-B* 之 PCR 電泳結果

圖為不同次實驗 *MLL-F2-B*(3' 端) LS PCR 之結果，相同來源的模板在不同次實驗顯現出的 pattern 不同。三角形標示處為經選殖定序後偵測到的染色體轉位，共有 6 個。1 為 t(2;11)(p24.3;q23.3)，PCR 產物大小為 441 bp；2 為 t(4;11)(p12;q23.3)，PCR 產物大小為 1039 bp；3 為 t(11;18)(q23.3;q12.2)，PCR 產物大小為 1450 bp；4 為 t(4;11)(q26;q23.3)，PCR 產物大小為 1051 bp；5 為 t(11;14)(q23.3;q31.1)，PCR 產物大小為 744 bp；6 為 t(11;16)(q23.3;q21)，PCR 產物大小為 257 bp。

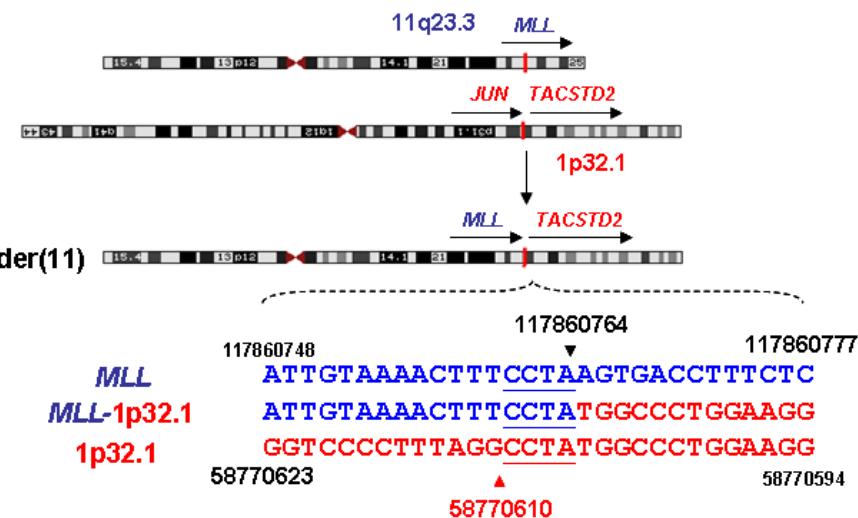


圖二十三之一、*MLL-F*(5'端)所偵測出的染色體轉位 *t*(10;11)(p12.1;q23.3)

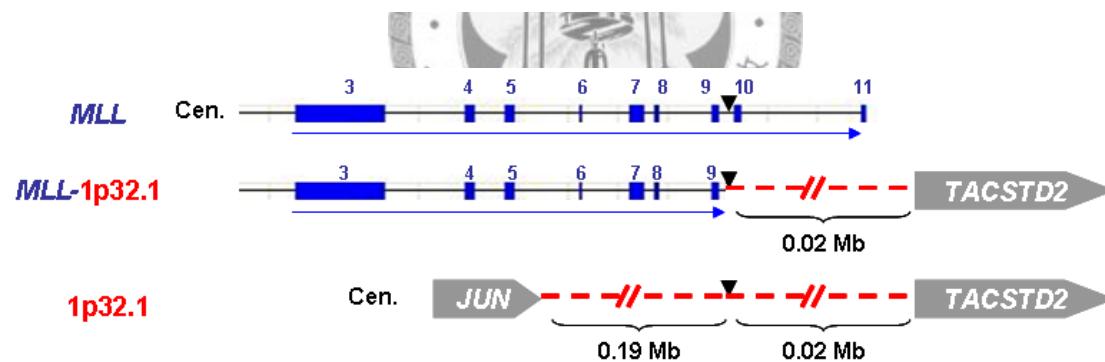


圖二十三之二、*MLL-F(5'端)*所偵測出的染色體轉位  $t(3;11)(p22.1;q23.3)$

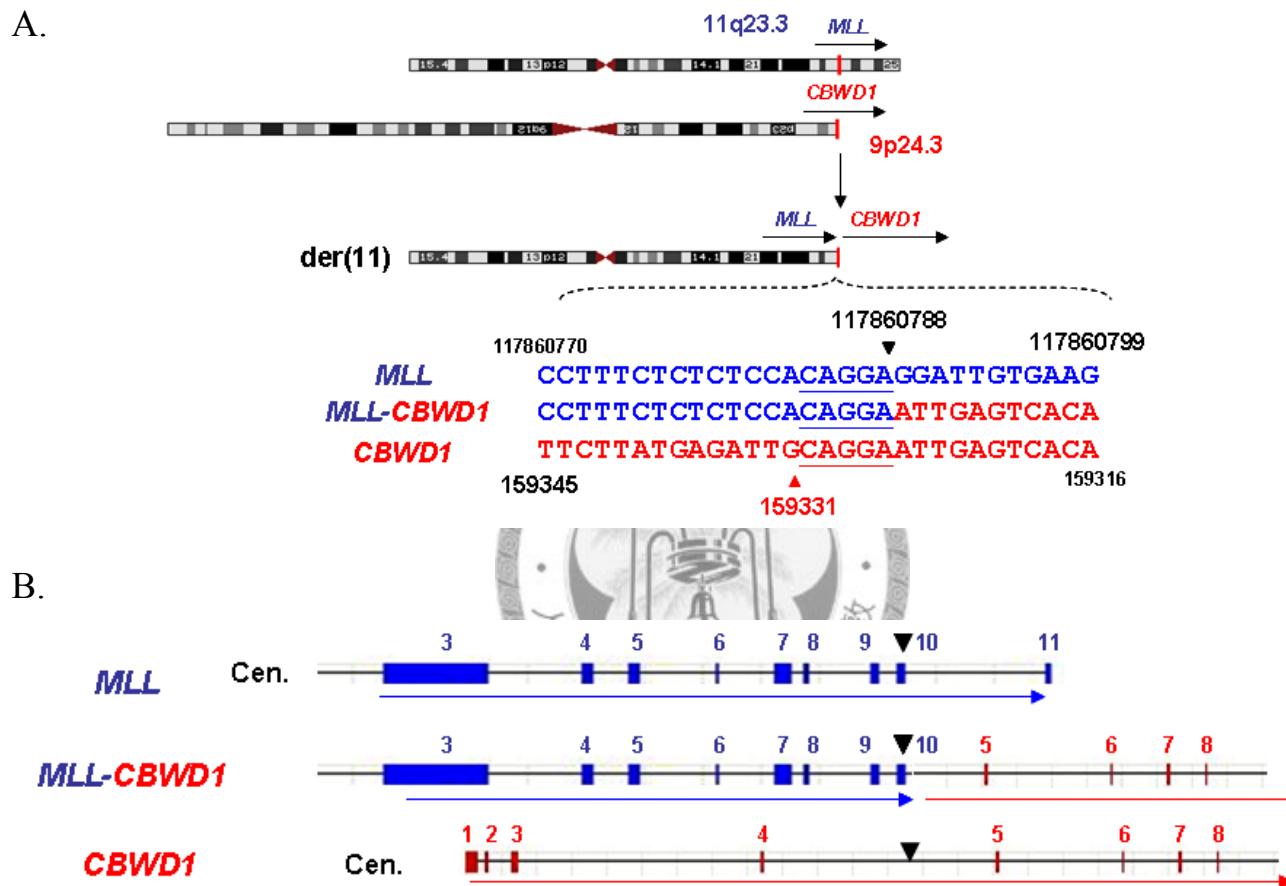
A.



