

國立臺灣大學醫學院臨床醫學研究所

博士論文

Graduate Institute of Clinical Medicine

College of Medicine

National Taiwan University

Doctoral Dissertation

探討基因選擇性剪接對骨髓性血癌的臨床與生物學效應

**Exploring the clinical and biological effects of alternative
splicing on myeloid malignancies**

楊鎰聰

Yi-Tsung Yang

指導教授：周文堅教授，陳培哲教授

Advisor: Wen-Chien Chou, M.D., Ph.D., Pei-Jer Chen, M.D., Ph.D.

中華民國 113 年 6 月

June 2024



國立臺灣大學博士學位論文 口試委員會審定書

探討基因選擇性剪接對骨髓性血癌的臨床與生物學效應

Exploring the clinical and biological effects of
alternative splicing on myeloid malignancies

本論文係楊鎰聰君（學號 Q04421022）在國立臺灣大學
臨床醫學研究所完成之博士學位論文，於民國 113 年 06 月
07 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

周文堅

陳培哲 (簽名)

何三青

(指導教授)

王弘毅

譚如云

楊宏志

徐信安

周祖述

系主任、所長

(簽名)

(是否須簽章依各院系所規定)

誌謝

在過去幾年的研究所生涯中，我很榮幸能參與這個研究主題。我衷心感謝我的臨床指導老師周文堅教授，田蕙芬教授，以及臨醫所導師陳培哲教授，不僅提供了這個寶貴的機會，還在我的博士生涯中給予了無私的支持。他們花了大量時間與我討論研究計畫，令我深深感激。此外，我還要感謝我的博士論文委員會成員，包括譚婉玉老師、涂玉青老師、王弘毅老師、楊宏志老師及侯信安老師，他們在討論會和口試會議上提供了許多寶貴的建議和反饋。

接著，我要感謝周教授和田教授實驗室的成員，感謝他們在我剛加入實驗室時的熱情接納。特別感謝郭遠燁老師、李一泓博士、徐悅淳博士、林建嶽博士、高千珩研究助理、姚啟元醫師和莊博涵博士生。謝謝你們傳授寶貴的智慧和研究實驗的技巧，並感謝你們長久以來的友誼。

另外，我要感謝國立臺灣大學醫學院附設醫院新竹臺大分院支持我在職進修研究，並給予我莫大的幫助。我也要感謝臺灣國家衛生研究院基因微陣列核心實驗室，基龍米克斯生物科技股份有限公司，國立臺灣大學醫學院附設醫院流式細胞儀分析和分類核心設施，以及國立臺灣大學醫學院第一共同研究室生醫資源中心提供的技術服務。

我的研究得到了新竹臺大分院研究資助(107-HCH026)，臺大醫院研究資助(108-N4273)，臺灣國家科學及技術委員會研究資助(104-2923-B-002-001, 106-2314-B-002-224-MY3, 106-2314-B-002-231, 111-2314-B-002-267)，以及臺灣衛生和福利部研究資助(MOHW106-TDU-B-211-144005, MOHW107-TDU-B-211-114009, MOHW108-TDU-B-211-124009, and MOHW109-TDU-B-211-134009)。

最後，我要感謝我的家人，特別是我的太太，她悉心照顧我們的家庭。感謝我的父母和妹妹，你們無條件的愛和支持令我無比感激。此外，還要感謝我博士期間的同學和好友，感謝你們的友誼和支持。

中文摘要



幾乎所有人類的基因都會有選擇性剪接(alternative splicing)，一旦這些原本自然現象的選擇性剪接產生變異(aberrant alternative splicing)，就有可能改變生物體原本正常的生理功能，甚至癌化。根據近幾年的研究，選擇性剪接產生變異可能是致癌的一個重要機轉。

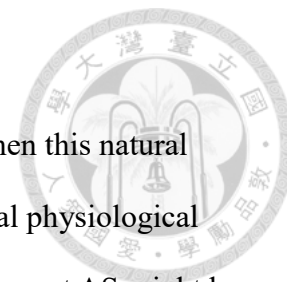
本研究計畫首先利用基因轉錄體微陣列(transcriptome array)的方式，分析了 176 個骨髓造血不良症候群(myelodysplastic syndrome; MDS)病患的骨髓細胞，以 20 個正常捐髓者的骨髓造血細胞檢體當對照組，發現 MDS 患者相對於正常人，在骨髓細胞整個基因體平均約有 25-30%的基因發生了變異選擇性剪接，若整個轉錄體變異選擇性剪接的程度越厲害，疾病的預後就越差。接著，利用 RNA 定序的方式(RNA sequencing)分析了 341 個急性骨髓性白血病(acute myeloid leukemia; AML)患者的骨髓檢體，進一步發現這些血癌細胞整個轉錄體整體選擇性剪接程度的改變會影響化療的療效與病患預後。因此，變異選擇性剪接可能在血癌致病機制中具有重要作用。

更進一步分析這些血癌患者發生變異選擇性剪接的基因，發現了一個候選基因 Nuclear Transcription Factor Y Subunit Alpha (*NFYA*)可能與血癌的致病機制及預後有關聯。*NFYA* 為 NFY 轉錄因子的最重要成分，它與基因上游的啟動子(promoter)結合後啟動其標的基因的表達。另外，*NFYA* 基因主要有兩個因選擇性剪接產生的轉錄異構物(isoforms)：*NFYA-L*[長形式，包含所有 10 個外顯子(exons)]和 *NFYA-S*[短形式，略過外顯子 3 (skipping exon 3)]。研究結果顯示，從臍帶血樣本(cord blood sample)中獲取的 CD34⁺正常造血幹細胞主要表達 *NFYA-S*，而 AML 癌細胞中可能同時表達 *NFYA-S* 和 *NFYA-L*。且較高的 *NFYA-L* 和較低的 *NFYA-S* 表達(*NFYA-L* 高表達模式)會與不好的臨床特徵與較差的存活率有關。相反，有較高的 *NFYA-S* 和較低的 *NFYA-L* 表達(*NFYA-S* 高表達模式)與化療後更高的緩解率和較好的存活率相關。這些血癌病患的轉錄體資料分析顯

示，*NFYA*-S 的高表達模式與細胞週期相關基因的高度表達相關，而 *NFYA*-L 的高表達模式會與 *KMT2* 基因重組的血癌細胞表現類似。這一發現與細胞實驗結果相符合，即 OCI-AML3 細胞中 *NFYA*-S 過度表達會增強細胞增殖能力，並增加對化療藥物 Ara-C 敏感性，但 *NFYA*-L 的過度表達卻會造成血癌細胞對化療藥物的抗藥性。這些研究結果突顯了 *NFYA* 基因選擇性剪接對 AML 的臨床與生物學重要性。

關鍵詞：選擇性剪接，變異選擇性剪接，骨髓造血不良症候群，急性骨髓性白血病，基因轉錄體微陣列，RNA 定序，*NFYA* 基因

英文摘要



Almost all human genes undergo alternative splicing (AS). When this natural phenomenon results in aberrant AS, it can alter an organism's normal physiological functions and even lead to cancer. Recent research indicates that aberrant AS might be a crucial mechanism in carcinogenesis.

In this research project, we first employed Affymetrix Human Transcriptome Array (HTA) 2.0 to identify global AS differences in 176 myelodysplastic syndrome (MDS) patients and 20 normal marrow donors. We found that, compared to normal individuals, MDS patients had approximately 25-30% of their genes undergoing aberrant AS in their bone marrow cells. The extent of global aberrant AS in MDS patients correlated with shorter leukemia-free survival, highlighting its involvement in the progression to acute leukemia. Subsequently, a comprehensive analysis of global AS in 341 *de novo* non-M3 acute myeloid leukemia (AML) patients using the RNA-seq platform demonstrated that the degree of global splicing pattern in AML patients could predict treatment outcomes independently of other well-established prognostic factors. These findings suggest that global AS aberrations may play a role in leukemogenesis, reflecting transcriptome complexity and instability, with potential implications for distinct clinical outcomes—a possible link between AS in specific genes and the pathogenesis of myeloid malignancies.

Further investigation into the Nuclear Transcription Factor Y Subunit Alpha (*NFYA*) gene in AML, which has two major isoforms: *NFYA*-L (long form, with all 10 exons) and *NFYA*-S (short form, skipping exon 3), revealed that the *NFYA*-L/*NFYA*-S ratios were higher in AML cells compared to CD34⁺ normal hematopoietic stem cells (HSCs) from cord blood samples. Patients with *NFYA*-L predominance (higher *NFYA*-L and lower *NFYA*-S expression) had worse prognostic features and clinical outcomes

after standard intensive chemotherapy compared to those with *NFYA*-S predominance (higher *NFYA*-S and lower *NFYA*-L expression). This prognostic effect was consistent regardless of age and the 2022 European LeukemiaNet (ELN) risk classification, as validated by The Cancer Genome Atlas (TCGA) cohort. Transcriptome analysis showed that *NFYA*-S predominance was associated with upregulated cell cycle-related genes, similar to those in active HSCs, indicating chemosensitivity. Conversely, *NFYA*-L predominance, as seen in *KMT2A*-rearranged leukemia, was linked to chemoresistance. This finding was supported by experiments showing enhanced cell proliferation and increased vulnerability to cytarabine in OCI-AML3 cells with *NFYA*-S overexpression. This study highlights the clinical significance of *NFYA* gene AS in AML, suggesting its potential as a prognostic biomarker with distinctive biological pathways in AML cells.

Key words : alternative splicing, aberrant alternative splicing, myelodysplastic syndrome, acute myeloid leukemia, transcriptome microarray, RNA sequencing, *NFYA* gene

目次



口試委員會審定書.....	i
誌謝.....	ii
中文摘要.....	iii
英文摘要.....	v
第一章 緒論(Introduction).....	1
1.1 前言.....	1
1.2 文獻回顧.....	3
1.2.1 基因選擇性剪接與疾病的關係.....	3
1.2.2 基因選擇性剪接與癌症的關係.....	3
1.2.3 基因選擇性剪接與血液惡性疾病的關係.....	4
1.3 研究的問題及其重要性.....	8
1.3.1 過去研究仍未解決的問題.....	8
1.3.2 研究問題的重要性.....	10
1.4 研究的假說與特定目的.....	10
1.4.1 假說1：整個轉錄體整體mRNA選擇性剪接模式.....	10
1.4.2 假說2：整個轉錄體整體mRNA選擇性剪接模式的程度與血癌治療的關係.....	11
1.4.3 假說3：整個轉錄體整體mRNA變異選擇性剪接模式的程度與血癌的發生.....	12
1.4.4 研究假說目的.....	12
第二章 材料與方法(Materials and methods).....	13
2.1 研究的材料.....	13
2.1.1 研究樣本族群(study sample cohorts).....	13
2.1.2 血癌細胞株(leukemia cell line).....	14
2.2 研究平台(study platform)與研究方法.....	15
2.2.1 microarray platform.....	15
2.2.2 RNA sequencing platform.....	17
2.2.3 細胞實驗.....	18
2.3 統計分析.....	23



第三章 結果(Results).....	24
3.1 基因選擇性剪接在骨髓造血不良症候群(MDS)病患上的角色.....	24
3.1.1 MDS患者族群的臨床基本資料.....	24
3.1.2 整個轉錄體上的變異選擇性剪接.....	24
3.1.3 變異選擇性剪接與臨床表徵的相關性.....	24
3.1.4 變異選擇性剪接對臨床預後的影響.....	26
3.1.5 變異選擇性剪接對細胞生理路徑的影響.....	26
3.1.6 驗證在microarray平台上所預測的變異選擇性剪接.....	27
3.2 基因選擇性剪接在急性骨髓性白血病(AML)病患上的角色.....	27
3.2.1 AML患者族群的臨床基本資料.....	28
3.2.2 整個轉錄體上的基因選擇性剪接.....	28
3.2.3 轉錄體整體選擇性剪接模式與臨床表徵的相關性.....	29
3.2.4 轉錄體整體選擇性剪接模式對臨床預後的影響.....	30
3.2.5 基因選擇性剪接對細胞生理路徑的影響.....	31
3.3 NFYA基因選擇性剪接對急性骨髓性白血病(AML)病患臨床預後的影響.....	31
3.3.1 NFYA基因的表現量和選擇性剪接與AML病患臨床表徵的關係及對預後的影響.....	31
3.3.2 NFYA基因選擇性剪接對AML病患血癌細胞潛在的生物學效應.....	33
3.3.3 NFYA基因選擇性剪接在白血病細胞株中的生物效應.....	33
第四章 討論(Discussion).....	36
4.1 基因選擇性剪接在MDS與AML病患上的臨床效應.....	36
4.2 NFYA基因選擇性剪接對急性骨髓性白血病(AML)的臨床與生物學效應.....	38
第五章 展望(Perspectives).....	43



圖次

圖 1：正常生理狀況下，基因選擇性剪接示意圖.....	46
圖 2：基因變異選擇性剪接示意圖.....	47
圖 3：比較 microarray platform 及 RNA-seq platform 在偵測基因選擇性剪接的差異.....	48
圖 4：剪接體的組裝(Spliceosome assembly).....	49
圖 5：Affymetrix Human Transcriptome Array 2.0 (HTA 2.0)示意圖.....	50
圖 6：MDS 轉錄體中 aberrant alternative splicing events 分析的流程圖.....	51
圖 7：microarray 平台上定義 aberrant alternative splicing event 的示意圖.....	52
圖 8：RNA-seq 平台上定義 alternative splicing event 的示意圖.....	53
圖 9：AML 轉錄體中 alternative splicing events 分析的流程圖.....	54
圖 10：lentiviral vector 示意圖.....	55
圖 11：分析MDS病患整個轉錄體變異選擇性剪接的類型(event type).....	56
圖 12：Aberrant alternative splicing score.....	57
圖 13：MDS病患與normal donor骨髓細胞轉錄體整體變異選擇性剪接模式(pattern).....	58
圖 14：比較MDS病患在變異選擇性剪接模式高變異度及低變異度臨床預後的差異.....	59
圖 15：利用PCR分析MDS病患AUP1, PVR2, GRIK5基因的變異選擇性剪接，以驗證在 microarray平台上所預測的變異選擇性剪接.....	60
圖 16：藉由Geneious軟體預測MDS患者AUP1的變異選擇性剪接，將導致蛋白質截斷 (protein truncation).....	60
圖 17：比較轉錄體整體選擇性剪接模式(pattern)在AML病患血癌細胞與臍帶血造血細胞間 的差異.....	61
圖 18：比較轉錄體整體選擇性剪接模式(pattern)在不同AML subtype(ICC classification)病患 間的差異.....	62
圖 19：分析AML病患及cord blood samples整個轉錄體選擇性剪接的類型(event type).....	63
圖 20：AML細胞基因體的基因突變數量似乎與轉錄體內所有基因選擇性剪接事件的數目 沒有相關性(NTUH_AML cohort).....	64

圖21：AML細胞基因體的基因突變數量似乎與轉錄體內所有基因選擇性剪接事件的數目沒有相關性(TCGA_AML cohort).....	64
圖22：分析AML病患轉錄體中所有基因選擇性剪接事件的加總數目，結果呈現鐘型曲線，近乎常態分佈.....	65
圖23：不同的global splicing pattern對NTUH_AML病患survival的影響.....	66
圖24：不同的global splicing pattern對TCGA_AML病患survival的影響.....	67
圖25：分析NFYA基因在AML上mRNA及protein的表達.....	68
圖26：比較NFYA基因在AML病患的血癌細胞與臍帶血的造血幹細胞(HSC)在mRNA表達量及splicing pattern上的差異.....	69
圖27：比較NTUH_AML病患，不同的NFYA表達量及不同的NFYA轉錄異構體表達量對survival的影響.....	70
圖28：比較AML病患不同的NFYA轉錄異構體表達型式(splicing pattern)對survival的影響.....	71
圖29：AML病患不同的NFYA轉錄異構體表達型式對整體基因的表達差異.....	72
圖30：AML病患不同的NFYA轉錄異構體表達型式潛在的生物學效應.....	73
圖31：OCI-AML3細胞實驗的示意圖.....	74
圖32：NFYA基因選擇性剪接在OCI-AML3白血病細胞株中的生物效應.....	75
圖33：OCI-AML3不同的NFYA基因選擇性剪接表達型式(splicing pattern)對細胞的遷移能力(migration ability)的影響.....	76
圖34：OCI-AML3不同的NFYA基因選擇性剪接表達型式(splicing pattern)對細胞週期分佈的影響.....	77
圖35：OCI-AML3不同的NFYA基因選擇性剪接表達型式(splicing pattern)潛在的生物學效應.....	78
圖36：Nuclear Transcription Factor Y (NFY)轉錄因子對目標基因及細胞生理路徑的影響示意圖；NFYA基因兩個主要的剪接異構體(alternatively spliced isoforms)示意圖.....	79
圖37：比較NTUH_AML兩群不同NFYA-L及NFYA-S表達的白血病患族群中特定蛋白質酪胺酸激酶基因表達量差異.....	80

表次

表1：MDS患者的臨床基本資料.....	81
表2：分析MDS病患整個轉錄體變異選擇性剪接在基因體上所影響的位置.....	82
表3：比較MDS病患在變異選擇性剪接模式高變異度及低變異度的臨床表徵差異.....	83
表4：多變數分析(multivariate analysis) MDS病患的預後因子，顯示變異選擇性剪接模式是一個獨立的預後因子.....	84
表5：分析MDS病患整個轉錄體變異選擇性剪接對細胞生理路徑的影響.....	85
表6：NTUH_AML患者的臨床基本資料.....	86
表7：TCGA_AML患者的臨床基本資.....	88
表8：比較NTUH_AML cohort不同的global splicing pattern間基因突變的差異.....	89
表9：比較TCGA_AML cohort不同的global splicing pattern間基因突變的差異.....	90
表10：單變數及多變數分析(univariate and multivariate analysis) NTUH_AML病患的預後因子，顯示轉錄體整體的選擇性剪接模式是一個獨立的預後因子.....	91
表11：分析NTUH_AML病患整個轉錄體選擇性剪接對細胞生理路徑的影響.....	92
表12：比較NTUH_AML cohort不同的NFYA基因splicing pattern間臨床表徵的差異.....	93
表13：比較NTUH_AML cohort不同的NFYA基因splicing pattern間基因突變的差異.....	94
表14：比較NTUH_AML病患，不同的NFYA-S表達量對誘導性化學治療(induction chemotherapy)的影響.....	95
表15：單變數及多變數分析(univariate and multivariate analysis) NTUH_AML病患的預後因子，顯示NFYA轉錄異構體表達型式(splicing pattern)是一個獨立的預後因子.....	96
參考文獻(References).....	97
附錄：個人在修業期間所發表之論文清冊.....	103

第一章、緒論(Introduction)

1.1 前言

真核生物的基因序列中，包含了內含子(intron)與外顯子(exon)，兩者交互穿插，組成基因。其中 intron 在基因轉錄成前信使核糖核酸(pre-mRNA)後會被剪接體(spliceosome)移除，剩下的 exon 才是能夠存在於信使核糖核酸(mRNA)(之後再進一步轉譯成蛋白質)的片段。而選擇性剪接(alternative splicing)便是利用這樣的特性，將同一基因中的 exon 以不同的組合方式來表現，使一個基因在不同時間，不同環境中能夠製造出不同的蛋白質，可增加生理狀況下系統的複雜性/適應性(圖 1)。因此，選擇性剪接是真核生物自然的生理現象，大約 70%以上的人類基因都會有選擇性剪接，尤其是有多個 exon 的基因。一旦原本正常調控這些自然現象的選擇性剪接的機轉發生變異，就會導致變異選擇性剪接(aberrant alternative splicing)的發生，因而改變生物體原本許多正常的生理功能，例如這些變異選擇性剪接若發生在調控生長，代謝，分化，細胞凋零等基因，就有機會使細胞一步一步走向癌化的過程(圖 2)¹⁻³。

成人的血癌多數是來自骨髓性血液惡性疾病，其中主要包含了骨髓增生性疾病(myeloproliferative neoplasm, MPN)，骨髓造血不良症候群(myelodysplastic syndrome, MDS)及急性骨髓性白血病(acute myeloid leukemia, AML)。不管是骨髓增生性疾病，骨髓造血不良症候群，或是急性骨髓性白血病皆是很複雜的骨髓性惡性疾病，包含了各式各樣的造血疾病，有多樣的臨床表現。骨髓增生性疾病主要會造成成熟血球的增高，增加血栓栓塞的風險；而骨髓造血不良症候群主要可能會發生因無效造血所導致的全血球低下外，並可能進一步衍生惡化成急性骨髓性白血病。根據過去的研究，血液惡性疾病主要是來自骨髓內同一株癌化的造血幹細胞(clonal hematologic progenitor or stem cell disorders)，且 80-90%癌化的造血幹細胞與正常的造血幹細胞相比，在基因體中至少有一個基因突變⁴⁻⁶。近幾年，由於高通量定序(high-throughput sequencing)的發展，癌化

的造血細胞除了基因體可能有基因突變外，轉錄體的變異(transcriptome aberrancy)也逐漸被我們關注。轉錄體的變異除了可能來自 mRNA，小分子核糖核酸(microRNA)及長鏈非編碼核糖核酸 LncRNAs(long non-coding RNAs)不正常的表達外，變異選擇性剪接可能也是一個重要的致病機轉^{3,7-24}。然而變異選擇性剪接在血液惡性疾病的臨床與生物學效應目前仍未被廣泛的研究探討。因此本論文想探討這種變異選擇性剪接在骨髓性血液惡性疾病致病機轉中的角色。

早期研究選擇性剪接的方法有表達序列標籤 EST(Expressed Sequence Tag)或交聯免疫沉澱法 CLIP(cross-linking immunoprecipitation)。EST 是從一組由某一種特定的細胞或組織在特定的環境條件下，所有基因表達的序列，再轉換成許多片段互補去氧核糖核酸 cDNA(complementary DNA)序列所集結而成。一般這些小片段序列約在 60-500bp，因此直接研究 EST 可獲取基因表現訊息或基因是否有產生選擇性剪接，但此種研究方式會受限於只能探索目前已經存在有大量 EST 的組織或細胞²⁵。另外 CLIP 是利用 RNA 和蛋白質在 256 nm 紫外光照射下會發生共價交聯的特性，藉此來研究與剪接因子(splicing regular protein)調控相關的 mRNA 片段序列，但實驗方式較為繁瑣，有時 RNA 與蛋白質交聯的效率太差，也會影響後續的實驗結果判斷²⁶。因此近期研究選擇性剪接的方法主要為微陣列晶片分析(microarray)及 RNA 定序法(RNA sequencing)。Microarray 是利用微陣列生物晶片上數萬個不同的探針(probe)，來偵測檢體中所有基因表達的 mRNA 片段(transcripts)，藉以研究基因選擇性剪接的發生²⁷。另外 RNA sequencing 是利用次世代的定序技術，對樣品中的 mRNA 片段定序，可以得到全部(在轉錄層面)基因表現的資訊²⁸。目前 microarray 或 RNA sequencing(RNA-seq)來研究基因選擇性剪接的方式都已是標準化的實驗方法，所需的檢體量也不多，一般只需 1 µg RNA。最近有研究論文比較使用 microarray 及 RNA sequencing 在研究基因選擇性剪接上的差異，結果發現若把 *P* value 訂在小於 0.001 為有意義的基因選擇性剪接事件(significant alternative splicing event)，而 *P*

value 大於 0.2 視為無意義的基因選擇性剪接事件，則此兩研究平台得到的研究結果會高度一致(圖 3)^{29,30}。



1.2 文獻回顧(Literature review)

1.2.1 基因選擇性剪接與疾病的關係

基因選擇性剪接的調控是一個非常精細的過程，其中剪接體(spliceosome)扮演了非常重要的角色。剪接體由小核核糖核酸 small nuclear RNA(snRNA)和超過 200 種蛋白質 small nuclear ribonucleoprotein(snRNP)與 pre-mRNA 一系列動態相互作用組裝而成(圖 4)³¹。根據過去幾年的研究，許多 snRNP 譯碼基因(coding gene)被發現產生突變，可能因此改變了 spliceosome 正常的生理功能，因而使生物體原本正常生理功能的基因選擇性剪接發生了變異，造成許多良性與惡性疾病的發生。例如許多神經系統的疾病，如阿茲海默症(Alzheimer's disease)，帕金森氏症(Parkinson's disease)，精神分裂症(Schizophrenia)，脊髓性肌肉萎縮症(spinal muscular atrophy)或其他良性疾病，如感染症，自體免疫性疾病及代謝性疾病...等，可能都與 snRNP 譯碼基因突變所造成的變異選擇性剪接有關^{24,32,33}。除了上述良性疾病外，目前已有許多腫瘤的致病機轉也被發現與變異選擇性剪接有關，例如固態腫瘤中頭頸癌，肺癌，腦癌，乳癌，肝癌，大腸癌，卵巢癌...等，另外許多血液惡性疾病，例如骨髓造血不良症候群及急性骨髓性白血病也被發現可能與此致病機轉相關^{34,35}。

1.2.2 基因選擇性剪接與癌症的關係

近期的文獻報告指出，基因選擇性剪接原本正常的調控機制發生異常，幾乎是所有腫瘤類型的共同分子特徵。與癌症相關的基因選擇性剪接變異主要是來自上述剪接體蛋白質(snRNP)的譯碼基因(spliceosome protein coding gene)突變，或影響基因選擇性剪接的催化和調控作用因子異常表達(altered expression

of trans-acting factors governing splicing catalysis and regulation)(圖 2)^{3,35}。人體正常的細胞在癌化的過程中可能會牽涉很多正常生理功能路徑發生改變，例如細胞生長(cell proliferation)，細胞凋零死亡(cell apoptosis)，血管新生(angiogenesis)，細胞代謝(metabolism)，及細胞侵襲且轉移(invasion and metastasis)。如果調控這些原本正常生理功能路徑的基因表現產生變異，就有可能使這些正常生理功能失去平衡，造成細胞增殖增加，降低細胞凋亡，增強侵襲和轉移能力，對化療產生抗藥性以及逃避正常免疫系統監視，而導致細胞一步一步癌化，而變異選擇性剪接可能就是造成這些基因表現異常的重要致病機轉。例如 integrin subunit $\alpha 6$ (*ITGA6*)基因因選擇性剪接可以產生兩種 mRNA 異構物(isoforms)：*ITGA6A* 及 *ITGA6B*，而 *ITGA6A* 會促進細胞生長，在大腸癌細胞內被大量表現，這可能與調控此基因選擇性剪接的剪接體蛋白質異常有關。此外 *VEGF-A* 基因因變異選擇性剪接而導致下游 mRNA 異構物表現異常，將會使原本組織內正常血管新生的機制失去平衡，導致組織異常的生長癌化。另外 *PKM* 基因下游的兩個 mRNA 異構物(*PKM1* and *PKM2*)會因變異選擇性剪接而導致細胞異常的代謝反應，*PKM1* 主要存在正常細胞，而 *PKM2* 在癌細胞中會被大量表達³⁵。

目前癌症的治療策略，除了手術，放射線治療，化療，標靶藥物，到近期最夯的免疫或細胞治療外，最新的研究已經有一些小分子的化合物被開發用來調節或抑制 mRNA 變異選擇性剪接²⁴，期待未來這也能成為另一個新的抗癌利器。

1.2.3 基因選擇性剪接與血液惡性疾病的關係

a. 骨髓性血液惡性病

在 2011 年，隨著次世代定序技術的進步，有一群剪接體蛋白質(snRNP)譯碼基因(spliceosome protein coding genes)突變首先在骨髓造血不良症候群的患者



上被發現，如上述，此類基因突變之後也陸續在其他固態腫瘤(solid cancer)被發現。在血液惡性疾病中主要有以下幾個基因突變：*SF3B1*，*SRSF2*，*U2AF1*，*ZRSR2*。約有 7-28%骨髓造血不良症候群的患者有 *SF3B1* 基因突變，尤其是在 RARS(Refractory anemia with ring sideroblasts)亞型。目前的研究指出，*SF3B1* 基因突變會造成 pre-mRNA 3'剪接位點的選擇異常(aberrant 3' splice-site selection)以及增加降解含有早期終止碼(nonsense codon)的 mRNA 分子(nonsense-mediated decay)。因此 *SF3B1* 基因突變可能對基因選擇性剪接的過程和 mRNA 的成熟產生重要影響，進而影響細胞功能和可能導致疾病的發生。有此基因突變的骨髓造血不良症候群患者的臨床表現可能有貧血，但卻有較好的整體存活期(overall survival)。*SRSF2* 是一種富含絲氨酸/精氨酸的剪接因子(splicing factor)，它通過與外顯子剪接增強子(exonic splicing enhancers)結合，能夠調節外顯子的包含或排除(inclusion or exclusion of exons)，從而影響 mRNA 的剪接模式，因此 *SRSF2* 基因突變將導致變異選擇性剪接的發生。*SRSF2* 基因突變約占 5-15%骨髓造血不良症候群的患者，會造成白血球低下及貧血的臨床表現，且整體的存活期較沒有此基因突變的病患短。*U2AF1* 基因突變約占 5-10%骨髓造血不良症候群的患者，也會造成 pre-mRNA 3'剪接位點的選擇異常，干擾正常選擇性剪接的過程，發生變異選擇性剪接，臨床上會造成白血球低下，貧血及血小板低下，也會有較差的整體存活期。*ZRSR2* 剪接因子在辨識和剪接 U12 型內含子(U12-type introns)上扮演重要角色，因此 *ZRSR2* 突變可能導致 U12 型內含子 RNA 剪接的變異(dysregulated RNA splicing of U12-type introns)，*ZRSR2* 基因突變約只占 1-11%骨髓造血不良症候群的患者，也會有較差的預後。由於這類剪接體蛋白質譯碼基因突變在骨髓造血不良症候群患者被反覆(recurrent)且大量的發現(約 50%的患者有此類基因突變)，隨後這些基因突變的轉殖鼠(*SF3B1* and *SRSF2* mutations knockin mouse model)也會出現像骨髓造血不良症候群患者的臨床表現，因此這些基因突變可能是造成骨髓造血不良症候群的重要因素。這些基因

突變導致剪接體的正常功能受到影響，或許因而導致變異選擇性剪接的發生。然而這些調控剪接體功能的基因突變後，是否真的會影響及改變整個轉錄體整體的 mRNA 選擇性剪接模式(global mRNA alternative splicing pattern)，以及這些基因變異選擇性剪接是否會造成臨床與生物學效應，例如是否會影響骨髓造血不良症候群病患的預後，目前仍然未被廣泛的研究與探討³⁶。

2014 年在 clinical cancer research 有一篇在探討急性骨髓性白血病患者(acute myeloid leukemia)的癌細胞與正常人的造血細胞在轉錄體間所有基因選擇性剪接的差異，文中的 228 個急性骨髓性白血病患是來自兩大醫學中心，美國的 Dana-Farber Cancer Institute (DFCI)及法國的 University Hospital de Nantes (UHN)，最後從中挑選了 66 個急性骨髓性白血病患(31 個來自 DFCI，35 個來自 UHN)與另外 8 個來自正常人的骨髓細胞檢體當對照組，在 Affymetrix Human Exon 1.0ST Array 的平台上分析這些變異選擇性剪接，結果顯示急性骨髓性白血病患在基因體中會經常性的發生變異選擇性剪接，在整個基因體中有 29%基因發生了變異選擇性剪接，在這些發生變異選擇性剪接的基因，大約八成是落在蛋白質編碼區域(protein coding region)，將近兩成是蛋白質非編碼區域(protein noncoding region)。另外進一步分析發現這些變異選擇性剪接分布在一些致癌基因(oncogene)，腫瘤抑制基因(tumor suppressor gene)，影響剪接體(spliceosome)的基因，調控細胞增殖及凋零死亡(apoptosis)的基因...等，因此將來或許可考慮研發抑制這些變異選擇性剪接的藥物來治療急性骨髓性白血病患¹⁶。

台大血癌團隊在 2016 年曾經發表大約 10.8%急性骨髓性白血病患具有剪接體蛋白質(snRNP)譯碼基因的突變(spliceosome protein coding gene mutations, spliceosome gene mutations)，且若有這些基因突變的病患，會有較差的臨床預後。由於剪接體是調控正常生理選擇性剪接的一個重要關鍵因素，一旦剪接體功能異常，將導致變異選擇性剪接的產生，這可能暗示著選擇性剪接的變異將