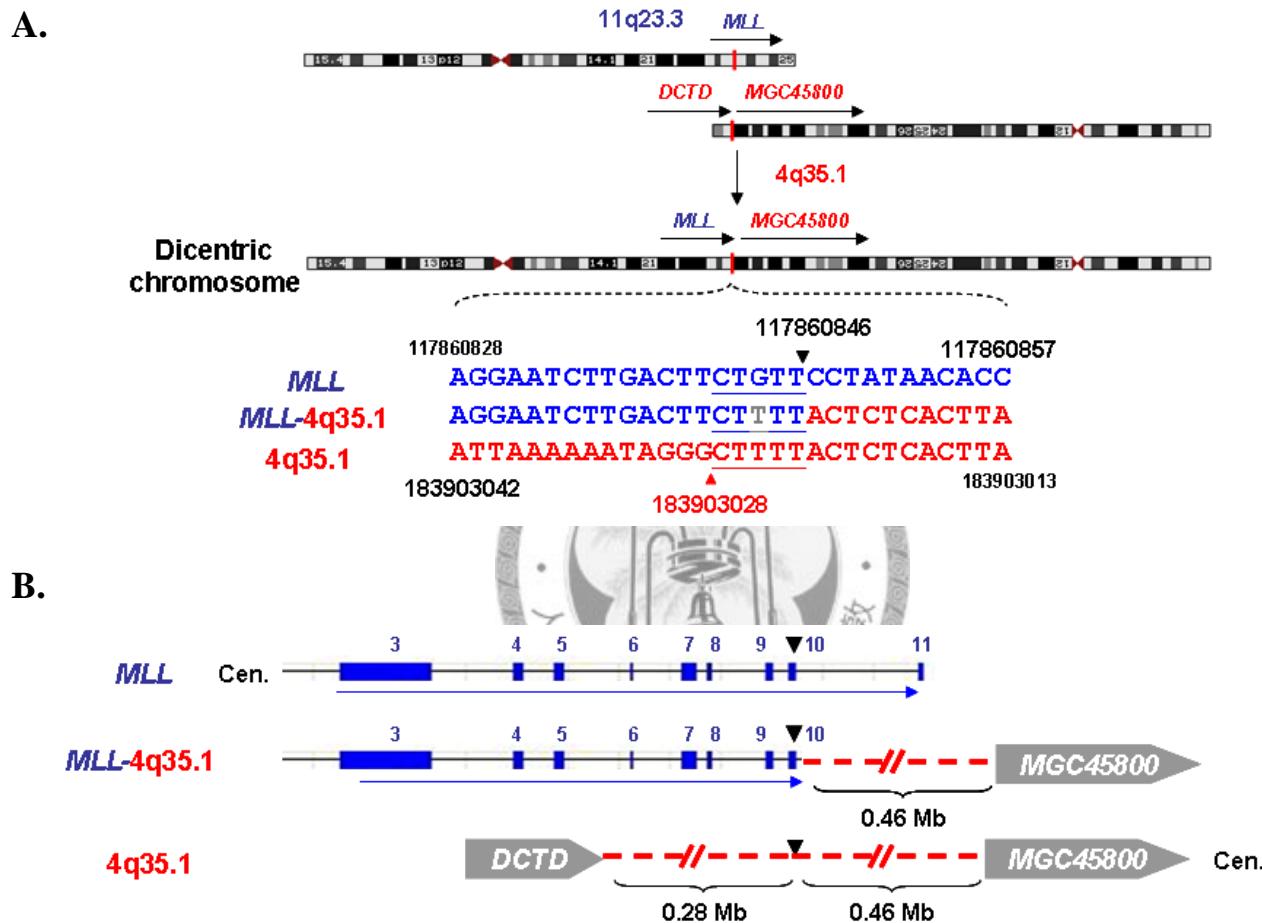
B.



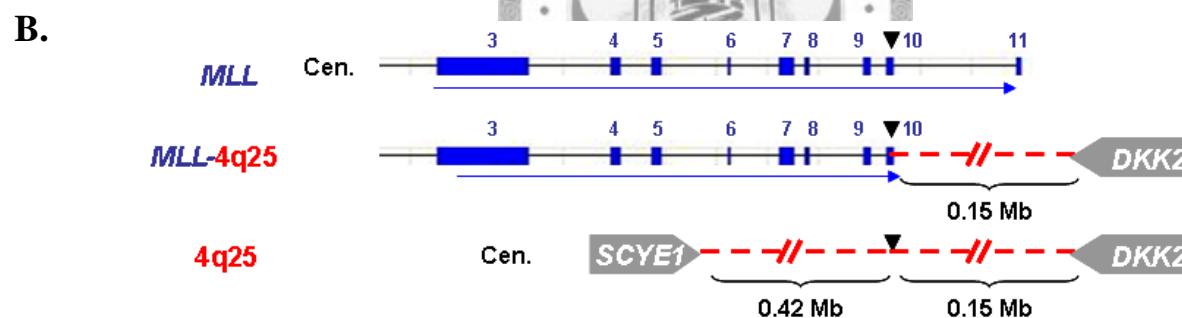
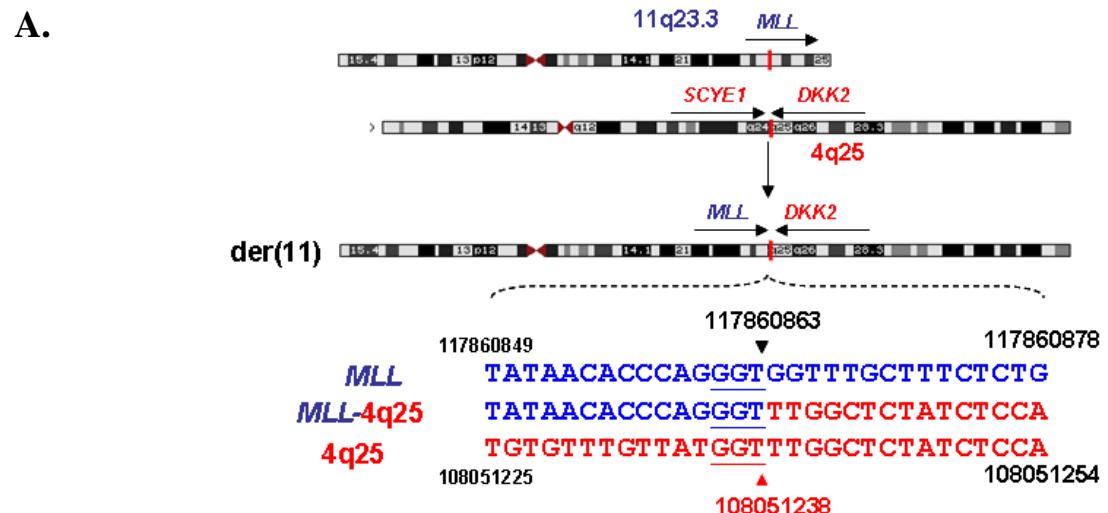
圖二十三之三、*MLL*-F(5'端)所偵測出的染色體轉位 *t*(1;11)(p32.1;q23.3)



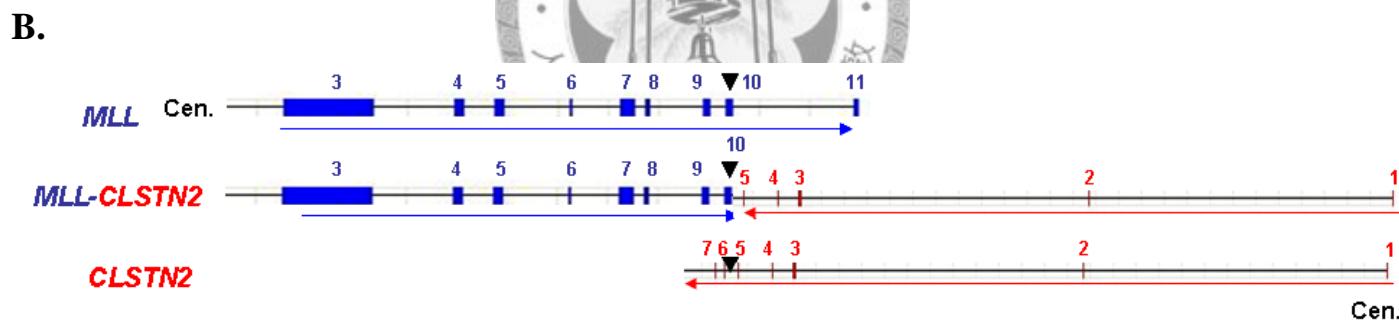
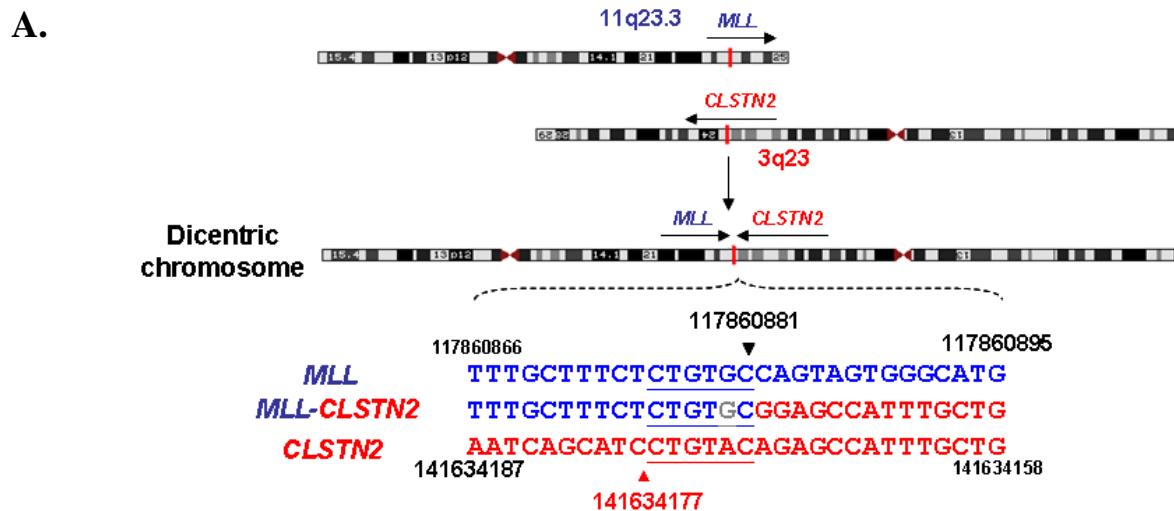
圖二十三之四、*MLL*-F(5'端)所偵測出的染色體轉位t(9;11)(p24.3;q23.3)