可能改變血癌患者原本的預後結果³⁷。

目前確切的血癌致病機轉仍不太清楚，但多數的研究指出血癌的發生至少需要兩類基因突變：一個突變導致細胞增殖優勢，另一個突變則導致造血細胞在分化過程的損害("雙擊"假說)(the "two-hit" hypothesis)^{38,39}。Omar Abdel-Wahab 研究團隊在 2019 年 nature 雜誌中指出，分析 982 名急性骨髓性白血病患者的轉錄體，發現了 *IDH2* 和 *SRSF2* 的基因突變經常重疊出現(frequent overlap of mutations in *IDH2* and *SRSF2*)，*IDH2* 基因主要會影響表觀遺傳調控(epigenetic regulation)，而 *SRSF2* 是剪接體蛋白質譯碼基因(spliceosome protein coding gene)之一，如之前所述，會影響 pre-mRNA 選擇性剪接的正常生理過程。因此 *IDH2* 與 *SRSF2* 兩基因共同突變將影響原本正常 epigenetic regulation 和 RNA splicing，並產生協同效應，促進白血病發生。儘管 *IDH2* 或 *SRSF2* 的單一突變導致不同的 pre-mRNA 選擇性剪接變化，但 *IDH2* 與 *SRSF2* 的共同突變會改變原本 *SRSF2* 所造成的選擇性剪接效應，並且比單獨任何一種突變更深刻地改變了原本正常生理功能的選擇性剪接。另外，*IDH2* 與 *SRSF2* 的共同突變也會導致造血細胞致命的增殖異常和增強的自我更新能力 (self-renewal)，而單獨突變則未觀察到這種情況。研究團隊更進一步的發現，*IDH2* 和 *SRSF2* 雙突變細胞除了會造成變異選擇性剪接的發生，也會降低整合複合物成員 INTS3(a member of the integrator complex)的表達，進一步會造成 RNA 聚合酶 II(RNA polymerase II, RNAPII)的阻滯(stalling)增加。異常的 *INTS3* 選擇性剪接與 *IDH2* 突變共同促進了白血病發生。這些實驗結果證明了在某些類型的白血病中，epigenetic regulation 和 RNA splicing 之間存在病理性的交互作用，提供了剪接因子突變(spliceosome gene mutations)驅動骨髓系惡性腫瘤發展的功能證據，並將剪接體變化(spliceosomal change)確定為 *IDH2* 突變引起白血病發生的調節因子⁴⁰。

近幾年由於高通量定序技術(high throughput technology)的進步，及公開的血癌轉錄體定序資料庫(public transcriptome database)，例如 The Cancer Genome

Atlas(TCGA)，Beat，Therapeutically Applicable Research to Generate Effective Treatments(TARGET)，及 Clinseq_AML，目前已經有一些報告指出某些特定的基因發生變異選擇性剪接會影響急性骨髓性白血病患的預後，有些基因的變異選擇性剪接會有較好的臨床結果，但有些卻會有相反較差的結果¹⁸⁻²⁰。另外有一篇研究比較 European LeukemiaNet(ELN) 2017 預後好及預後差的兩群急性骨髓性白血病患轉錄體中選擇性剪接的差異，發現這些基因選擇性剪接的差異會導致細胞壓力反應(cellular stress response)的生理路徑活化，導致急性骨髓性白血病患的預後變差。此篇研究也發現這些有差異選擇性剪接的發生並非一定在有剪接體蛋白質譯碼基因突變(spliceosome gene mutations)的病患上，且大約45%這些差異選擇性剪接主要會影響下游轉譯蛋白質的高度保留區域(highly conserved protein domains)²¹。

b. 淋巴性血液惡性疾病

如同骨髓性血液惡性疾病，淋巴性血液惡性疾病的致病機轉也可能來自於變異選擇性剪接。以最常見的淋巴瘤：瀰漫性大 B 細胞淋巴瘤(Diffuse large B cell lymphoma)為例，發生變異選擇性剪接的原因可以源自於上述的剪接體蛋白質譯碼基因的突變(spliceosome gene mutations)，另外也有可能是正常剪接的位點序列產生突變(splicing site mutations)，這些改變會造成變異選擇性剪接的發生，進而影響了細胞原來正常的生理路徑，如腫瘤抑制，DNA 修復，細胞週期，細胞分化，細胞增殖和細胞凋亡。因此，在淋巴器官生殖中心的 B 細胞中發生了惡性轉化、癌症進展和轉移。根據目前的研究顯示，受剪接突變(splicing mutations)影響最常見的基因包括 *BCL7A*，*CD79B*，*MYD88*，*TP53*，*STAT*，*SGK1*，*POU2AF1*，*NOTCH1*⁴¹。

1.3 研究的問題及其重要性

1.3.1 過去研究仍未解決的問題

血液惡性疾病主要是來自骨髓內同一株癌化的造血幹細胞(clonal hematologic progenitor or stem cell disorders)，過去研究血癌的致病機轉主要集中在基因體的變異(genome variation)，包含基因突變，單一核苷酸多型性 (single nucleotide polymorphism, SNP)，拷貝數變異(copy number variation, CNV)，雜合性缺失(loss of heterozygosity, LOH)，基因重組(genomic rearrangement)...等⁴²，目前已知骨髓造血不良症候群或急性骨髓性白血病患者體內大約 80-90%的基因體有發生變異^{5,6,43}。除了基因體的變異外，隨著高通量次世代定序技術的進步，轉錄體變異的研究也逐漸蓬勃發展，這包含基因的表達量(gene expression)差異，非編碼核糖核酸(non-coding RNA, ncRNA)，小分子 RNA(small RNA, sRNA)，另外也包含了基因異常的選擇性剪接^{7-14,16-22}。

骨髓造血不良症候群是一群很複雜的骨髓性惡性疾病，它包含了各式各樣的造血疾病，有多樣的臨床表現，除了可能會發生無效造血所導致的全血球低下外，並可能進一步衍生惡化成急性骨髓性白血病。目前文獻對於骨髓造血不良症候群致病機轉的研究大多侷限於基因體的變異，尤其是剪接體蛋白質譯碼基因的突變(spliceosome gene mutations)(如上所述)³⁶，然而這些調控剪接體功能的基因突變後，是否真的會影響及改變整個轉錄體整體的 mRNA 選擇性剪接模式(global mRNA alternative splicing pattern)，另外不同病患間 global mRNA alternative splicing pattern 是否有差異，不同的 global mRNA alternative splicing pattern 是否會影響病患的臨床預後。以及是否有哪些特定的基因變異選擇性剪接具備某種臨床效應或生物意義，是否會影響哪些正常生理功能路徑的調控，而進一步加速惡化成急性骨髓性白血病，或影響病患的預後，目前仍然是一個需要進一步探究的議題。

近幾年有越來越多的研究文獻使用公開的轉錄體資料庫(TCGA, Beat, TARGET, Clinseq_AML public transcriptome databases)探索基因選擇性剪接對急性骨髓性白血病的影響¹⁸⁻²²，然而，這仍存在一些問題未被謹慎考慮到的。首

先，這些資料庫包含許多相對異質性患者族群，如原發性(*de novo* AML)，次發性(*secondary* AML)或是與治療相關的血癌患者(*therapy related* AML)；所收集的檢體樣本來自不同的時間點，如初診斷(*at initial diagnosis*)，復發/難治狀態(*relapsed/refractory status*)，或接受不一致治療方式(如高強度化療或僅接受支持性治療)。其次，這些不同研究學者利用這些公開資料庫分析後所得出對預後有相關的特定基因選擇性剪接事件(*prognosis associated alternative splicing events*)並不一致。第三，這些研究只找出一些單獨的選擇性剪接事件(*alternative splicing events*)，並未探討整個轉錄體整體的 mRNA 選擇性剪接模式(*global mRNA alternative splicing pattern*)對臨床所產生的影響。最後，這些研究並無“正常”的造血幹細胞(*hematopoietic stem cell*)當作對照組。另外值得注意的是，由於這些公開的轉錄體資料庫主要來自歐美患者的檢體，因此我們也想知道對這些西方患者族群的分析結果是否也適用於亞洲患者。

1.3.2 研究問題的重要性

基因選擇性剪接的變異可能參與了血癌致病機轉的過程，同時也可能影響疾病的臨床表現和臨床預後結果。另外，我們也對會影響正常生理功能路徑或疾病預後的基因選擇性剪接事件(*alternative splicing event*)感到興趣。目前血癌的治療主要在化療藥物或特定基因上的標靶藥物，但目前血癌治療的成績仍令人十分不滿意，尤其在老人血癌的病患。希望這個研究計畫成果會帶來科學上的重大影響，而開發影響基因選擇性剪接藥物可能是血癌治療之新策略。

1.4 研究假說與特定目的

1.4.1 假說 1：本研究論文以轉錄體中所有基因選擇性剪接事件的數目加總(*the number of alternative splicing events*)，來代表整個轉錄體整體 mRNA 選擇性剪接模式(*global mRNA alternative splicing pattern*)

本篇論文想探討不同病患間整個轉錄體整體的 mRNA 選擇性剪接模式是否會有差異，但如何評估整個轉錄體整體的選擇性剪接形式？回顧大多數的文獻，是把所有轉錄體中所有基因 alternative splicing events 的數目加總起來，在比較個案在不同狀態下(例如 at disease status or disease recovery status; at birth or during maturation)的差異^{16,21,44-47}。因此本論文利用轉錄體中所有基因 alternative splicing events 的數目加總，來代表此病患整個轉錄體整體的選擇性剪接模式。本論文進一步定義若轉錄體中所有基因選擇性剪接事件(alternative splicing events)的數目加總較高，則轉錄體 mRNA 選擇性剪接的程度較厲害，代表整個轉錄體中有較多不同的 mRNA transcripts；若轉錄體中所有基因選擇性剪接事件(alternative splicing events)的數目加總較低，則轉錄體 mRNA 選擇性剪接的程度較輕微，整個轉錄體不同的 mRNA transcripts 數量較少。藉由分析不同病患間轉錄體整體的選擇性剪接模式差異，配合不同病患間完整的臨床及基因突變(genetic mutation status)資訊，便有一個很好的機會可以來探討整個轉錄體整體的 mRNA 選擇性剪接模式在血液惡性疾病患者上可能的臨床生物意義。

1.4.2 假說 2：整個轉錄體整體 mRNA 選擇性剪接模式的程度(the degree of global alternative splicing pattern)會影響白血病患接受標準高強度化療(standard intensive chemotherapy)的療效，因而影響病患的預後。

本論文推測整個轉錄體中所有基因選擇性剪接事件(alternative splicing events)的數目加總較高，則轉錄體 mRNA 選擇性剪接的程度較厲害，代表整個轉錄體中有較多不同的 mRNA transcripts，這可能會讓轉錄體較不穩定(transcriptome instability)⁴⁴，進而改變標準高強度化療的療效，影響這些病患的預後。

1.4.3 假說 3：整個轉錄體整體 mRNA 變異選擇性剪接模式的程度(the degree of global aberrant alternative splicing pattern)會促進白血病的發生，影響病患的預後。

若以正常人骨髓造血細胞的轉錄體當對照組，比較病患血癌細胞與正常人骨髓造血細胞在轉錄體中所有選擇性剪接間的差異，把所有具統計意義的變異選擇性剪接事件的數目加總起來作為整體變異選擇性剪接模式(global aberrant alternative splicing pattern)。根據之前的一些研究結果，剪接體蛋白質譯碼基因(spliceosome protein coding genes)突變可能是造成骨髓造血不良症候群疾病或急性骨髓性白血病的重要致病機轉。由於剪接體是調控轉錄體 pre-mRNA 選擇性剪接的重要關鍵因子，因此本論文推測這些病患轉錄體中整體變異選擇性剪接的程度會影響許多正常生理功能路徑的改變，促進血癌的發生，更進一步影響病患的預後。

1.4.4 研究假說目的

目前有許多影響基因選擇性剪接藥物正在如火如荼的研發中，本研究論文主要是想探討基因選擇性剪接對骨髓性血癌的臨床與生物學效應，期待這些影響基因選擇性剪接的藥物未來能成為血癌治療之新策略。

第二章、材料與方法(Materials and methods)



2.1 研究的材料

2.1.1 研究樣本族群(study sample cohorts)

為了研究基因選擇性剪接對骨髓性血癌的臨床與生物學效應，本研究論文共收集了兩類病患族群(骨髓造血不良症候群及急性骨髓性白血病)

a. 骨髓造血不良症候群(myelodysplastic syndrome, MDS)

本研究論文收集了從 1991 年 11 月至 2010 年 12 月，在台大醫院(National Taiwan University Hospital, NTUH)新診斷出骨髓造血不良症候群(MDS)患者，並且這些患者皆有足夠可用的樣本檢體和完整的臨床和基因突變資料，納入這項研究。此外，本研究論文還收集了另一群在 2011 年 1 月至 2012 年 5 月期間在台大醫院被新診斷的骨髓造血不良症候群(MDS)患者族群，作為驗證群體(validation cohort)。

b. 急性骨髓性白血病(acute myeloid leukemia, AML)

本研究論文納入了 2004 年 2 月至 2016 年 12 月間在台大醫院新診斷，原發性(*de-novo*)，非 M3 類型急性骨髓性白血病(non-M3 AML)患者，排除了具有先前的血液疾病或治療相關的 AML (antecedent hematological diseases or therapy-related AML)患者。這些患者接受了標準高強度的化療，並且有完整的臨床和基因資訊。此外，本研究論文也跟 National Center for Biotechnology Information The Database of Genotypes and Phenotypes(NCBI dbGaP)(基因型和表型數據庫)申請權限，從 Genomic Data Commons 數據門戶 (<https://portal.gdc.cancer.gov/>)下載了美國癌症基因組圖譜計劃-急性骨髓性白血病患族群(TCGA AML)的 RNA-seq 原始數據(raw data)，作為驗證群體(validation cohort)。

c. 對照組族群(control cohort)

在上述病患檢體收集的時間，本研究論文額外收集了 20 個正常骨髓移植幹細胞捐贈者的骨髓檢體作為 MDS 患者的對照組，另外也跟台大醫院婦產部合作，共收集了 6 個臍帶血樣本的檢體當 AML 患者研究參考的對照組。

本研究論文計畫已獲得國立台灣大學醫院(NTUH)的機構審查委員會批准(批准號碼：201603034RINA、201709072RINC)。

2.1.2 血癌細胞株(leukemia cell line)

除了上述病患的骨髓檢體外，目前實驗室也有許多骨髓性血癌的細胞株，例如 OCI-AML3, TF-1, HEL, K562, KG-1, KG-1a (DSMZ, Braunschweig, Germany), MV4-11, 及 HL-60 (ATCC, Manassas, VA, USA)，我們將選擇 1-2 株 myeloid leukemia cell lines 作為後續細胞實驗使用，以驗證本研究計劃我們所觀察到的一些結果。

HEL、K562、KG-1 和 HL-60 細胞在 RPMI 1640 培養基(GIBCO, Life Technologies Corporation, NY, USA)中培養，加入 10% heat-inactivated fetal bovine serum (FBS)和抗生素(青黴素 100 U/ml，鏈黴素 100 µg/ml) (GIBCO, Life Technologies Corporation, NY, USA)。OCI-AML3、KG-1a 和 MV4-11 細胞在 RPMI 1640 培養基中培養，加入 20% FBS 和抗生素(青黴素 100 U/ml，鏈黴素 100 µg/ml)。TF-1 細胞在 RPMI 1640 培養基中培養，加入 20% FBS、抗生素(青黴素 100 U/ml，鏈黴素 100 µg/ml)和 4 ng/ml 重組 GM-CSF (GIBCO, Life Technologies Corporation, NY, USA)。所有細胞株在 37°C，5% CO₂ 濕潤的環境中培養。定期使用 Lonza MycoAlert Mycoplasma Detection Kit 檢測是否有黴漿菌污染。這些細胞株的基因型皆通過臺灣新竹財團法人食品工業發展研究所進行 16-marker short tandem repeat analysis。結果證實，基因型(genetic profiles)與其報告的通用型號(generic profiles)相同。

2.2 研究平台(study platform)與研究方法

2.2.1 microarray platform

本研究計畫初期，一開始使用 microarray 平台來分析骨髓造血不良症候群病患與正常骨髓捐贈者骨髓造血細胞(當對照組)，此兩者整個轉錄體選擇性剪接的差異。為了回答這個問題，本研究論文使用了 2013 年 Affymetrix 開發出來的微陣列晶片平台 (Affymetrix Human Transcriptome Array 2.0)，此種高解析度的晶片上有超過 600 萬個不同的探針(probe)，這些探針涵蓋了蛋白質編碼區域 (protein coding region)及蛋白質非編碼區域 (protein noncoding region)(如圖 5 所示)，與傳統的 exon array 或 gene array 相比，此種晶片 probe 涵蓋的轉錄體區域更密更廣，最重要的是還包含了外顯子間的區域(exon-exon junction)。所以這個晶片平台是目前唯一可以較精準的偵測選擇性剪接的發生。本研究抽取這些骨髓檢體內單核細胞中的 RNA，經過處理後並進一步與 Affymetrix Human Transcriptome Array 2.0 晶片上的探針進行雜交，之後透過晶片所釋出的訊號，掃描成影像檔，經過一系列的軟體轉換處理，最後再使用 Affymetrix Transcriptome Analysis Console version 3.0 software 來分析骨髓造血不良症候群病患與正常骨髓捐贈者骨髓細胞間的變異選擇性剪接(aberrant alternative splicing)(簡易的流程圖如圖 6 所示)。本研究論文報告中的 Affymetrix Human Transcriptome Array 2.0 原始和標準化微陣列數據已存放在 Gene Expression Omnibus 數據庫中，存取編號為 GSE107400。