圖二十三之五、*MLL-F(5'端)*所偵測出的染色體轉位 t(4;11)(q35.1;q23.3)

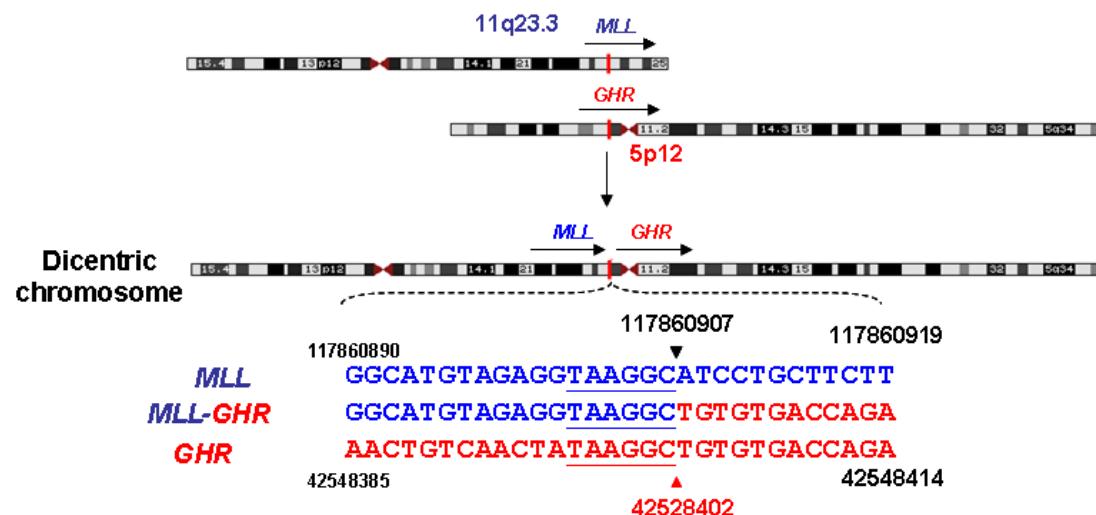


圖二十三之六、*MLL-F(5'端)*所偵測出的染色體轉位 t(4;11)(q25;q23.3)

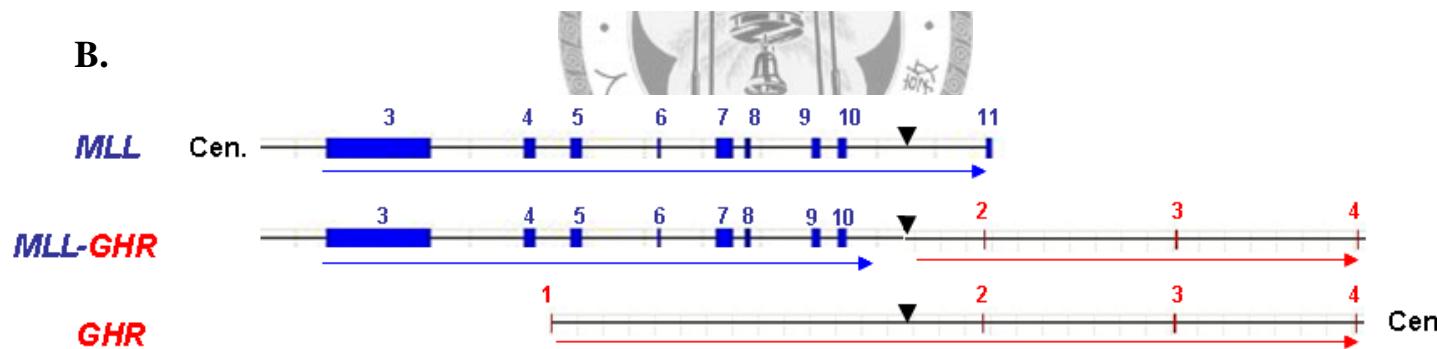


圖二十三之七、 *MLL*-F(5'端)所偵測出的染色體轉位 t(3;11)(q23;q23.3)

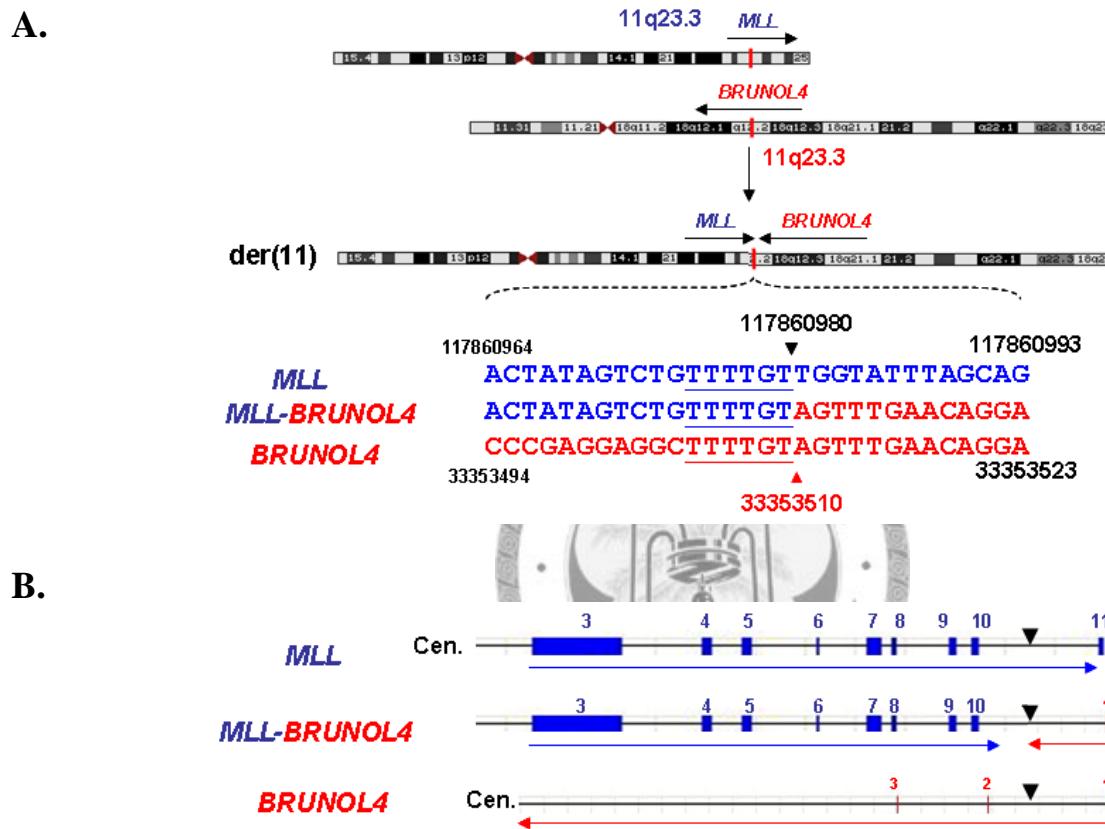
A.



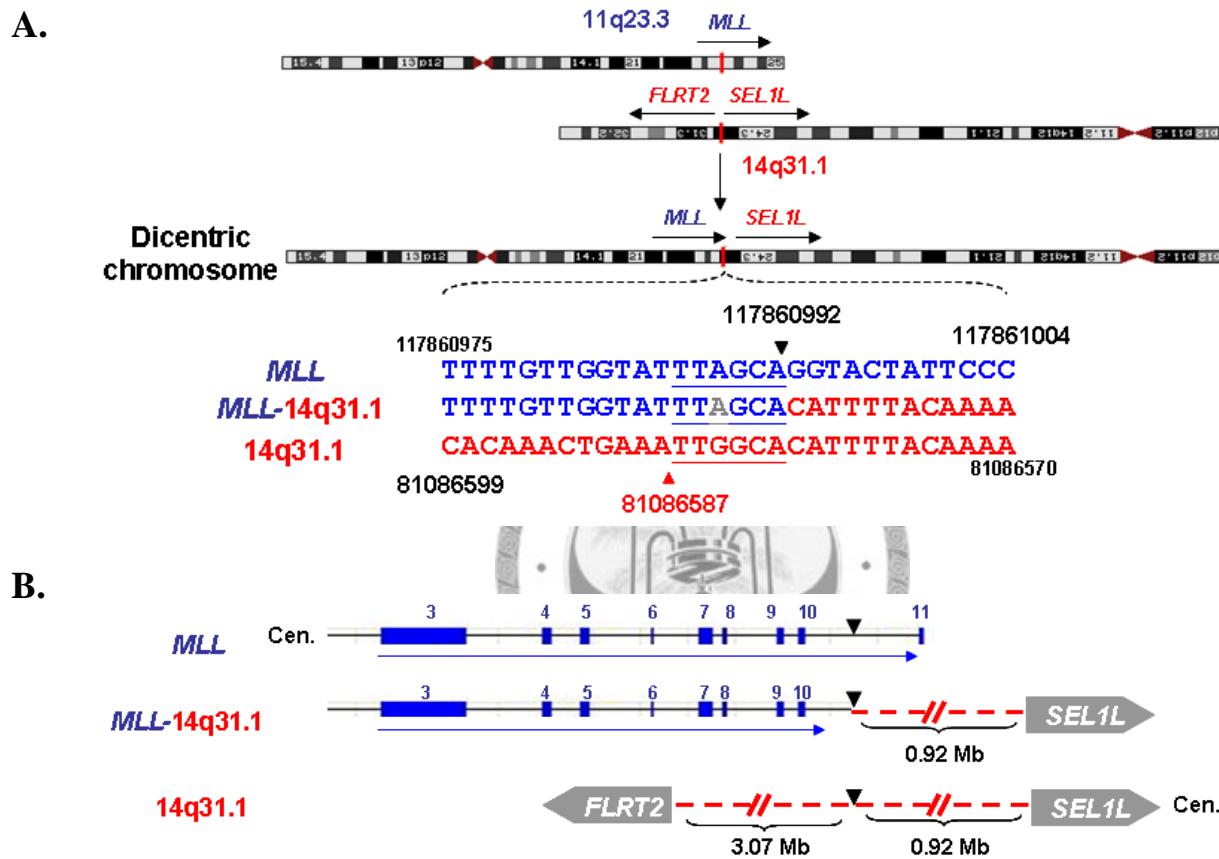
B.



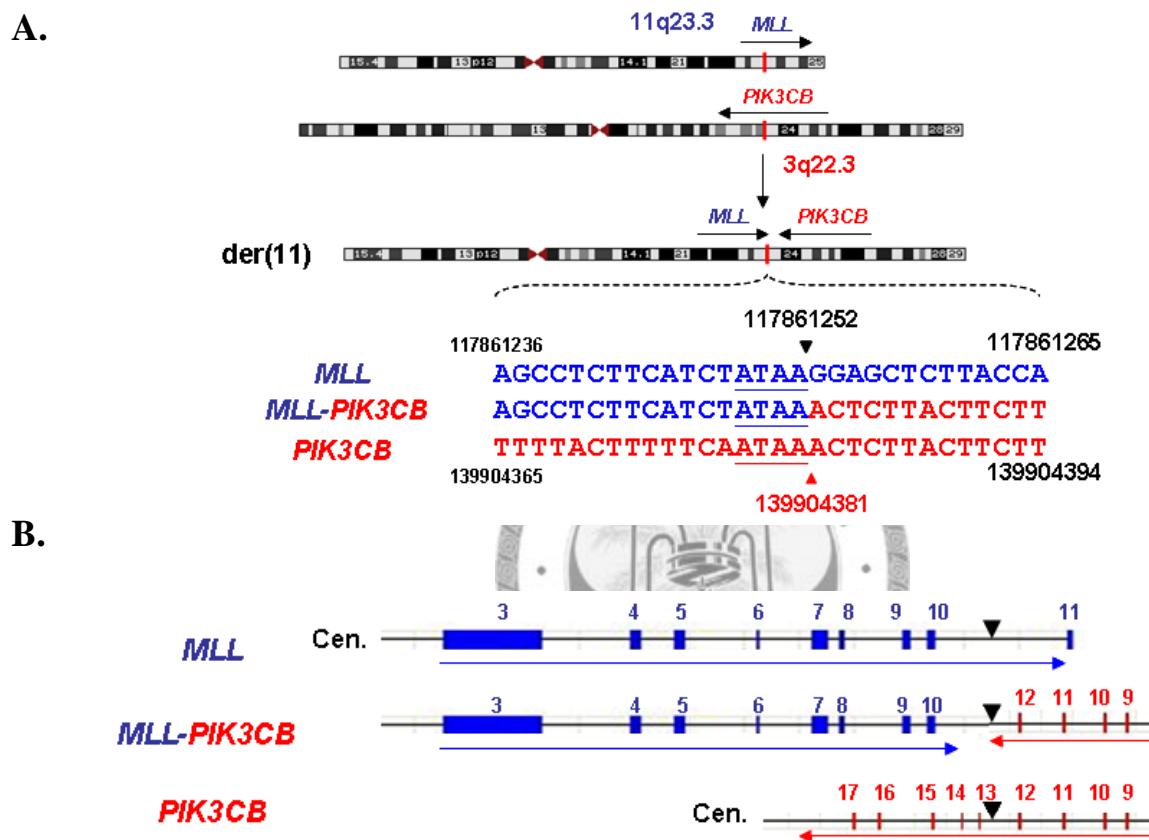
圖二十三之八、 *MLL-F(5'端)*所偵測出的染色體轉位  $t(5;11)(p12;q23.3)$



圖二十三之九、*MLL-F(5'端)*所偵測出的染色體轉位 t(11;18)(q23.3;q12.2)

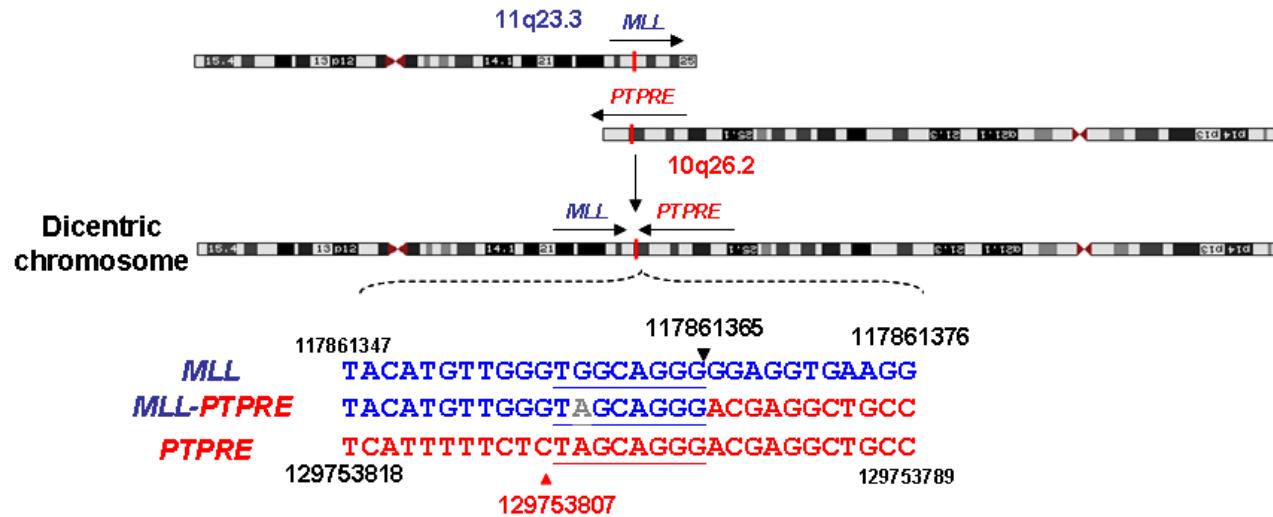


圖二十三之十、 *MLL-F*(5'端)所偵測出的染色體轉位 t(11;14)(q23.3;q31.1)

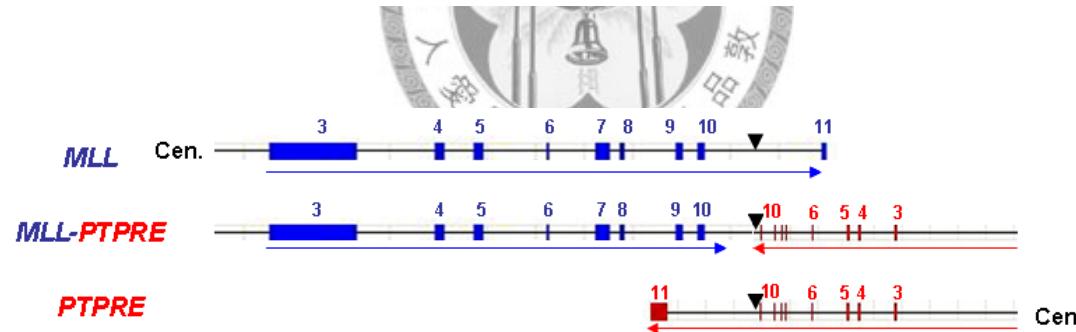


圖二十三之十一、 *MLL-F(5'端)*所偵測出的染色體轉位  $t(3;11)(q22.3;q23.3)$

A.