本研究論文透過 Affymetrix Transcriptome Analysis Console version 3.0 software 的 splicing index (SI) default setting 來分析 MDS 病患與正常捐髓者 (normal control donor)骨髓細胞間的變異選擇性剪接 (aberrant alternative splicing)。剪接指數(SI)是一個用於評估基因選擇性剪接事件(gene alternative splicing event)在不同條件下變化的指標^{48,49}。如果 splicing index 的數值大於 2，或小於-2 即表示某基因在此處的 exon 發生了變異選擇性剪接事件 (aberrant

alternative splicing event) (+ : condition 1 > condition 2 ; - : condition 1 < condition 2)。例如，我們想知道 gene A 中的 exon 2 在 MDS 病患中是否有發生變異選擇性剪接事件， $SI = (\text{exon 2 在 MDS 的表達量} / \text{gene A 在 MDS 的表達量}) / (\text{exon 2 在 normal control donor 的表達量} / \text{gene A 在 normal control donor 的表達量})$ (如圖 7 所示)。接著本研究把整個基因體中上述所得到的所有變異選擇性剪接事件個數(the number of aberrant alternative splicing events)與所影響(involve)的全部基因數目(the number of aberrant alternative splicing genes)相除，便可得到平均每個發生變異選擇性剪接的基因 (aberrant alternative splicing gene) 共發生了多少次變異選擇性剪接事件(aberrant alternative splicing events)，本研究論文藉此數值(aberrant alternative splicing score)來代表該個案整個轉錄體發生變異選擇性剪接的模式(global aberrant alternative splicing pattern)。

接著可以使用上述本研究論文新定義個案發生變異選擇性剪接的數值(aberrant alternative splicing score)，使用統計軟體 SPSS 22 version 中的 t 檢定，或 ANOVA 檢定來分析本研究想提出的第一個問題：這些變異選擇性剪接(aberrant alternative splicing)與骨髓造血不良症候群的疾病 FAB risk 分類或不同的基因突變型式(genetic mutation status)是否有什麼關聯性？若進一步搭配每個病患的臨床基本資料，使用統計軟體存活分析中的 Kaplan-Meier survival curve，Multivariate analysis 就可嘗試回答第二個問題：變異選擇性剪接(aberrant alternative splicing)是否也為骨髓造血不良症候群病患的獨立預後因子(independent prognostic factor)之一。更進一步利用分析軟體 Transcriptome Analysis Console version 3.0 software 內 *P* 值的設定，找出這些真正非常顯著有意義發生變異選擇性剪接的事件或基因(aberrant alternative splicing events or genes)，再進一步比對 DAVID (the Database for Annotation, Visualization and Integrated Discovery)或 KEGG 資料庫內的功能路徑(functional annotation analysis)，就可嘗試回答第三個問題：這些變異選擇性剪接(aberrant alternative

splicing)是分布在哪些基因，影響了那些訊息傳遞途徑。最後，本研究論文選擇幾個基因來進行 bench work，驗證本研究所觀察到的 array data，並嘗試探尋變異選擇性剪接(aberrant alternative splicing)與白血病發生(leukemogenesis)的關聯性。

2.2.2 RNA sequencing platform

近期隨著次世代定序技術的進步，RNA sequencing(RNA-seq)逐漸成為目前研究轉錄體的主要工具。因此本研究論文收錄了新診斷 *de novo*，non-M3 AML 患者的骨髓檢體，及臍帶血樣本的檢體。本研究從 AML 骨髓細胞及臍帶血細胞中抽取了 RNA。接著使用 ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE)和 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA)對 RNA 的濃度和完整性進行了評估。然後，本研究從每個樣本抽取四微克(4 µg)RNA，進一步純化成 mRNA，又進一步轉化成 cDNA library。接著使用 2100 Bioanalyzer system 和 real-time polymerase chain reaction system 評估 cDNA library 品質。合格的 library 接著使用 Illumina NovaSeq 6000 平台進行序列讀取(sequencing)，最後會得到 150 bp 的 paired-end reads。本研究論文委託 Genomics (BioSci & Tech Co., New Taipei City, Taiwan)來完成上述 RNA-seq 的實驗步驟。另外，本研究論文也嘗試分析 TCGA AML 的 RNA-seq 原始資料，作為驗證群體(validation cohort)。所有本研究中的 mRNA 序列資料已經存儲在 NCBI BioProject 數據庫中(存取編號：PRJNA921731、PRJNA924552、PRJNA928097、PRJNA928392、PRJNA929107)。

接著本研究論文使用了 fastp 軟體版本 0.21.0(<https://github.com/OpenGene/fastp>)來檢查 RNA-seq 原始資料的品質。本研究去除了樣本中的轉接子污染(adapter contamination)和低品質的序列讀取(low-quality reads)。接下來，本研究論文使用 SpliceSeq 軟體(版本 2.1)來分析 RNA-

seq 資料。這個軟體是由 MD Anderson 癌症中心開發的網路公開資源，主要用於分析選擇性剪接事件(alternative splicing events)。首先，將樣本 reads 與從人類參考基因組(human reference genome hg38) (UCSC Genome Browser 資料庫)構建的基因選擇性剪接示意圖(splice graphs)進行比對。每個基因的 alternative splicing event 藉由 "percentage spliced in (PSI)" 的值來量化，該值範圍在 0 到 1 之間。PSI 值是通過計算該基因在七種 alternative splicing 類型(包括 exon skipping, alternative donor sites, alternative acceptor sites, retained intron, mutually exclusive exons, alternate terminator, and alternate promoter)的剪接位點上(splice site)，所轉錄的 mRNA reads 比例來測量的。以一個包含三個 exons(E1、E2 和 E3)的基因為例，如果 isoform 1 和 isoform 2 分別包含 E1 + E2 + E3 和 E1 + E3，若該基因在 E2 處發生 exon skipping，且 PSI 值為 0.5，即表示在轉錄體中該基因的 mRNA transcripts 有 50% 包含 exon E2。因此，E2 的 PSI 值越高，mRNA transcripts 主要是 isoform 1；相反，E2 的 PSI 值越低，mRNA transcripts 主要是 isoform 2 (圖 8)^{50,51}。為了確定該基因真的有發生 alternative splicing event(同時有 isoform1 及 isoform2)，因此，我們將 PSI 值訂在 0.1 和 0.9 之間的 alternative splicing event 視為可進一步分析的對象。RNA sequencing 的研究流程圖如圖 9。

2.2.3 細胞實驗

藉由上述的 microarray 或 RNA sequencing 平台，本研究論文期待找到一個與臨床預後有相關的基因選擇性剪接事件(prognosis associated alternative splicing event)，作為本研究論文接下來研究的對象。進一步利用細胞分子生物學技術來探討此基因的 aberrant alternative splicing event 在血癌上的致病機轉，是否會改變哪些重要的生理功能路徑，因而進一步影響了病患的臨床預後。

本研究論文將選用 OCI-AML3 血癌細胞株做進一步實驗。假設所要研究的

基因 G 因選擇性剪接產生兩個 isoforms，分別是 g1 及 g2，本研究論文的細胞實驗包含了以下數個重要步驟：

a. 建立體外試驗系統(*in vitro* assay system)，透過轉殖入 isoform g1 或 g2 的互補的 DNA (complementary DNA，cDNA)到 OCI-AML3 細胞株內：

本研究論文利用 PCR 的技術把 isoform g1 及 isoform g2 互補的 DNA (complementary DNA，cDNA)從人骨髓(BM)正常細胞中進行放大。隨後，isoform g1 cDNA 和 isoform g2 cDNA 將分別被 clone 到 pTRIP_dU3_IRES_GFP 慢病毒載體(lentiviral vector)中(圖 10A)。之後分別把攜帶不同 cDNA 的病毒載體暫時 transfect 到 HEK293T 慢病毒包裝細胞中(lentiviral packaging cell line)，將生成攜帶 isoform g1 cDNA，isoform g2 cDNA 或空載體(empty vector)的慢病毒。轉染(transfect)後的三天，通過收集 HEK293T 細胞培養液來收集慢病毒。培養液通過 0.22 微米的針筒過濾器，然後通過 Amicon Ultra-15 離心式過濾器進行濃縮。濃縮的病毒液再 transfect 到 OCI-AML3 細胞中，並在培養基下進行培養。由於病毒載體序列上會轉譯出綠螢光蛋白，所以經過 72 小時後，可以透過流式細胞技術(flow cytometry)選出(sort)會發出綠色螢光分子的細胞用於進一步的實驗。

b. 使用短髮小核糖核酸(short hairpin RNA，shRNA)的技術，先抑制 (knockdown) OCI-AML3 細胞株內原本內源性 G 基因的表達，之後再轉殖入 isoform g1 或 g2 的互補的 DNA (complementary DNA，cDNA)到 OCI-AML3 細胞株內：

為了避免來自基因 G 內源性 isoforms 的干擾，本研究論文將先進行血癌細胞 G 基因表現的抑制實驗(gene knockdown)，然後再分別把轉殖入血癌細胞中 isoform g1 或 g2 互補的 DNA (complementary DNA，cDNA)表達，這樣一來，更可以好好地研究這兩個 isoforms 的生物學效應。

本研究論文先設計能夠針對 G 基因的 shRNA 序列，稱為 shG，且確保這些



shRNA 序列是特異性的，並將此 shRNA 序列設計在能針對 G 基因的 3' UTR(未轉譯區)區域，主要是要 knockdown 內源性 G 基因的表達，但不會干擾到接下來轉殖入的 isoform g1 或 g2 cDNA 表達。本研究論文另外設計了 shRNA 可針對 *LacZ* 基因(稱為 sh*LacZ*)當作對照組。接下來，將 shG 和 sh*LacZ* 克隆(clone)到 pLKO_TRC019 慢病毒載體中(lentiviral vector)(圖 10B)。然後，將這些病毒載體通過轉染(transfect)到 HEK293T 慢病毒包裝細胞株中(lentiviral packaging cell line)，生成攜帶 shG 或對照組(sh*LacZ*)的慢病毒。轉染後的三天，通過收集 HEK293T 細胞的培養液，收集慢病毒。培養液通過 0.22 微米的針筒過濾器(syringe filter)，然後通過 Amicon Ultra-15 離心式過濾器進行濃縮。濃縮的病毒液再 transfect OCI-AML3 細胞中，並在培養基下進行培養 72 小時。

接著利用 PCR 的技術把 isoform g1 及 isoform g2 互補的 DNA (complementary DNA, cDNA)從人骨髓(BM)正常細胞中進行放大。隨後，isoform g1 cDNA 和 isoform g2 cDNA 將分別被 clone 到 pTRIP_dU3_IRES_mCherry 慢病毒載體中(圖 10C,D)。之後分別把攜帶不同 cDNA 的病毒載體暫時 transfect 到 HEK293T 慢病毒包裝細胞中(lentiviral packaging cell line)，將生成攜帶 isoform g1 cDNA，isoform g2 cDNA 或空載體(empty vector)的慢病毒。轉染後的三天，通過收集 HEK293T 細胞培養液來收集慢病毒。培養液通過 0.22 微米的針筒過濾器，然後通過 Amicon Ultra-15 離心式過濾器進行濃縮。濃縮的病毒液再 transfect 到 OCI-AML3 細胞中，並在培養基下進行培養。由於病毒載體序列 pLKO-TRC019-shRNA-IRES-GFP 會轉譯出綠螢光蛋白，而 pTRIP_dU3_IRES_mCherry 會轉譯出紅螢光蛋白，所以經過 72 小時後，可以透過流式細胞技術(flow cytometry)選出(sort)會同時發出綠色及紅色螢光分子的細胞用於進一步的實驗。

c. 西方墨點法(Western Blotting)：

把上述 OCI-AML3 細胞裂解(lyse)，這些細胞攜帶了 shRNA 及不同 isoform

g cDNA 的慢病毒，這些細胞總裂解物(total lysate)將用於 western blot 分析，用來確認 gene knockdown 及 isoform g cDNA expression 是否有成功。

d. 分析原有的細胞生理路徑是否有改變(biological behavior changes analysis)

→ 將上述處理過的 OCI-AML3 細胞株經培養收集後做以下的實驗：

I. 細胞增殖和存活試驗(cell proliferation and viability assays)

將 OCI-AML3 經培養後的細胞液以每孔 5×10^3 個細胞的密度加入到 96 孔板中(96 well plate)，然後將 MTT(3-(4,5-Dimethylthiazol-2-yl)-2,5-Diphenyltetrazolium Bromide)溶液(5mg/ml) 200 uL 加入到每個孔中，接著在 37°C 下培養 4 小時，然後在 570 奈米(nm)波長下測量光密度，這個試驗將持續進行 5 天。在這個 MTT 試驗中，活細胞中的活性輔酶會將 MTT 還原成紫色結晶形式，形成可檢測的紫色產物。通過測量光密度，可以評估細胞的增殖和存活情況。在培養的五天內，可以觀察細胞的生長情況，並檢測因不同處理組之間所造成的細胞生長及活性是否有差異。

II. 細胞週期分析(cell cycle analysis)

首先，選用將上述經過改造後能穩定表達 isoform g1 或 isoform g2 的 1×10^5 個 OCI-AML3 細胞，將這些細胞在 4°C 下用 70%乙醇固定並過夜。然後將細胞用 RNase A(20 ug/mL)處理，接著使用 50 ug/mL 的草薺素酮碘化物(propidium iodide, PI)進行染色。最後使用 BD FACSVerse 流式細胞儀並使用軟體 FlowJo software version 10 進行細胞週期分佈分析。通過流式細胞儀及軟體分析，可以得到 G 基因不同異構物(isoform)表達的細胞在 G0、G1、S、G2 和 M 階段的比例，從而進一步瞭解 isoform g1 和 isoform g2 對細胞增殖和生長的調節機制。這對於本研究論文探索 G 基因選擇性剪接的變異在白血病致病機轉中具有重要的意義。

III. 分析 G 基因因變異選擇性剪接對其他基因的表達是否有影響(analyzing downstream genes regulated by G gene)

將上述處理過的 OCI-AML3 細胞株收集 RNA，接著將對這些 RNA 進行全轉錄組分析，使用 RNA sequencing 技術來進行。之後就可挑選有興趣的基因進一步分析其表達量是否會受到 G 基因的影響。



IV. 化療抗藥性檢測(chemoresistance detection)

使用經過上述改造後能穩定表達 isoform g1 或 isoform g2 的 OCI-AML3 細胞以每孔 1×10^4 個細胞的密度加入到 96 孔板中。然後將逐漸增加不同濃度的化療藥物 Cytarabine (Ara-c) 加入到每個孔中。接著在細胞培養 48 小時後分別使用上述提到的 MTT 試驗來測量細胞的存活率。

V. 細胞遷移實驗(cell migration assay)

使用 24 孔轉膜插入板(transwell insert)，利用直徑 $5\mu\text{m}$ 的聚碳酸酯薄膜(polycarbonate membrane) (Cell Biolabs, San Diego, CA) 進行細胞遷移實驗。在上室(upper chamber)中，我們加入 1×10^5 個經改造後的 OCI-AML3 細胞，並添加 200 μL 的無血清培養液(serum-free medium)，而下室(lower chamber)則收到 300 μL 的無血清培養基(serum-free conditioned medium)。然後培養在 37°C 下，5% CO_2 的溼氣中持續 4 小時。實驗結束後，我們將遷移的細胞從薄膜上分離，然後使用 Invitrogen 的專利 CyQuant® GR 染料進行檢測。

VI. 生物信息學分析(Bioinformatics analysis)

在上述研究中，當獲得了 RNA-Seq 數據之後，除了可以分析轉錄體中各個基因的表達量，本研究論文也會使用基因集豐富分析(Gene Set Enrichment Analysis, GSEA)方法⁵²。GSEA 是一種用於分析基因表達數據的生物信息學方法，它通過將基因分組成相關的功能性基因集合(例如某個生理功能路徑或疾病相關的基因集合)，來評估在特定生物學條件下，哪些功能性或致病性基因集合在基因表達中是顯著的，有助於提供整體性的生物學意義，並讓本研究計畫有機會深入研究與白血病相關的重要功能性基因集合。