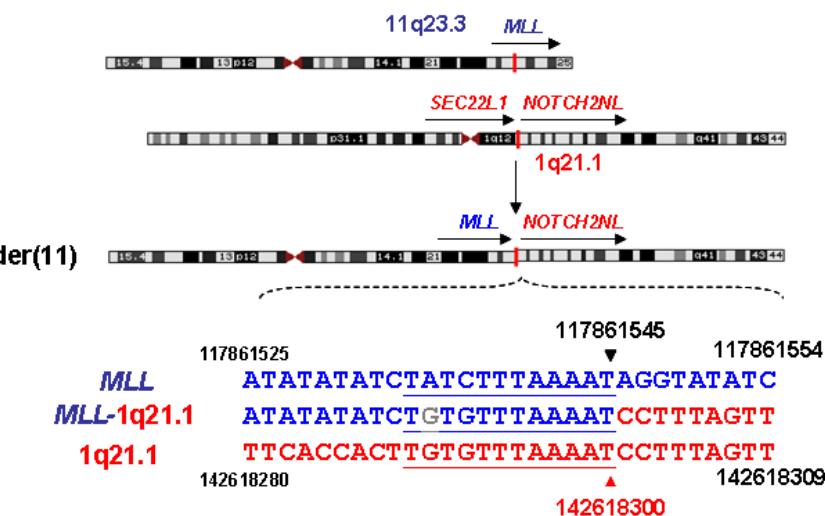


B.

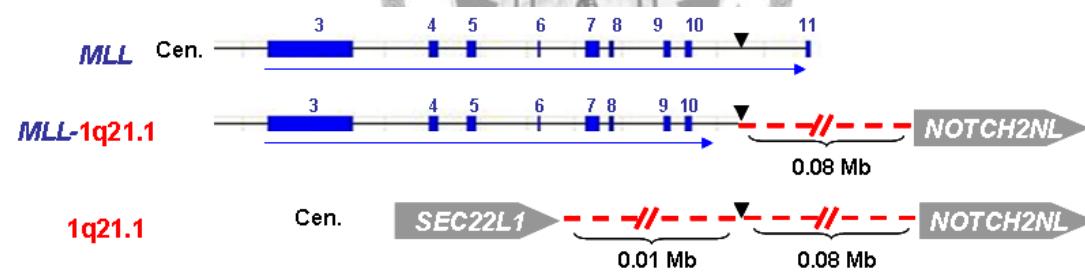


圖二十三之十二、 *MLL-F(5'端)*所偵測出的染色體轉位  $t(10;11)(q26.2;q23.3)$

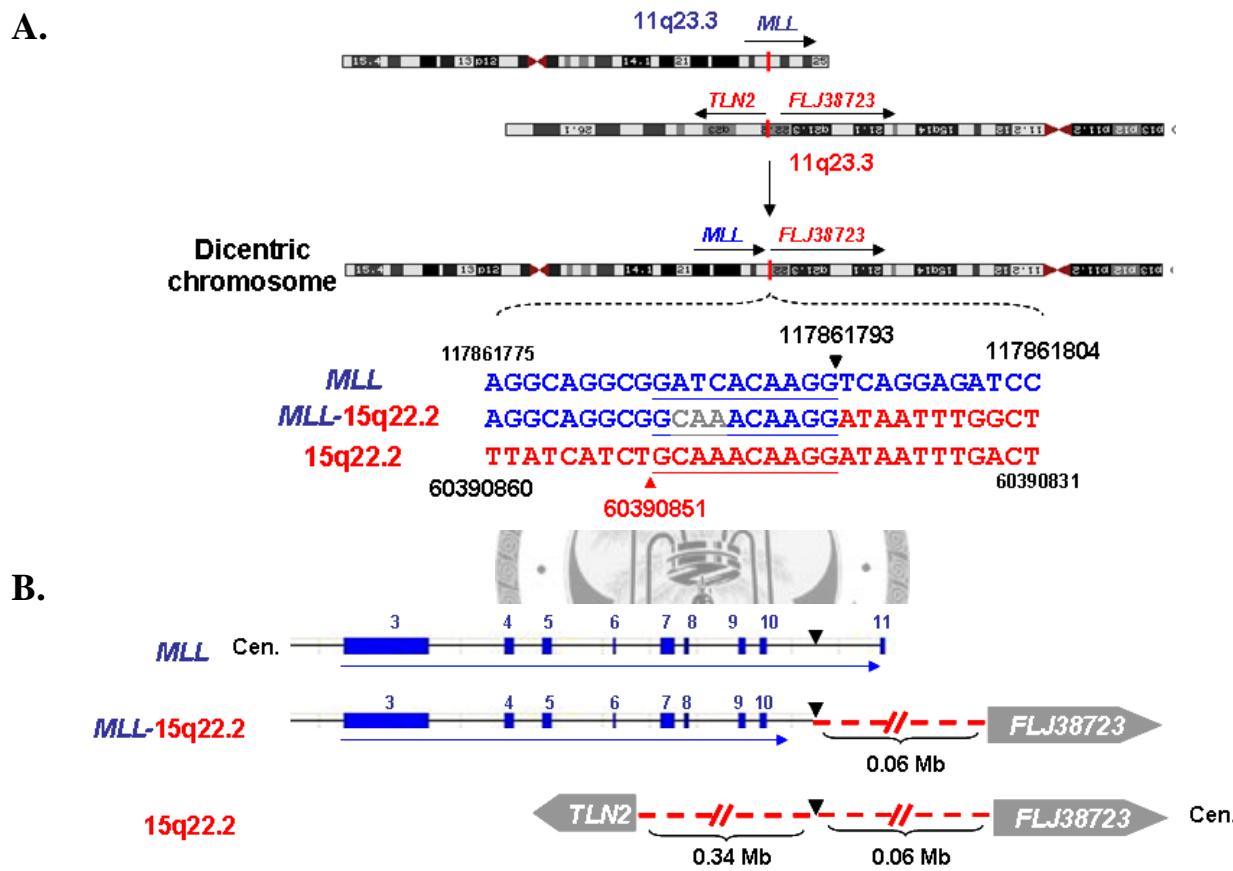
A.



B.

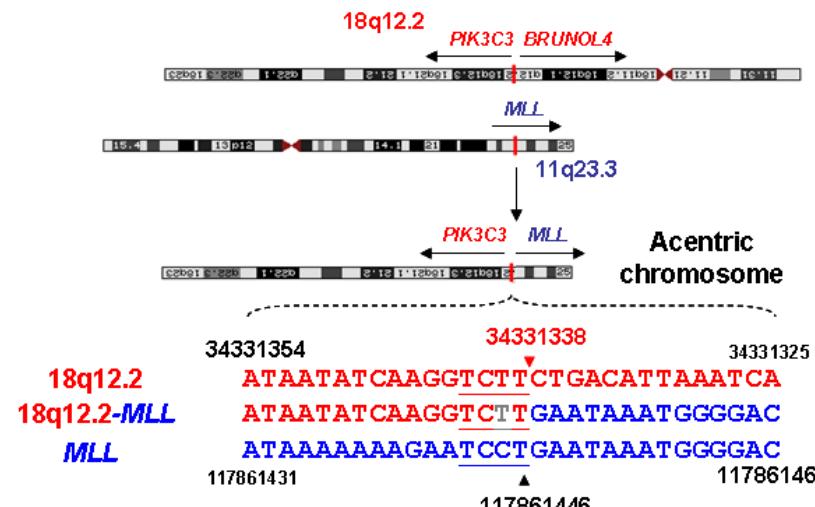


圖二十三之十三、*MLL*-F(5'端)所偵測出的染色體轉位  $t(1;11)(q21.1;q23.3)$

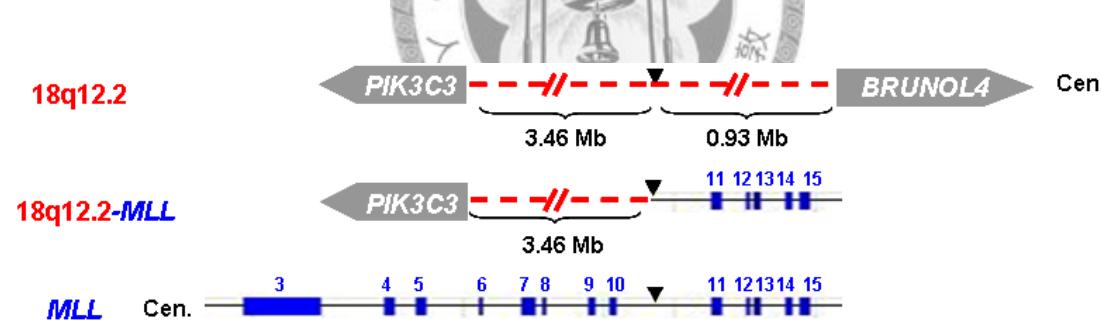


圖二十三之十四、*MLL-F(5'端)*所偵測出的染色體轉位 *t(11;15)(q23.3;q22.2)*

A.

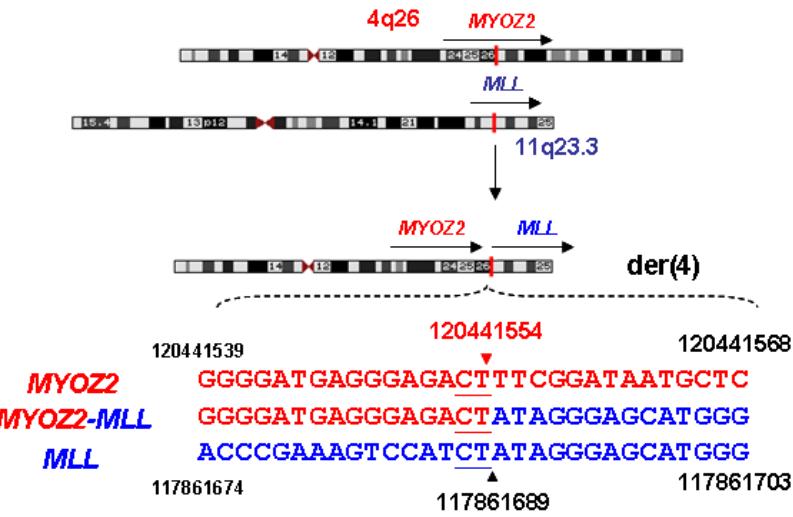


B.

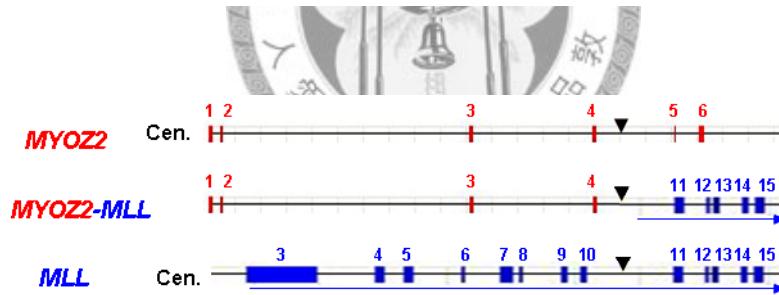


圖二十四之一、*MLL-B*(3'端)所偵測出的染色體轉位  $t(11;18)(q23.3;q12.2)$

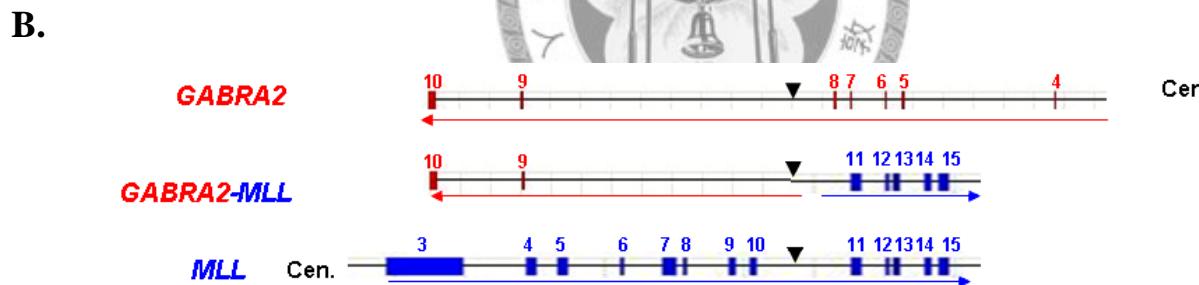
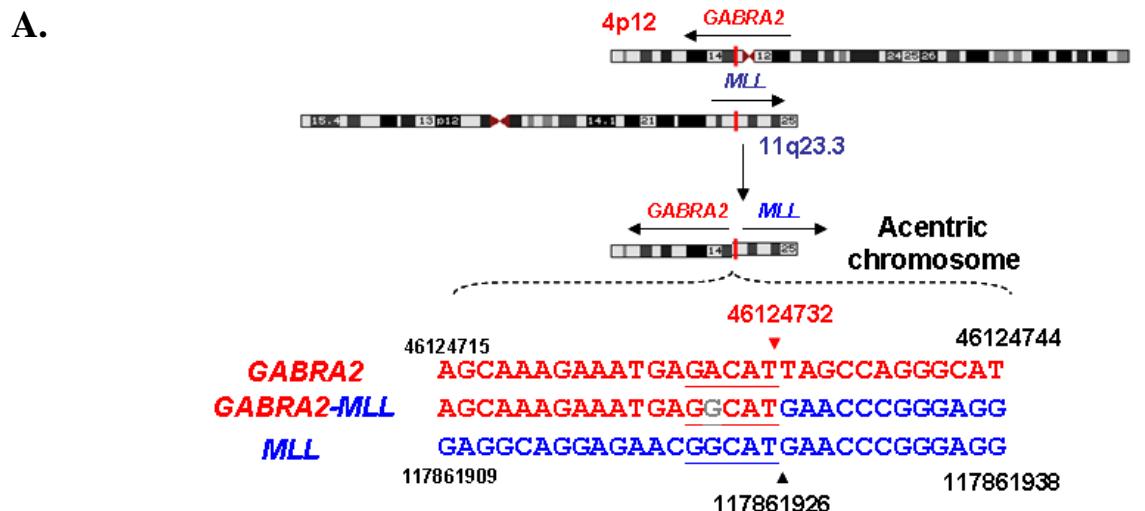
A.



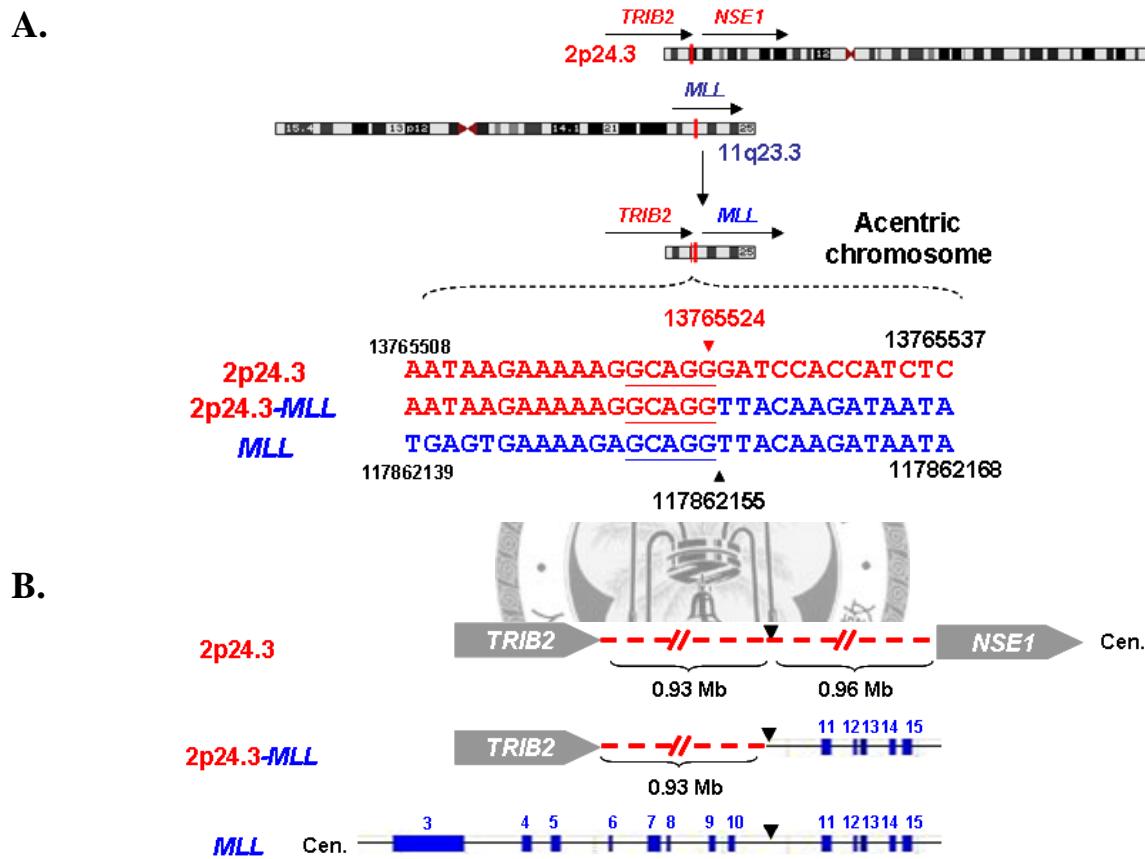
B.