2.3 統計分析

本研究中的統計分析是使用軟體 SPSS version 22 (IBM Corporation, Armonk, NY)來進行。本研究論文採用了 Kolmogorov-Smirnov 檢驗來驗證研究結果中所分析的變量是否呈現常態分佈。利用 t 檢驗來比較兩群連續變量間的差異顯著性，ANOVA 檢驗用於比較三群或更多連續變量的分佈。Pearson 相關分析用於檢定研究的變量與臨床和生物學數據之間的關係。使用 Kaplan-Meier 方法及 log-rank 檢驗對病患進行存活分析。單變量和多變量 Cox 比例風險回歸分析用於確定特定與生存相關的研究變量。 P 值 <0.05 被視為具有統計學意義。

第三章、結果(Results)



3.1 基因選擇性剪接在骨髓造血不良症候群(MDS)病患上的角色

3.1.1 本研究論文 MDS 患者的臨床基本資料(Clinical characteristics of the MDS patients)

本研究共收集了從 1991 年 11 月至 2010 年 12 月在台大醫院新診斷的 176 名 MDS 患者，大多數(84.1%)的患者僅接受支持性治療(supportive care)。另外，4 名患者(2.3%)接受了高強度化學治療(intensive chemotherapy)，2 名患者(1.1%)接受了去甲基化劑治療(hypomethylating agent)，22 名患者(12.5%)進行了異體造血幹細胞移植(allogeneic hematopoietic stem cell transplantation)。這些病患的詳細資料如表 1。此外，本研究還收集了另一群在 2011 年 1 月至 2012 年 5 月期間在台大醫院被新診斷的 31 名骨髓造血不良症候群(MDS)患者族群，作為驗證群體(validation cohort)。

3.1.2 整個轉錄體上的變異選擇性剪接(Aberrant alternative splicing of whole transcriptome)

Affymetrix Human Transcriptome Array 2.0 (HTA 2.0)上可偵測 67539 個基因，但整體來說，骨髓造血不良症候群患者與正常對照組的骨髓細胞共只有 **52730 (78.1%)** 的基因在 HTA 2.0 array 上同時被表達；更進一步，有 **17565 (33%)**在 HTA 2.0 array 上這些同時有被表達的基因，發生了變異選擇性剪接(aberrant alternative splicing)，大約 **88%**是影響了蛋白質編碼區域 (protein coding region)，**12%**是蛋白質非編碼區域 (protein noncoding region) (表 2)；此外這些 aberrant alternative splicing events 可歸納成五種常見的剪接類型(splicing event type)，以 exon skipping 的類型最常見(如圖 11 所示)。

3.1.3 變異選擇性剪接與臨床表徵的相關性(Associations of clinical features with

global aberrant alternative splicing)



本研究論文定義 **aberrant alternative splicing score** 為轉錄體中所有 aberrant alternative splicing events 與所 involve 的基因(aberrant alternative splicing genes)相除。此數值代表著一個變異選擇性剪接基因平均含有多少個變異選擇性剪接事件。進一步，本研究發現 aberrant alternative splicing score 與 the total number of aberrant alternative splicing events 呈現高度相關性(圖 12)。如本研究論文之前的假說，因此若 aberrant alternative splicing score 數值越高，則轉錄體中所有 aberrant alternative splicing events 也會越多，即代表整個轉錄體整體 mRNA 變異選擇性剪接模式的程度(the degree of global aberrant alternative splicing pattern)就越厲害，反之越小代表該轉錄體僅輕微的發生了變異選擇性剪接。分別把這 176 個 MDS 病患的 aberrant alternative splicing score 計算出來，本研究發現 MDS 患者 aberrant alternative splicing score 的數值介於 2.05-5.48 (中位數 2.6)，與 normal donors 間有很顯著的差異(圖 13A)。進一步比較不同 FAB subtype(圖 13B)及 WHO risk(圖 13C) MDS 病患各組間的 global aberrant alternative splicing pattern 的差異(ANOVA $P = 0.068$; ANOVA $P = 0.448$ respectively)，不同 genetic mutation status (Splicing factor mutations, Epigenetic mutations, Neither splicing factor nor epigenetic mutations)(圖 13D)之間的差異(ANOVA $P = 0.708$)，本研究發現 global aberrant alternative splicing pattern 並不會因不同的 MDS sub-classification 而有顯著的差異。

接著本研究使用每個病患的 aberrant alternative splicing score 與存活關係的 ROC 曲線來預測 MDS 患者的生存情況，並獲得了一個最佳區分患者生存的截斷點值(cut-off point value)，為 2.45。根據這個值，本研究將變異選擇性剪接的模式定義為低和高。在這個 176 個病患的 cohort 中，若 aberrant alternative splicing score ≥ 2.45 ，表示 high degree of global aberrant alternative splicing pattern，若 aberrant alternative splicing score < 2.45 為 low degree of global

aberrant alternative splicing pattern。研究結果顯示 high degree of global aberrant alternative splicing pattern 與男性性別和 *U2AF1* 基因突變之間存在顯著相關(表 3)。



3.1.4 變異選擇性剪接對臨床預後的影響(The prognostic impact of global aberrant alternative splicing)

本研究論文進一步分析 overall survival，發現 high degree of global aberrant alternative splicing pattern 會有統計顯著較差的預後，但反之預後較好(median OS: 21.7 months vs 69.9 months, $P = 0.009$)(圖 14A)；接著以 time to acute leukemia transformation 來看，在這個 176 個病患的 cohort 也有顯著的差異($P = 0.002$)(圖 14B)；除此之外，MDS 的病患若無任何 spliceosome gene mutations 及 epigenetic modification gene mutations，high degree of global aberrant alternative splicing pattern 的患者也會有較差的 overall survival (median OS: 42.4 months vs not reached, $P = 0.007$) (圖 14C)。進一步分析 validation cohort ($n=31$)，也可以驗證 high degree of global aberrant alternative splicing pattern 會有統計顯著較差的預後(圖 14D)。最後經由多變數分析(multivariate analysis)，本研究論文的資料顯示 global aberrant alternative splicing pattern 可能是 MDS 的一個新的獨立預後因子 ($P = 0.042$) (independent of age, IPSS-R risk score, and mutations of *SRSF2*, *ZRSR2*, *ASXL1*, *TP53*, and *EZH2*) (表 4)。

3.1.5 變異選擇性剪接對細胞生理路徑的影響

為了更進一步探討這些變異選擇性剪接對細胞生理路徑的影響，本研究使用 DAVID (The Database for Annotation, Visualization and Integrated Discovery) 資料庫對受變異選擇性剪接影響的基因(aberrant alternative splicing genes)進行了分析。本研究確定了 71 個訊號傳遞路徑(signaling pathways)因變異選擇性剪接而

受到顯著影響，其 P 值小於 0.001，虛假發現率(FDR)小於 0.05。本研究論文按照 fold enrichment 列出了前 20 個訊號傳遞路徑(表 5)，並發現這些受變異選擇性剪接影響的細胞生理路徑與細胞增殖，黏附和蛋白質降解(cell proliferation, adhesion, and protein degradation)的調控有關。這些發現可證明在 MDS 患者中，正常的選擇性剪接機制被破壞，而信號網絡中的變異選擇性剪接似乎是普遍存在的，影響了許多正常的生理路徑。

3.1.6 驗證在 microarray 平台上所預測的變異選擇性剪接

為了驗證在 microarray 平台上所預測的變異選擇性剪接，本研究選擇了 3 個基因的 transcripts 進行驗證。*AUP1* 基因位於 2P13.1，包含 12 個外顯子(exons)；在 HTA 2.0 平台上，大多數 MDS 患者(71.6%) 此基因的 transcript(NM_181575)顯示出異常保留第 9 內含子(aberrant retention of intron 9)(mean SI level 3.54)。進一步做 RT-PCR，與 3 名正常對照者(normal donors)相比，4 名 MDS 患者中的 *AUP1* transcript 在 204 bp 處明顯地有額外的 band 存在(圖 15A)。本研究透過直接定序(direct sequencing)確認了這些 MDS 患者中 *AUP1* 第 9 內含子的保留(圖 15B-D)。另外，*PVRL2* 基因(NM_001042724)位於 19q13.32，包含 9 個外顯子；microarray HTA 2.0 data 顯示本研究族群大多數的 MDS 患者(98.3%) *PVRL2* 基因發生 exon 4 skipping。*GRIK5* 基因(NM_002088)位於 19q13.2，包含 19 個外顯子；microarray HTA 2.0 data 顯示本研究族群大多數的 MDS 患者(84.1%) *GRIK5* 基因發生 exon 19 skipping。透過 RT-PCR 方法，本研究驗證了這 4 名 MDS 患者中 *PVRL2* 基因 exon 4 skipping 以及 *GRIK5* 基因 exon 19 skipping(圖 15E,F)。此外藉由 Geneious 軟體的預測(圖 16)，MDS 患者 *AUP1* 的變異選擇性剪接將導致蛋白質截斷(protein truncation)。

3.2 基因選擇性剪接在急性骨髓性白血病(AML)病患上的角色

3.2.1 本研究論文所研究 AML 患者的臨床基本資料(Clinical characteristics of the AML patients)

本研究收集了 341 名在台大醫院新診斷的非 M3 型原發性 AML 患者，這些患者需有足夠的檢體可進行 RNA 序列研究，並且有完整的臨床資料。他們都接受了標準的誘導化療，包括 idarubicin 每日 12 mg/m²(第 1-3 天)和 cytarabine 每日 100 mg/m²(第 1-7 天)，如果患者在引導治療後可達到完全緩解(complete remission)，之後會進行 2-4 個週期的高劑量 cytarabine(high dose Ara-C，每 12 小時 2000 mg/m²，第 1-4 天，總共 8 劑)鞏固化療。在這個研究族群有近一半的患者(44%)有接受了異體造血幹細胞移植(allogeneic hematopoietic stem cell transplantation)。詳細的病患臨床特徵如表 6。另外，本研究也納入了 141 名非 M3 TCGA_AML 患者當作驗證群體(validation cohort)，這些病患的基本臨床資料如表 7。

3.2.2 整個轉錄體上的基因選擇性剪接 (Identification of global splicing)

本研究使用 SpliceSeq 軟體分析了轉錄體中 19036 個基因的選擇性剪接事件(alternative splicing event)。本研究分別計算了兩個 AML 病患族群(NTUH 和 TCGA)中每個血癌患者骨髓內的血癌細胞，在這些基因中所有基因選擇性剪接事件的數目(PSI 值在 0.1 到 0.9 之間的選擇性剪接事件)以及六個臍帶血樣本內 CD34⁺/CD38⁻細胞基因選擇性剪接事件的數目。在 NTUH_AML 病患族群中，每個患者所有基因選擇性剪接事件的數目介於 13520 到 25168，中位數為 18869，這些選擇性剪接事件共來自於 6438 個基因；而在 TCGA 病患族群中，基因選擇性剪接事件的數目範圍從 11750 到 20794，中位數為 17832，共來自於 6404 個基因。在六個臍帶血樣本中，CD34⁺/CD38⁻細胞基因選擇性剪接事件的數目介於從 6327 到 10708，中位數為 7497，共來自於 3367 個基因(圖 17)。總結，在 AML 細胞的轉錄體中，大約有 34%(6404-6438 of 19036)基因(包括

NTUH 和 TCGA 病患族群)發生了選擇性剪接事件(基因有選擇性剪接事件 PSI 值在 0.1 到 0.9 之間)，相較於臍帶血 CD34⁺/CD38⁻的造血細胞只有 18%(3367 of 19036)，有顯著的差異(圖 17)。

進一步分析，在兩個 AML 病患族群中，具有 *TP53* 基因突變的 AML 細胞，所有基因選擇性剪接事件的數目顯著高於其他 AML 亞型(圖 18A,B)。在七種常見的 alternative splicing event 類型中(exon skipping, alternate donor sites, alternate acceptor sites, retained intron, mutually exclusive exons, alternate terminator, alternate promoter)不管是 AML 細胞或臍帶血 CD34⁺/CD38⁻的造血細胞，exon skipping 都是最常見的類型(圖 19)。另外，AML 細胞基因體的基因突變數量似乎與轉錄體內所有基因選擇性剪接事件的數目沒有相關性(圖 20, 21)。

3.2.3 轉錄體整體選擇性剪接模式與臨床表徵的相關性(Clinical features associated with differential global splicing patterns)

本研究發現 NTUH_AML 患者轉錄體中所有基因選擇性剪接事件的數目呈現近似鐘形曲線分佈(圖 22)，且根據 Kolmogorov-Smirnov 常態分佈檢驗($P = 0.083$)，本研究假設 AML 病患轉錄體中所有基因選擇性剪接事件的數目呈現常態曲線分佈。因此，轉錄體中基因選擇性剪接事件數目的中位數值與平均數值相似。為了平均地將患者分為兩組來進一步分析，本研究使用轉錄體中基因選擇性剪接事件數目的中位數值作為截點，將 AML 患者分為兩組，high or low degree of global alternative splicing pattern；在 NTUH_AML 病患中，low degree of global alternative splicing pattern 的患者白血球計數較高、乳酸脫氫酶(LDH)數值較高、有預後較好的染色體變化、完全緩解(CR)率較高以及較少誘導期死亡(induction death)的情況。global alternative splicing pattern 與性別、年齡或根據 FAB 分類定義的 AML 亞型之間沒有相關性。然而，在 WHO 分類第 5 版或

ICC 分類中，具有 *t(8;21)/RUNX1::RUNX1T1* 和 *inv(16)/CBFB::MYH11* 的 AML 病患與 low degree of global alternative splicing pattern 相關(表 6)。此外，*KIT*、*U2AF1*、*SMC3* 的基因突變，及 *CEBA* double mutations 與 low degree of global alternative splicing pattern 相關，而 *TP53* 基因突變則與 high degree of global alternative splicing pattern 相關(表 8)。然而，TCGA_AML 病患 low degree of global alternative splicing pattern 的臨床表徵與 NTUH_AML 病患有些不同，但 *TP53* 突變與 high degree of global alternative splicing pattern 也是有統計顯著關聯性(表 7,9)。

3.2.4 轉錄體整體選擇性剪接模式對臨床預後的影響(The clinical impact of differential global splicing patterns)

在 NTUH_AML 病患中，進一步發現 high degree of global alternative splicing pattern 相較於 low degree of global alternative splicing pattern 會有統計顯著較差的整體存活期(median overall survival：24.6 個月 vs 161.6 個月， $P = 0.001$)和疾病無復發存活期(relapse free survival：77.8 個月 vs not reached， $P = 0.035$)(圖 23A)。此外，若結合 global alternative splicing pattern 和 2022 European LeukemiaNet (ELN)風險分類更可進一步將 AML 患者區分為四組不同的預後(圖 23B,C)。最後經由單變數及多變數分析(univariate and multivariate analysis)，無論病患的年齡，白血球計數和 2022 ELN 風險分類如何，global alternative splicing pattern 和 the number of global alternative splicing events 都是顯著獨立的預後因子(表 10)。接著我們以 TCGA_AML 病患當驗證族群(validation cohort)，high degree of global alternative splicing pattern 相較於 low degree of global alternative splicing pattern 仍然會有統計顯著較差的整體存活期(median overall survival：366 天 vs 577 天， $P = 0.013$)(圖 24)。

3.2.5 基因選擇性剪接對細胞生理路徑的影響

為了更進一步探討這些基因選擇性剪接對細胞生理路徑的影響，本研究使用 DAVID (The Database for Annotation, Visualization and Integrated Discovery) 資料庫對受選擇性剪接影響的基因(alternative splicing genes)進行了分析。我們按照 fold enrichment 列出了前 20 個訊號傳遞路徑(其 P 值小於 0.005，虛假發現率(FDR)小於 0.05)，發現這些受到基因選擇性剪接變化影響的細胞生理路徑包括 G 蛋白調節(G protein regulation)，蛋白質降解，細胞週期以及 DNA 損傷/修復相關的途徑(表 11)。

3.3 *NFYA* 基因選擇性剪接對急性骨髓性白血病(AML)病患臨床預後的影響

3.3.1 *NFYA* 基因的表現量和選擇性剪接與 AML 病患臨床表徵的關係及對預後的影響

經由上述急性骨髓性白血病患者骨髓的血癌細胞 RNA-seq 資料，本研究發現有許多基因的選擇性剪接型式與正常骨髓造血細胞有明顯的不同。本研究注意到了一個有趣的基因 Nuclear Transcription Factor Y Subunit Alpha (*NFYA*)，並選擇此基因作為進一步研究的候選基因。