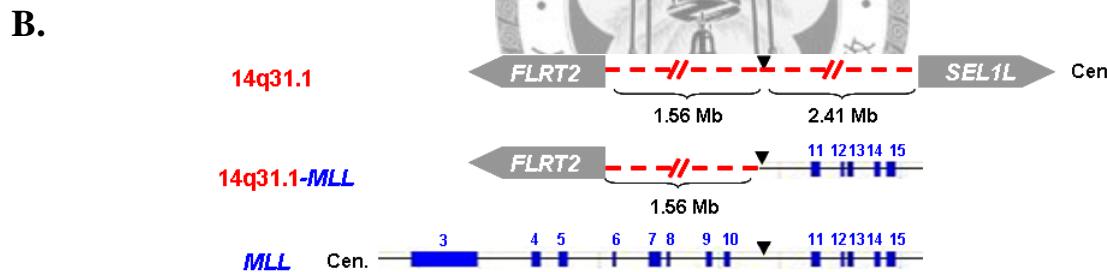
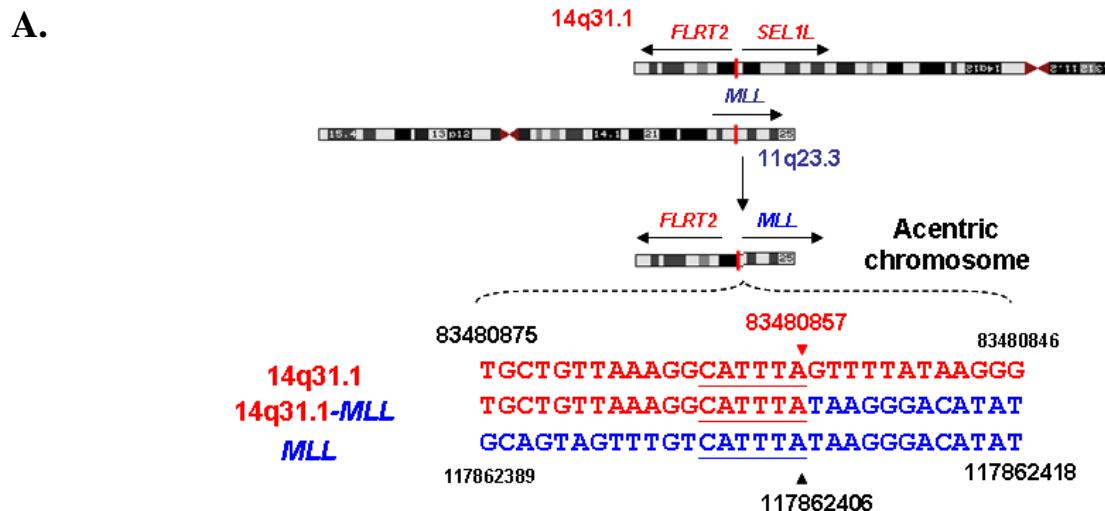
圖二十四之二、*MLL-B*(3'端)所偵測出的染色體轉位 *t(4;11)(q26;q23.3)*



圖二十四之三、*MLL-B(3'端)*所偵測出的染色體轉位  $t(4;11)(p12;q23.3)$

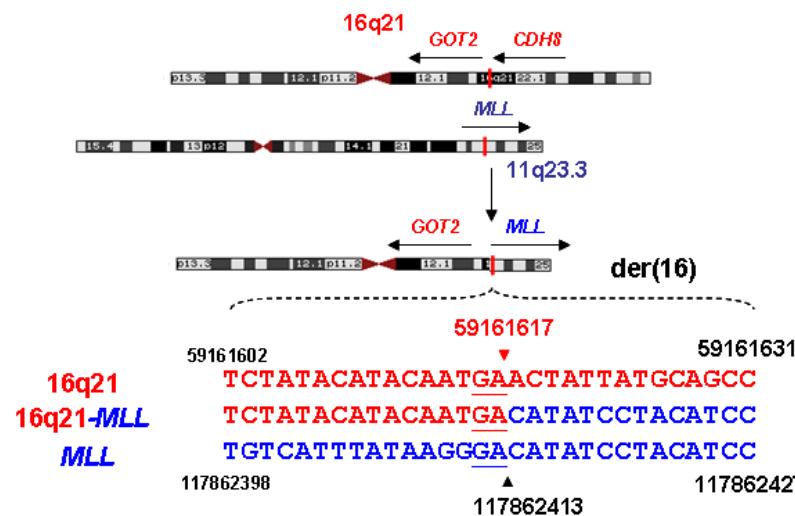


圖二十四之四、*MLL-B*(3'端)所偵測出的染色體轉位  $t(2;11)(p24.3;q23.3)$

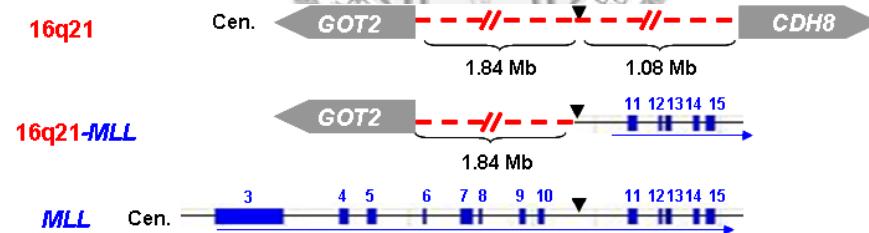


圖二十四之五、*MLL-B(3'端)*所偵測出的染色體轉位  $t(11;14)(q23.3;q31.1)$

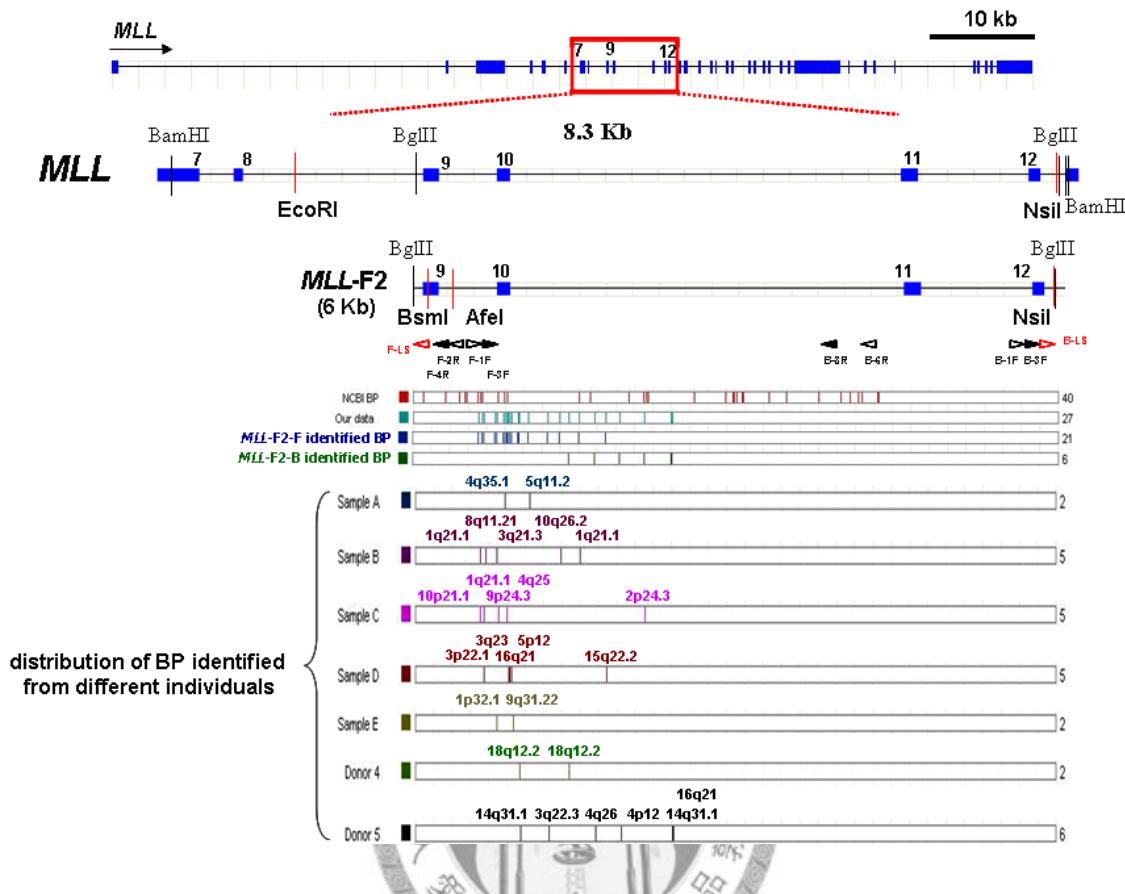
A.



B.

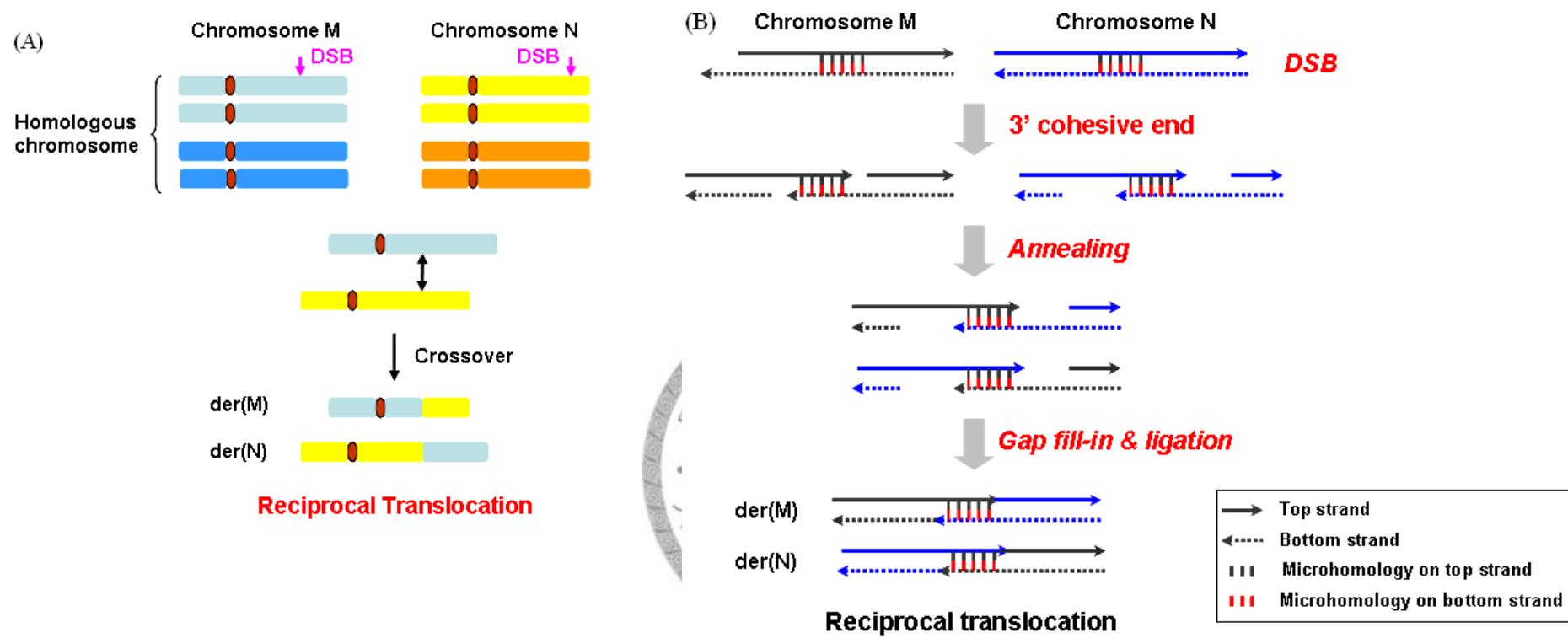


圖二十四之六、*MLL-B*(3'端)所偵測出的染色體轉位  $t(11;16)(q23.3;q21)$



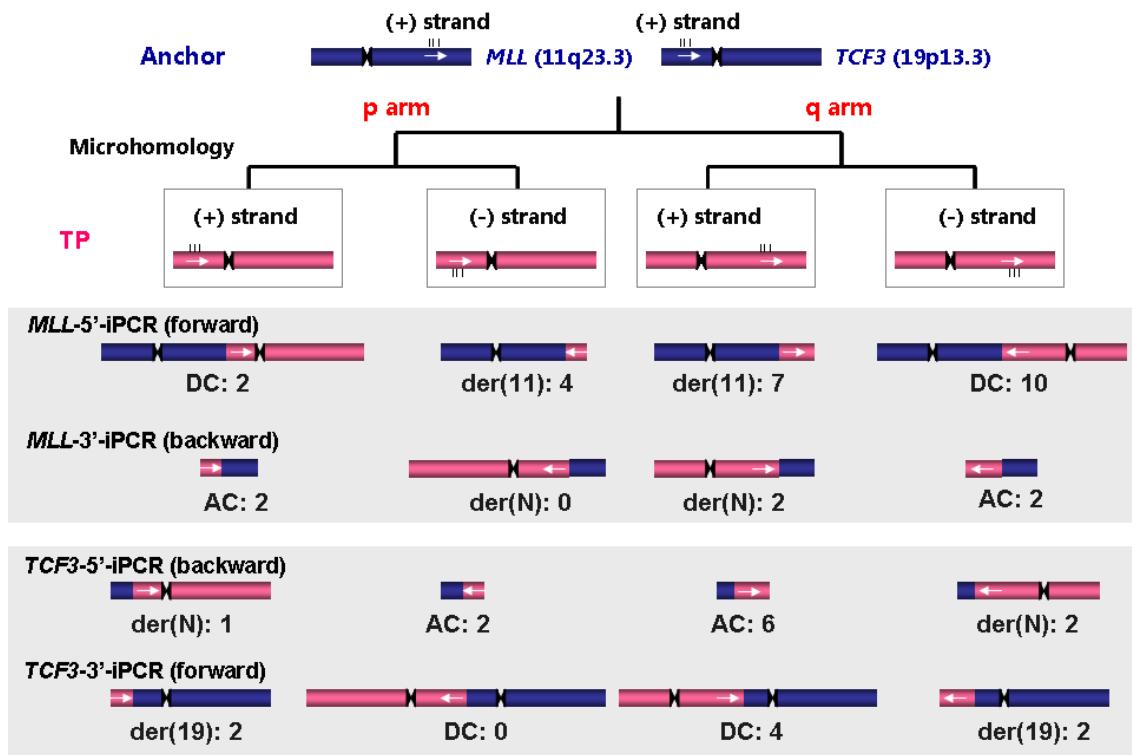
圖二十五、*MLL-F2* 之染色體轉位斷裂點分布

由前置引子(5'端)及後置引子(3'端)偵測出的染色體轉位斷裂點，在 *MLL* intron 9 至 intron 10 呈現密集分佈，且分布狀況與文獻所發表之斷裂點相似。箭頭方向表示 *MLL* 的轉錄方向。由不同來源精細胞 gDNA(between individuals)偵測出的染色體轉位，在 *TCF3* 上的斷裂點不同；而斷裂位置在個體本身(within individual)也有差異，且轉位對象亦不同。



圖二十三、MMIT(microhomology mediated interchromosomal translocation)

(A)兩個非同源性染色體上均發生雙股斷裂(DSBs)，經 MMIT 後，產生二種衍生性染色體，屬於相對互調(reciprocal translocation)的互換。(B)雙股斷裂後產生 3'黏性末端，以確保兩個染色體上的 microhomology 均能被保留，配對後產生二種衍生性染色體。



圖二十四、生殖細胞染色體轉位產物形式之分類

將目標基因序列固定為正股(top strand, +)，依據染色體轉位對象位於染色體長臂或短臂，及其使用的微同源性(microhomology)序列为正股(top strand, +)或負股(bottom strand, -)，可將轉位對象分為四類。箭號方向為 $5' \rightarrow 3'$ 。 $5'$ 及 $3'$ PCR分別可偵測到四種類型的產物。將偵測到之染色體轉位依據此方式分類，數字為該形式產物之數量。DC: dicentric chromosome; AC: acentric chromosome; N: 轉位對象之染色體號碼。