本研究論文分析了 341 個急性骨髓性白血病(AML)患者血癌細胞中 *NFYA* 基因表達及其基因選擇性剪接型式，另外也分析了 6 個臍帶血樣本中的造血幹細胞，以此作為對照組，比較兩者在 *NFYA* 基因表現及基因選擇性剪接型式間的差異。本研究進一步使用毛細管螢光電泳(capillary fluorescence electrophoresis)來驗證 AML 患者 RNA-seq 資料中基因選擇性剪接型式。*NFYA* 基因主要有兩個因選擇性剪接造成的轉錄異構體(transcript isoforms)：*NFYA* 長型(*NFYA-L*)包含完整的 10 個 exons，而 *NFYA* 短型(*NFYA-S*)在 pre-mRNA 剪接時 exon 3 skipping。本研究的實驗結果顯示，*NFYA* 基因的選擇性剪接可以在

AML 血癌細胞株和 NTUH_AML 患者的血癌細胞中得到驗證(圖 25A-E)。有趣的是，經由毛細管螢光電泳實驗及進一步的定序(Sanger sequencing)，本研究檢測到了一個新的轉錄異構體(a novel transcript isoform) p.Q23 del 相較於主要的轉錄異構體 *NFYA*-L(圖 25D-G)。此外，本研究發現，AML 血癌細胞若與正常臍帶血造血幹細胞相比，*NFYA* 基因在 AML 血癌細胞中被過度表達，而且 *NFYA* 基因的選擇性剪接型式在 AML 血癌細胞與正常臍帶血造血幹細胞間有很大的差異(圖 26A,B)。*NFYA*-S 在正常造血幹細胞(臍帶血)中似乎是主要的轉錄異構體。相較於正常造血幹細胞，AML 血癌細胞中的 *NFYA*-L/*NFYA*-S 比率會顯著增加(圖 26B)。接著若把病患分別以 *NFYA*-L 及 *NFYA*-S 表達量的中位數當切點，把病患區分成高 *NFYA*-L 但低 *NFYA*-S (*NFYA*-L predominance pattern)及高 *NFYA*-S 但低 *NFYA*-L 表達(*NFYA*-S predominance pattern)的兩群群病患，*NFYA*-L predominance pattern 的 AML 患者表現出較高的白血球(WBC)計數，較高的乳酸脫氫酶(LDH)，較高的比例為 FAB M4 或 M5 亞型，以及 *NRAS* 和 *U2AF1* 基因突變的頻率較高。相反地，*NFYA*-S predominance pattern 表達的患者則傾向具有較低的 WBC 計數，較低的 LDH，FAB M2 亞型的比例較高，以及較高的 *CEBA*_{double} 和 *GATA2* 基因的突變(表 12,13)。更進一步，本研究分析這 341 個接受標準高強度化療的 AML 患者存活預後，首先發現 *NFYA* 基因整體表達量及 *NFYA*-L 表達量的高低對 AML 患者的預後沒有顯著影響(圖 27A,B)。然而，*NFYA*-S 的表達量高低具有預後價值(圖 27C)。*NFYA*-S 表達量較高的 AML 患者在誘導化療後完全緩解的比例較高，上述結果可能與預後效應有相關(表 14)。接著，若 AML 癌細胞中有較高的 *NFYA*-L 但較低的 *NFYA*-S 表達(*NFYA*-L predominance)在接受標準化療後會有較差的存活率。相反，有較高的 *NFYA*-S 但較低的 *NFYA*-L 表達(*NFYA*-S predominance)與標準化療後更高的緩解率和較好的存活率相關。這些結果在一個獨立的血癌群體 TCGA_AML 群體中得到進一步驗證。該群體包括 139 個 AML 患者，若診斷時有較高的 *NFYA*-L 但較低的

NFYA-S 表達(*NFYA*-L predominance)也會與較短的整體存活期(overall survival)相關(圖 28A,B)。之後本研究對這 341 位 NTUH_AML 患者做單變量和多變量存活分析，結果顯示 *NFYA* 的剪接型式在其他因素如年齡、European LeukemiaNet (ELN) 2022 risk classification 等情況下仍保持其作為獨立預後因子的重要性(表 15)。

3.3.2 *NFYA* 基因選擇性剪接對 AML 病患血癌細胞潛在的生物學效應

接下來本研究進行了 functional enrichment analysis，以探求 NTUH_AML 患者 *NFYA* 基因選擇性剪接變異潛在的生物學效應。本研究將注意力集中在上述兩個不同的患者群體上：高 *NFYA*-L 但低 *NFYA*-S (*NFYA*-L predominance)及高 *NFYA*-S 但低 *NFYA*-L 表達(*NFYA*-S predominance)。整個基因體基因表達的熱圖(heat map)顯示出這兩個患者群體明顯不同的基因表達模式(圖 29)。基因集豐富分析(GSEA)顯示，在 *NFYA*-S predominance 的患者族群中，我們發現一些與造血幹細胞(HSC)有關的基因被高度表達，此外與細胞週期相關的基因也被高度表達(圖 30A-C)。另一方面，在 *NFYA*-L predominance 的患者族群中，則與 mixed lineage leukemia (MLL)相關、mitogen-activated protein kinase (MAPK)訊號傳遞途徑相關和單核細胞遷移(mononuclear cell migration)相關的基因被高度表達有關(圖 30D-F)。

3.3.3 *NFYA* 基因選擇性剪接在白血病細胞株中的生物效應

為了探討 *NFYA* 剪接異構體對白血病細胞產生的影響，本研究利用體外細胞模型進行實驗(*in vitro* cellular models)。本研究利用接受不同 *NFYA* 剪接異構體 lentiviral vector 處理後的 OCI-AML3 細胞株來進一步做實驗(圖 31A,B; 32A,D)。首先，本研究分析在不同 OCI-AML3 細胞株中過度表達不同 *NFYA* 剪接異構體對其增殖能力的影響。在 OCI-AML3 細胞中過度表達 *NFYA*-S 將導致

細胞增殖顯著增加，而過度表達 *NFYA*-L 對細胞增殖沒有明顯影響(圖 32B)。進一步，本研究評估這些不同 OCI-AML3 細胞株對 Ara-C 化療藥物的敏感性。在 48 小時不同濃度的 Ara-C 處理後，過度表達 *NFYA*-S 的 OCI-AML3 細胞表現出比對照組和 *NFYA*-L 過度表達細胞有更大的化療敏感性(圖 32C)。

此外，在細胞實驗中，為了消除 *NFYA* 原本內源性異構體的干擾，本研究先使用 shRNA 對 *NFYA* 基因進行敲除(knockdown)。在敲除後，再利用轉殖入 *NFYA* 轉錄異構體 cDNA 來表達 *NFYA*-S 或 *NFYA*-L。這種方法使本研究能夠更有效地研究這兩種剪接異構體的生物效應(圖 32D)。在本研究的實驗中，可以觀察到若只對 *LacZ* control 基因敲除，結果會如同先前的實驗結果(圖 32B,E)。但當 *NFYA* 基因被敲除時，OCI-AML3 細胞的增殖和遷移能力(migration ability)明顯減弱(圖 32F, 33)。然而，在表達 *NFYA*-S 後，OCI-AML3 細胞的增殖顯示出恢復的跡象，而在表達 *NFYA*-L 後卻未觀察到這種情況(圖 32G)。而且在 *NFYA* 敲除的條件下，*NFYA*-S 和 *NFYA*-L 的表達均能使 OCI-AML3 細胞的遷移能力恢復(圖 33)。這些結果顯示 *NFYA* 在調控 OCI-AML3 細胞增殖方面具有特定的作用，且這細胞增殖效應歸因於 *NFYA*-S 異構體。

另外，在本研究的細胞週期實驗分析中，觀察到 *NFYA* 基因的敲除(knockdown)會導致 OCI-AML3 細胞 S 期細胞比例顯著減少。然而，將 *NFYA*-S 表達後，S 期細胞的比例似乎明顯增加。但在表達 *NFYA*-L 後，S 期細胞的比例增加幅度較不明顯(圖 34A,B)。上述研究結果可以發現 *NFYA*-S 和 *NFYA*-L 對 OCI-AML3 細胞週期中 S 期的調控作用不同。

最後，本研究分析經 *NFYA* 基因的敲除後再分別表達 *NFYA*-S 或 *NFYA*-L 的 OCI-AML3 細胞 RNA-seq 資料，基因集富集分析(GSEA)顯示主要表達 *NFYA*-L 的細胞會抑制細胞週期中進展至 G2/M 調控的相關基因表達，這顯示 *NFYA*-L 表達將導致細胞週期進展的抑制，因此減少細胞增殖(圖 35A)，但會活化細胞代謝路徑 oxidative phosphorylation 的相關基因表達，導致血癌細胞對化療藥物

Ara-C 產生抗藥性(圖 35B)⁵³。



第四章、討論(Discussion)

4.1 基因選擇性剪接在骨髓造血不良症候群(MDS)與急性骨髓性白血病(AML)病患上的臨床效應

基因選擇性剪接對正常造血過程至關重要；這種生理過程的變異可能與白血病的發生有關。轉錄體中正常的基因選擇性剪接發生變異已在許多疾病中被報告，包括良性和惡性疾病，如囊性纖維化(cystic fibrosis)、阿滋海默症(Alzheimer disease)和許多癌症，也包括血液惡性疾病^{24,32-35}。一些特定基因的變異選擇性剪接已被認為可當作癌症的診斷或預後生物標記。儘管特定基因的變異選擇性剪接在骨髓性血液惡性疾病中已有報告^{16,17}，但這類疾病整個轉錄體整體 mRNA 變異選擇性剪接模式(global mRNA aberrant alternative splicing pattern)尚未得到很好的研究。因此本研究論文是第一篇探討 MDS 及 AML 病患轉錄體整體 mRNA 選擇性剪接模式的改變(changes in global alternative splicing patterns)所造成疾病的臨床效應。

在本研究論文中，使用微陣列(microarray)和 RNA-seq 平台研究了骨髓造血不良症候群(MDS)和非 M3 急性骨髓性白血病(non-M3 AML)患者的選擇性剪接(alternative splicing)，結果顯示選擇性剪接在疾病的進展中扮演重要的角色。在 MDS 中，相較於正常骨髓造血細胞，整個轉錄體整體 mRNA 變異選擇性剪接模式的程度(the degree of global aberrant alternative splicing pattern)會促進白血病的發生，影響病患的預後⁵⁴。而在 AML 患者中，整個轉錄體整體 mRNA 選擇性剪接模式的程度(the degree of global alternative splicing pattern)會影響白血病患接受標準高強度化療(standard intensive chemotherapy)的療效，因而影響病患的預後⁵⁵。這個研究論文結果顯示選擇性剪接(alternative splicing)在白血病發生中具有重要的角色，它可以反映轉錄體的複雜性和不穩定性，導致不同的疾病預後⁴⁴。

本研究論文在 MDS 的研究中使用的研究檢體為骨髓單核細胞(bone marrow

mononuclear cells)，而不是同質性更高，如經過細胞分選(cell sorting)的 CD34⁺細胞。一個質疑是骨髓單核細胞中，因不同類別的細胞，可能各基因表達和選擇性剪接型式差異很大，而產生異質性。然而，在最近的研究報告「MDS 基因表達和白血病轉化風險的相關性」，Shiozawa 等人發現，根據骨髓單核細胞的基因表達方式與經過細胞分選純化的 CD34⁺前驅細胞相似⁵⁶。但進一步的研究仍是必須的。

此外，本研究論文在 AML 的研究中使用的研究檢體也是整個骨髓單核細胞(bone marrow mononuclear cells)，而不是細胞分選後的白血病芽細胞(sorted leukemic blasts)。但在本研究的 NTUH_AML 病患中，骨髓內血癌芽細胞百分比為 20%至 98%，中位數為 67%，因此大多數(72%)病患骨髓檢體中的白血病芽細胞百分比(leukemia blast percentage)大於 50%。另外，本研究論文在 AML 的研究中顯示 AML 細胞中 34%的表達基因有發生明顯的選擇性剪接，顯著高於臍帶血造血幹細胞(只有表達基因的 18%)。AML 細胞和臍帶血細胞(cord blood cells)之間整個轉錄體整體 mRNA 選擇性剪接模式(global mRNA alternative splicing pattern)有顯著的差異(圖 17)。臍帶血前驅細胞(cord blood precursor cells)和 AML 細胞在整個轉錄體整體 mRNA 選擇性剪接模式的差異可能也與老化因素有關。但我們此研究中選擇臍帶血樣本而非年齡相符的對照組 CD34⁺細胞，原因包括(1)臍帶血樣本中的 CD34⁺細胞比未受 G-CSF 刺激且年齡相符的健康對照組骨髓中的 CD34⁺細胞更豐富。(2)很難找到未接受 G-CSF 刺激的骨髓捐贈者骨髓細胞檢體。沒有年齡相符的健康對照組確實是我們研究的一個限制，但之前有研究報告指出，儘管老化被認為會對基因選擇性剪接產生重大變化，但該研究指出基因選擇性剪接觀察到的戲劇性差異可能也與疾病相關的因素有關⁵⁷。

即使上述因研究檢體量與經費等因素導致本研究論文使用的檢體品質可能存在著一些問題，但本研究論文的研究結果顯示選擇性剪接的變異在血癌致病

過程中確實扮演著重要的角色，進一步推測這些變異選擇性剪接(aberrant alternative splicing)可能會改變轉錄體的複雜性和不穩定性，因而導致疾病的進展⁴⁴。另一方面，經由生物資訊學分析(bioinformatics analysis)，本研究發現 MDS 病患這些發生變異選擇性剪接的基因可能會涉及到與細胞增殖、細胞黏附和蛋白質降解相關的生理途徑，而 AML 病患變異選擇性剪接的基因也可能導致許多重要的細胞生理途徑改變，例如 G 蛋白介導的信號傳遞途徑(G protein-mediated signaling pathway)、蛋白質降解、細胞週期和 DNA 損傷/修復，甚至會影響血癌細胞對化療的反應。但本研究主要是從臨床的角度出發，主要想探討正常基因選擇性剪接的變異會導致臨床的效應，因此變異選擇性剪接(aberrant alternative splicing)在 MDS 或 AML cells 所導致的確切生物學效應(biological effects)仍有待進一步更多的研究。

4.2 *NFYA* 基因選擇性剪接對急性骨髓性白血病(AML)的臨床與生物學效應

接著，本研究論文使用先前提到的 RNA-seq 平台，進一步發現 Nuclear Transcription Factor Y Subunit Alpha (*NFYA*)基因獨特選擇性剪接型式(unique alternative splicing pattern)，在 AML 細胞和正常臍帶血造血細胞之間顯示出顯著差異(圖 26B)。這引起對這個基因的研究產生興趣，這可能在 AML 致病機轉中發揮作用。

Nuclear Transcription Factor Y (NFY)是一種普遍存在的異源三聚體轉錄因子(heterotrimeric transcription factor)，它會結合到其目標基因的啟動子序列 CCAAT box，讓目標基因表達(圖 36A)⁵⁸。越來越多細胞和動物模型的證據支持 NFY 在細胞週期調控以及細胞凋亡和增殖的平衡中發揮重要作用(圖 36B)⁵⁹。在許多人類疾病中觀察到 NFY 的調控發生變異，例如硬皮症、神經退化性疾病和惡性腫瘤⁵⁸⁻⁶⁰。根據之前的文獻報告，NFY 在果蠅複眼發育過程中佔了重要

的關鍵作用⁶¹。許多最近的研究發現，NFY 似乎會促進很多癌症的發生。另外，NFY 在細胞株和癌症樣本中很少發生突變，因此它在癌症發生中扮演的角色可能是透過異常的基因表達。事實上，*NFYA* 基因在許多癌症中過度表達，如肺癌、乳腺癌、子宮頸癌和肝癌⁶²⁻⁶⁹。NFY 由 3 個亞基(subunit)組成，即 NFY 亞基 α 、 β 和 γ (*NFYA*、*NFYB* 和 *NFYC*)；由於 *NFYA* 與 CCAAT box 的序列特異性相互作用，所以 *NFYA* 被認為是 NFY 中最重要的調控單元^{58,59}。在小鼠的動物實驗中，若把 *NFYA* 基因剔除，對於正常組織的發育、存活和代謝功能將受到嚴重影響⁷⁰。*NFYA* 還能刺激活化多個造血幹細胞相關的調節基因，並促進造血幹細胞的更新(renewal)。此外，*NFYA* 基因對造血幹細胞的增殖和存活至關重要^{71,72}。由於我們對 *NFYA* 基因對造血幹細胞的已知影響，以及本研究論文對 AML 樣本 *NFYA* 基因選擇性剪接型的分析，我們很好奇 *NFYA* 基因選擇性剪接是否對 AML 患者有任何的臨床和生物學效應。

NFYA 基因有兩個主要的剪接異構體(alternatively spliced isoforms)。*NFYA* 長型(*NFYA-L*)包含完整的 10 個 exons，而 *NFYA* 短型(*NFYA-S*)在 pre-mRNA 剪接時 exon 3 skipping，導致 N 端的轉錄啟動區域(transactivation domain)缺少了 29 個胺基酸(圖 36C)⁷³。這兩種異構體之間因選擇性剪接而造成的變異，導致它們在蛋白質結構和功能上存在差異。最初，*NFYA-L* 和 *NFYA-S* 這兩種剪接異構體被認為具有相等的生物功能。但最近的研究顯示，*NFYA-S* 在小鼠胚胎幹細胞中被大量表達，但隨著分化的過程中表達會逐漸減少⁷⁰。*NFYA-S* 是幹細胞轉錄網絡(transcriptional circuitry)的一部分，在胚胎幹細胞中高度表達，而 *NFYA-L* 則會促進細胞分化。這些發現顯示 *NFYA* 基因的不同剪接異構體在調控幹細胞增殖和細胞分化方面可能具有不同的生物學功能。這對於理解幹細胞的維持和分化以及相關疾病的致病機轉具有重要意義。其他研究指出，這兩個剪接異構體在肌肉細胞分化上有不同的影響。*NFYA-S* 可促進肌肉細胞增殖，而 *NFYA-L* 則抑制肌肉細胞增殖。*NFYA-L* 似乎有助於促進肌肉細胞分化^{74,75}。此外，在

先前的研究中，雖然有越來越多的證據顯示 *NFYA* 分子對癌細胞會產生影響，但對於其因選擇性剪接所產生不同異構體在致癌過程中所造成的影響，知之甚少，尤其是白血病。由於本研究的 RNA-seq 資料顯示 *NFYA* 基因選擇性剪接形式在 AML 血癌細胞與造血幹細胞(cord blood sample)有很顯著的差異(圖 26B)，因此我們對 *NFYA* 的變異選擇性剪接在 AML 的致病過程中所扮演的角色感到興趣。

本研究觀察到 AML 細胞與正常臍帶血造血細胞相比，*NFYA* 的表達量顯著增加(圖 26A)。儘管 *NFYA* 的表達增加與很多種腫瘤的不良預後有關，但本研究對於接受標準高強度化療的 AML 患者進行的分析並未得出相同的結果。然而，本研究對 *NFYA* 基因選擇性剪接的分析顯示，在接受標準高強度化療的 AML 患者中，較高的 *NFYA*-S 表達或 *NFYA*-S predominance pattern 與整體存活的提升密切相關。進一步分析，本研究發現 *NFYA*-S 表現量較高或 *NFYA*-S predominance pattern 的 AML 患者在接受標準前導性化療後，有較高的疾病完全緩解率(complete remission rate)。上述的研究結果可進一步推斷，在 AML 細胞中過度表達 *NFYA*-S transcript isoform 將增強對化療的敏感性。

基因集富集分析(GSEA)顯示 AML 病患若 *NFYA*-L 表達量低但 *NFYA*-S 表達量高(*NFYA*-S predominance pattern)將會導致與造血幹細胞(HSC)特徵和細胞週期相關的基因被高度表達。細胞週期調節基因的高度表達將可以使腫瘤細胞對化療藥物的敏感性提高。然而，值得注意的是，如果與造血幹細胞(hematopoietic stem cells, HSCs)相關的基因被高度表達是否會呈現類似白血病幹細胞(leukemia stem cells, LSCs)的特徵，是否也會表現出化療抗性。先前的研究顯示靜止的 HSCs 和 LSCs 在代謝方面的區別。靜止的 HSCs 表現出低代謝活性，依賴糖酵解(glycolysis)進行能量生產。但相反的是，靜止的 LSCs 則減少對糖酵解的依賴，更青睞粒線體氧化磷酸化(mitochondrial oxidative phosphorylation, OXPHOS)滿足其能量需求。提高的 OXPHOS 活性與 AML 化療藥物 Ara-C 抗性相關

^{53,76}。由於代謝狀況的不同，因此若有 HSC 的特徵可能不一定會與類似 LSC 的特徵相同的化療抗性。

相反地，AML 病患若 *NFYA-L* 表達量高但 *NFYA-S* 表達量低(*NFYA-L* predominance pattern)將導致與 mixed lineage leukemia (MLL)和 mitogen-activated protein kinase (MAPK)信號通路活化一致的基因被高度表達，同時也會高度表達與單核細胞遷移相關的基因。這種模式可能會導致相對的化療抗性。在 *MLL*-rearranged 白血病中，白血病幹細胞自我更新(self-renewing)的能力持續存在、獲得次要突變、細胞信號、細胞週期和 DNA 損傷修復機制的失調，都可能是化療抗性表現型的因素之一⁷⁷。此外，有文獻指出 *KMT2A (MLL)*重組白血病細胞的化療抗藥性與特定蛋白質酪氨酸激酶(PTK)基因的表達升高有高度關聯性，包括 *FLT3*、*BTK*、*SYK*、*JAK* 家族激酶以及各種 Src 家族 PTK⁷⁸。因此，本研究論文又進一步評估了兩群 AML 患者(*NFYA-L* predominance pattern 及 *NFYA-S* predominance pattern)的 *FLT3*、*BTK*、*SYK*、*JAK2* 和幾種 Src 家族 PTK 的基因表達量差異。在 *NFYA-L* predominance 的患者中(相對於 *NFYA-S* predominance 的患者)，*HCK* 表達量上升最顯著(fold change = 1.96, $P < 0.001$)，其次是 *FGR* (fold change = 1.95, $P < 0.001$)，*FLT3* (fold change = 1.39, $P = 0.0015$)，*SYK* (fold change = 1.36, $P < 0.001$)和 *JAK2* (fold change = 1.34, $P < 0.001$) (圖 37A-E)。另外，MAPK 訊號傳遞路徑被活化已被視為與各種實體腫瘤，包括 AML，的化療抗性相關聯⁷⁹。

為了驗證上述的假說，本研究論文設計了細胞實驗，透過在 OCI-AML3 血癌細胞株內過度表達不同的 *NFYA* transcript isoforms 來觀察是否有不同的細胞生物效應。本研究的實驗結果顯示，過度表達 *NFYA-S* transcript 顯著增加了細胞增殖及存活能力，並增強了對 Ara-C 化療的敏感性。此外，為了消除實驗中細胞內源性 *NFYA* transcript isoforms 的影響，我們先使用 shRNA 對 *NFYA* 基因敲除(knockdown)，這將會抑制細胞原有內源性 *NFYA-S* 及 *NFYA-L* 的表達。我

們觀察到 *NFYA* 基因的敲除會顯著降低了細胞的增殖及存活能力，以及細胞週期 S 期細胞比例(圖 32, 34)。當我們在 *NFYA* 基因敲除後再表達 *NFYA-S* transcript，我們注意到細胞的存活能力和增殖似乎有所恢復，並且細胞週期 S 期的細胞比例也明顯增加。上述我們在白血病細胞中的發現與其他類別的固態腫瘤(solid cancers)的觀察一致，暗示 *NFYA* 基因選擇性剪接變異可能對細胞週期相關基因產生不同的影響，從而在細胞增殖中發揮重要的作用。最後，基因集富集分析(GSEA)顯示，主要表達 *NFYA-L* 的 OCI-AML3 細胞會抑制細胞週期中進展至 G2/M 調控的相關基因表達，但會活化細胞代謝路徑 oxidative phosphorylation 的相關基因表達，這暗示著 *NFYA-L* 會抑制細胞週期進程，造成細胞的增殖下降，但會讓細胞對 Ara-C 化療藥物產生抗性 (圖 35)。

Gerd Bungartz 等人先前報告指出，*NFYA* 基因剔除(knockout)導致造血幹細胞(hematopoietic stem cells)在 G2/M 細胞週期中積聚，觸發凋亡，最終導致造血功能異常和動物死亡。這一發現突顯了 *NFYA* 在造血幹細胞增殖和存活中的重要作用^{71,72}。在我們的研究中，我們觀察到 *NFYA* 可能也在白血病細胞的增殖中發揮著關鍵作用，並且我們確定了這種效應歸因於 *NFYA* transcript isoform *NFYA-S*。

然而，本研究仍存在一些限制需要注意。首先，由於資源和技術的限制，本研究論文僅使用一個白血病細胞株進行實驗。考慮到 AML 及其不同亞型的複雜性，研究 *NFYA* transcript isoforms 對不同 AML 亞型的影響仍需要再進一步的研究。其次，本研究論文主要依賴初步的體外細胞實驗來驗證 *NFYA* splice variants 在 AML 中的功能角色。未來的研究應包括體內動物模型研究，以更深入了解所涉及的分子機制。

第五章、展望(Perspectives)

整體而言，本研究論文確認了基因選擇性剪接對骨髓性血癌的臨床與生物學效應。轉錄體中整體基因選擇性剪接變異的程度被證實對於骨髓性血癌疾病的進展與治療預後產生影響。透過進一步分析某些特定基因選擇性剪接的變異，我們確認了在骨髓性血癌的發生中，除了透過基因突變改變原有細胞正常的生理功能外，基因選擇性剪接的變異也是一個重要的致病機轉。因此，這項研究的結果不僅擴展了對骨髓性血癌發病機制的理解，同時凸顯了基因選擇性剪接在疾病的複雜性中所扮演的關鍵角色。對於治療的角度而言，本研究為開發影響基因選擇性剪接調控的藥物提供了有力的基礎。未來的治療策略或許可以針對特定基因選擇性剪接的變異進行精準調控，為患者帶來更有效的治療方案。這個新的治療契機有望為骨髓性血癌患者帶來更多希望與可能性。

如前述，本研究論文使用的研究檢體為骨髓單核細胞(bone marrow mononuclear cells)，而不是同質性更高，如經過細胞分選(cell sorting)的 $CD34^+$ 細胞。雖然已有研究顯示骨髓單核細胞的基因表達方式與經過細胞分選純化的 $CD34^+$ 前驅細胞相似⁵⁶，但將來如果能進一步使用經過細胞分選(cell sorting)的 $CD34^+$ 細胞，應更能提升研究的品質。在另一方面，本研究使用的是批量樣本(bulk sample)，可以顯示一組或一批細胞的整體特性，例如血癌細胞群體中的整體模式。然而，近期單細胞的研究已逐漸成為研究趨勢，例如單細胞 RNA 序列(scRNA-seq)的技術不僅擴展了研究範疇，還超越了對細胞狀態的描述性研究，我們可以想像將這項技術應用於眾多的醫學領域。腫瘤異質性是一種普遍現象，不僅可能發生在腫瘤內部，還可能存在於不同腫瘤之間。預期 scRNA-seq 可應用於揭示傳統批量樣本(bulk sample)轉錄體研究難以辨識的未知腫瘤特徵。例如，此技術可應用於評估癌細胞在抗藥性發展過程中的轉錄異質性，並分析特定途徑的表達情況。另一個引人注目的潛在應用是識別參與幹細胞調控網絡的基因。也許我們能夠深入了解幹細胞如何被啟動轉變為癌細胞，從而建

立癌症演變模型，這對於理解人類白血病疾病的基本致病過程應該至關重要。

此外，本研究論文聚焦於 MDS 或 AML 患者整個轉錄體 mRNA 整體選擇性剪接的差異對臨床預後的影響，發現整個轉錄體整體 mRNA 變異選擇性剪接模式的程度可以預測 MDS 或 AML 患者的預後，而且是獨立的預後因子。透過本研究的轉錄體數據分析，確實觀察到 MDS 或 AML 細胞因選擇性剪接產生 transcript isoforms，而使得轉錄體呈現複雜性，然而本研究論文並未深入研究蛋白質體(proteome)的多樣性或相關的生物學效應，因此，仍需進一步研究探討在 MDS 或 AML 細胞中變異選擇性剪接的生物學意義。

接著本研究論文從 transcriptome dataset 找到了 *NFYA* gene 的變異選擇性剪接會影響血癌病患接受標準高強度化療後的預後，並藉由血癌細胞的體外模型 (*in vitro* models) 來驗證不同的臨床結果可能來自於基因變異選擇性剪接可能會改變血癌細胞原有的生理路徑，因此造成不同的預後結果。儘管本研究結果呈現了有意義的發現，但必須承認仍存在一些限制，進而影響本研究論文結論的廣泛性及深度的應用。首先，由於資源及技術的限制，本研究論文僅使用了一個白血病細胞株進行實驗，但急性骨髓性白血病(AML)包含多樣的亞型，這使得將來仍有必要在多個細胞株上進行更全面的研究。其次，本研究論文主要仰賴初步的體外細胞實驗來驗證基因選擇性剪接變異在 AML 中的功能角色。儘管這些實驗提供了對潛在分子機制的初步證據，但我們必須承認體外條件可能無法完全模擬複雜的體內微環境。為了提升我們研究結果的臨床相關性，未來的研究應該包含體內動物模型(*in vivo* animal models)實驗，以更細緻地了解基因選擇性剪接變異在血癌細胞上更廣泛生理環境之間的相互作用。這樣的探索將有助於更全面地理解基因選擇性剪接變異在 AML 致病過程中的角色，同時提供更深入的基礎對於未來治療策略的制定。

另外，本研究的檢體收集於疾病初次診斷時。透過對疾病預後的仔細觀察，我們得以深入研究基因選擇性剪接的變異對疾病發展或治療效果的潛在影

響。然而，為了更全面地了解基因選擇性剪接的角色，未來的研究也可考慮在疾病治療後收集檢體。透過比較疾病初次診斷和治療後兩個不同時間點的基因選擇性剪接變化，我們能夠深入瞭解治療對基因選擇性剪接所帶來的影響。這樣的設計不僅有助於揭示基因調控的時空動態，也提供了評估治療效果及相應生物學改變的新視角。其次，目前有些新藥的開發是在調節 alternative splicing 的過程，如 splicing inhibitor or splicing modulator (例如 SF3B1 modulator H3B-8800 已被用於骨髓性血癌的臨床試驗)⁸⁰，我們之後也可以把這些新開發的藥物用於細胞實驗上，評估用藥前後對於血癌細胞 global alternative splicing pattern 是否會造成改變。

最後還有一些研究主體有待後續更進一步的研究：1. *NFYA* 基因選擇性剪接對骨髓造血不良症候群(MDS)病患上的角色。2. 本研究論文的主題聚焦於骨髓性血液惡性疾病(myeloid malignancies)，如同緒論所述，骨髓性血液惡性疾病主要由骨髓增生性疾病(myeloproliferative neoplasm, MPN)，骨髓造血不良症候群(myelodysplastic syndrome, MDS)及急性骨髓性白血病(acute myeloid leukemia, AML)，但本研究論文只探討了基因選擇性剪接(alternative splicing) 在 MDS 及 AML 的臨床與生物學效應，期待將來我們能更進一步研究探討 alternative splicing 在 MPN 上的角色。

圖 1：正常生理狀況下，基因選擇性剪接示意圖

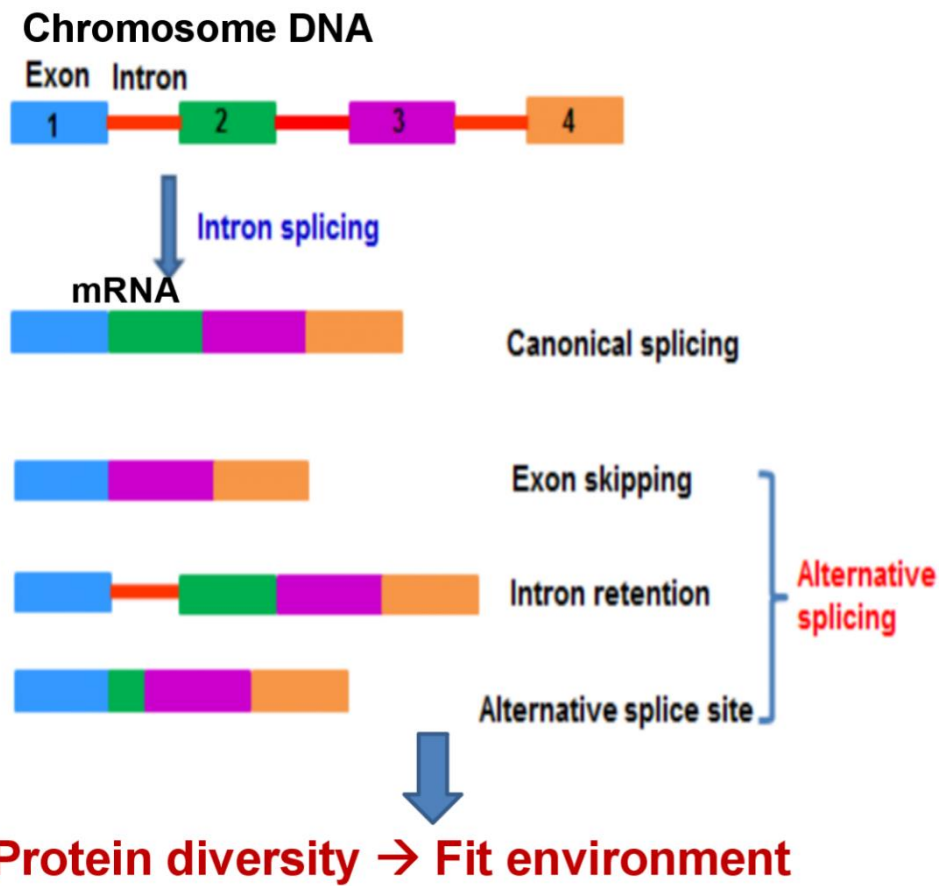


圖 2：基因變異選擇性剪接示意圖

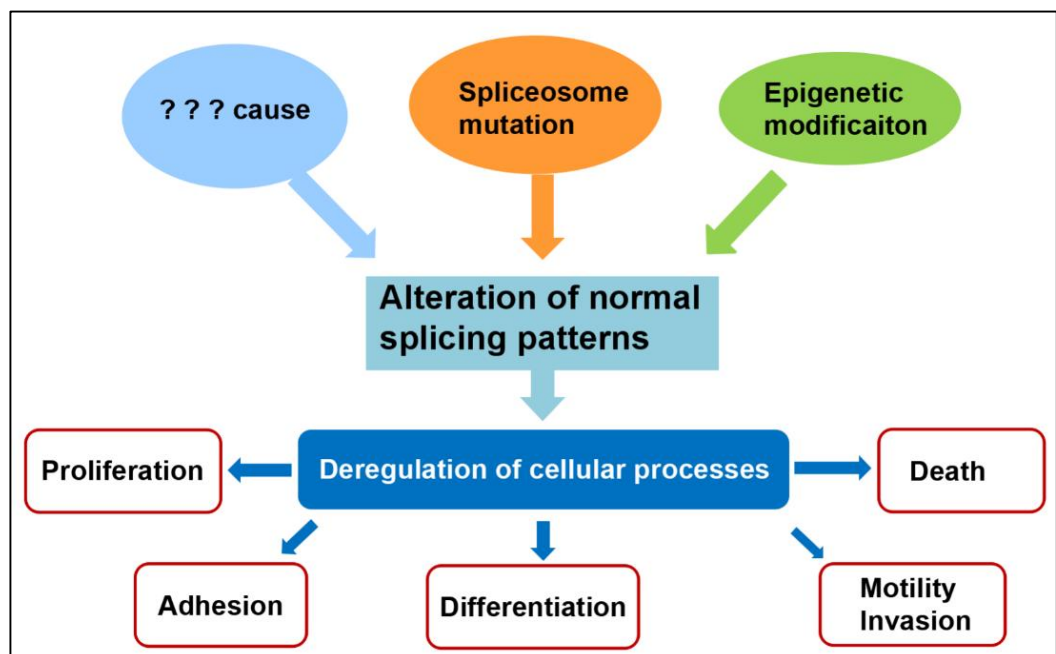
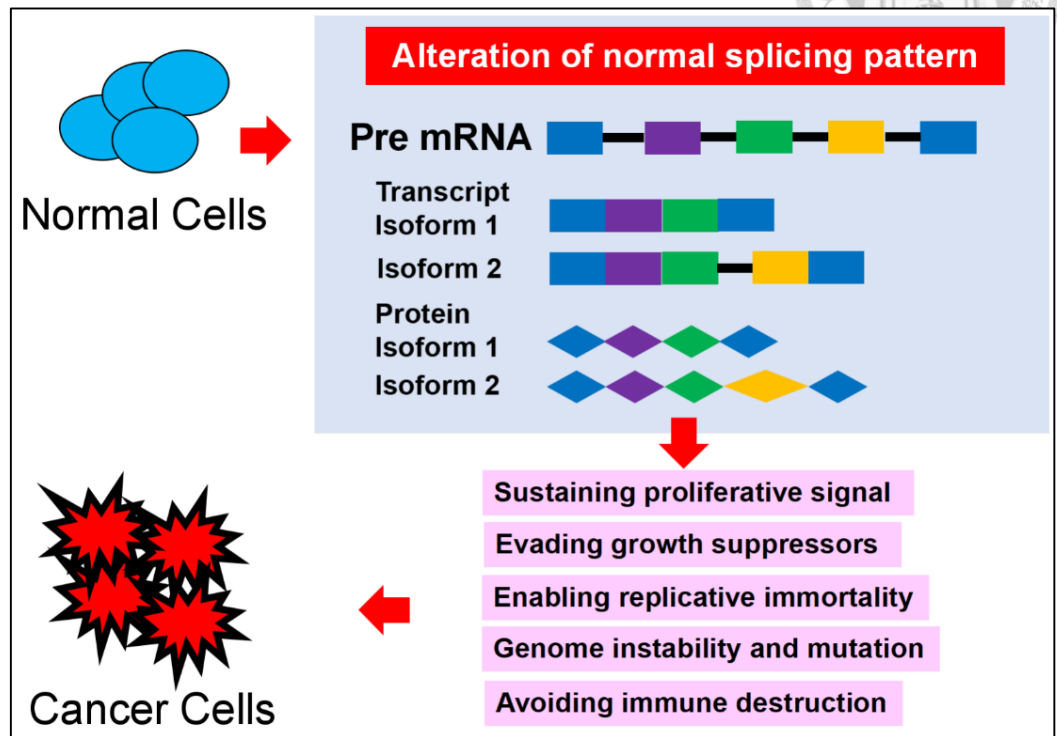
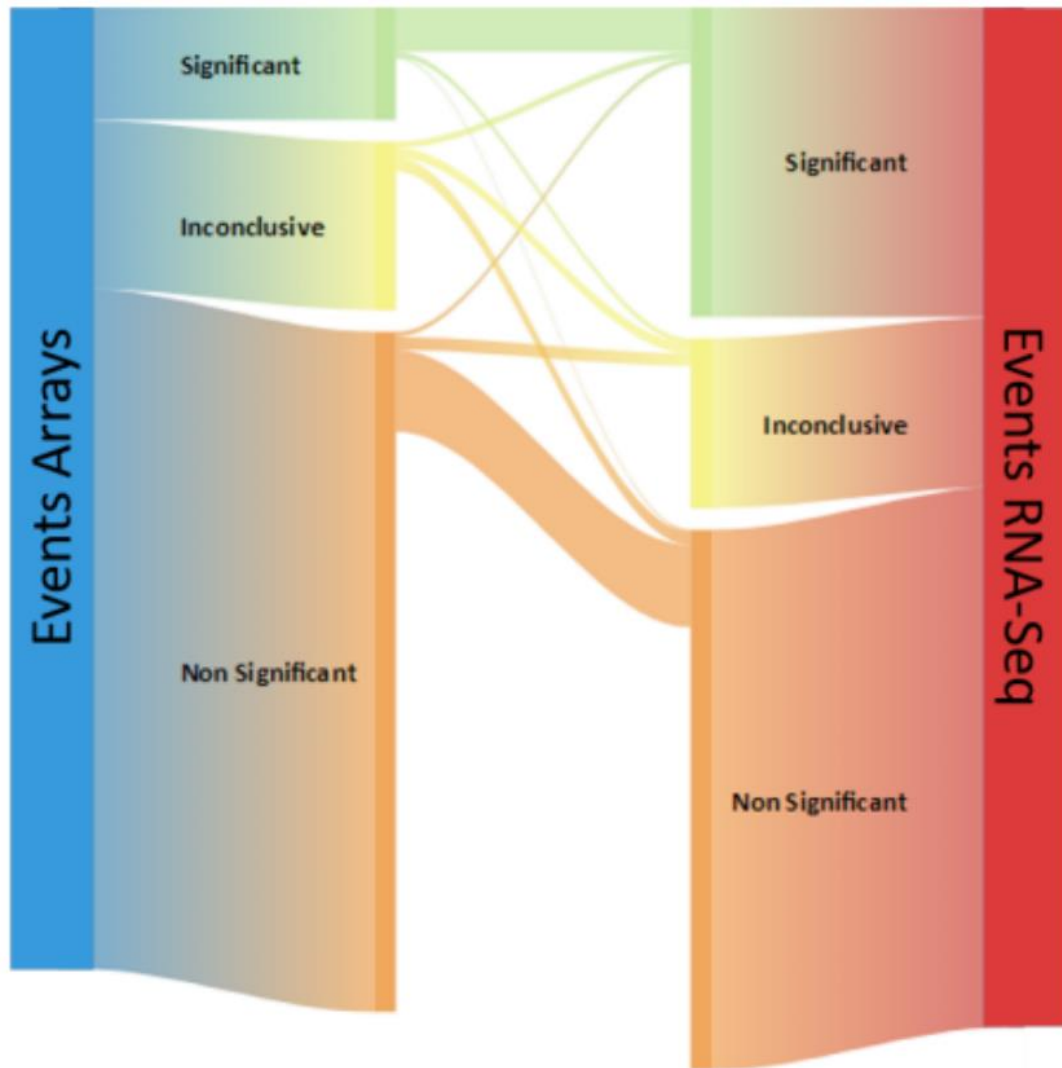


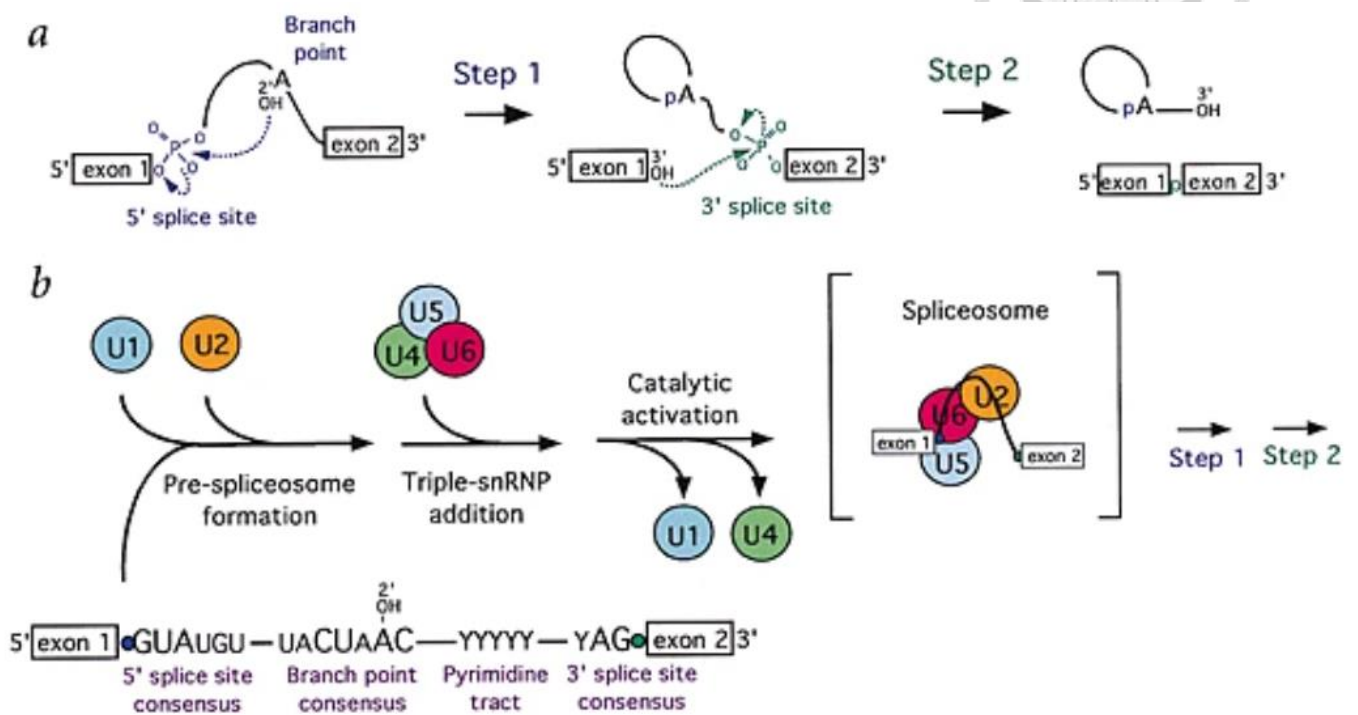
圖 3：比較使用 microarray platform 及 RNA-seq platform 在偵測基因選擇性剪接(alternative splicing event)的差異



Source:

Romero JP, et al. BMC Genomics. 2018;19(1):703.

圖 4：剪接體的組裝(spliceosome assembly)

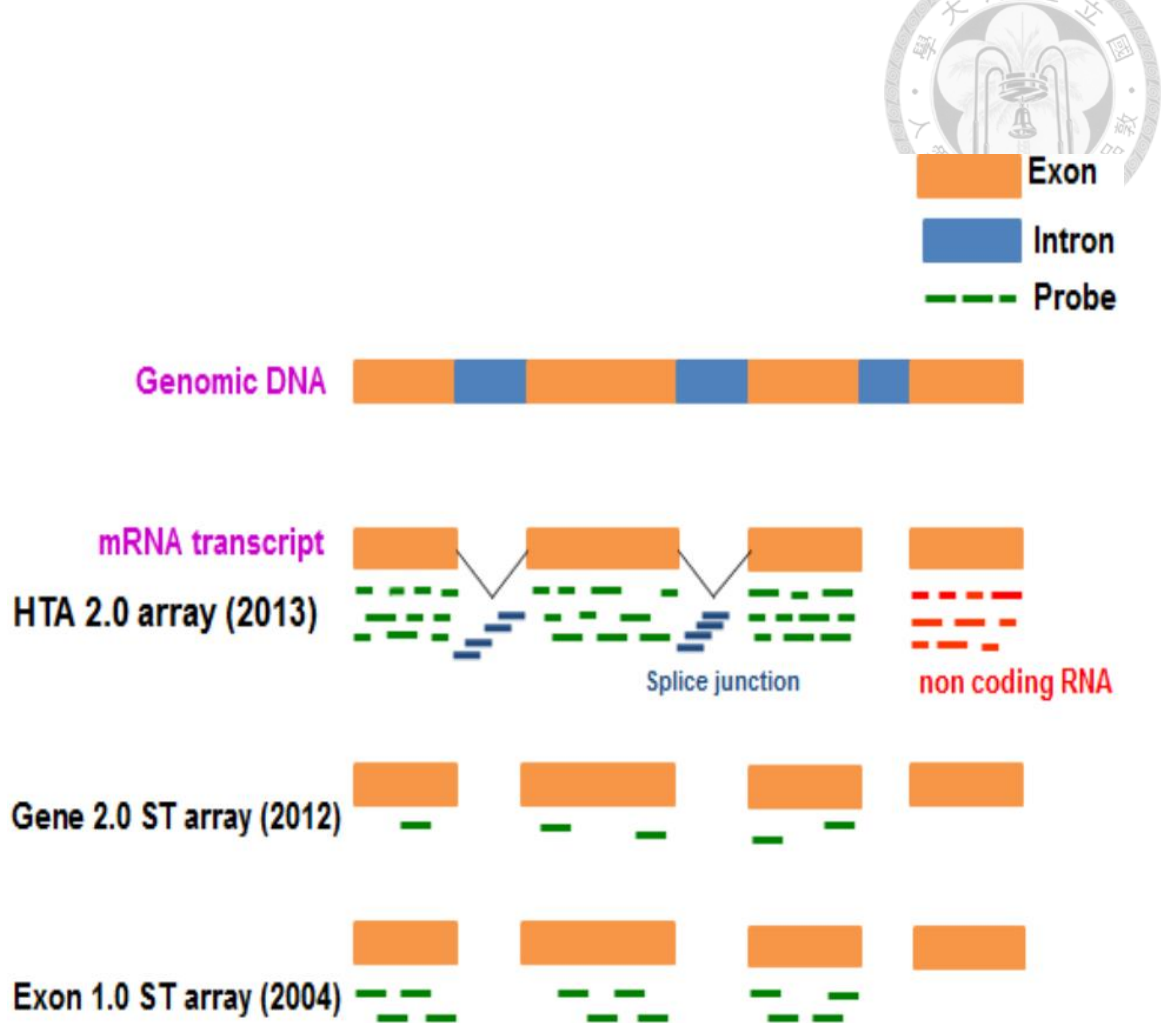


U1~U6: small nuclear RNA(snRNA)

Source:

Collins CA, et al. Nat Struct Biol. 2000;7(10):850-854.

圖 5：Affymetrix Human Transcriptome Array 2.0 (HTA 2.0)示意图



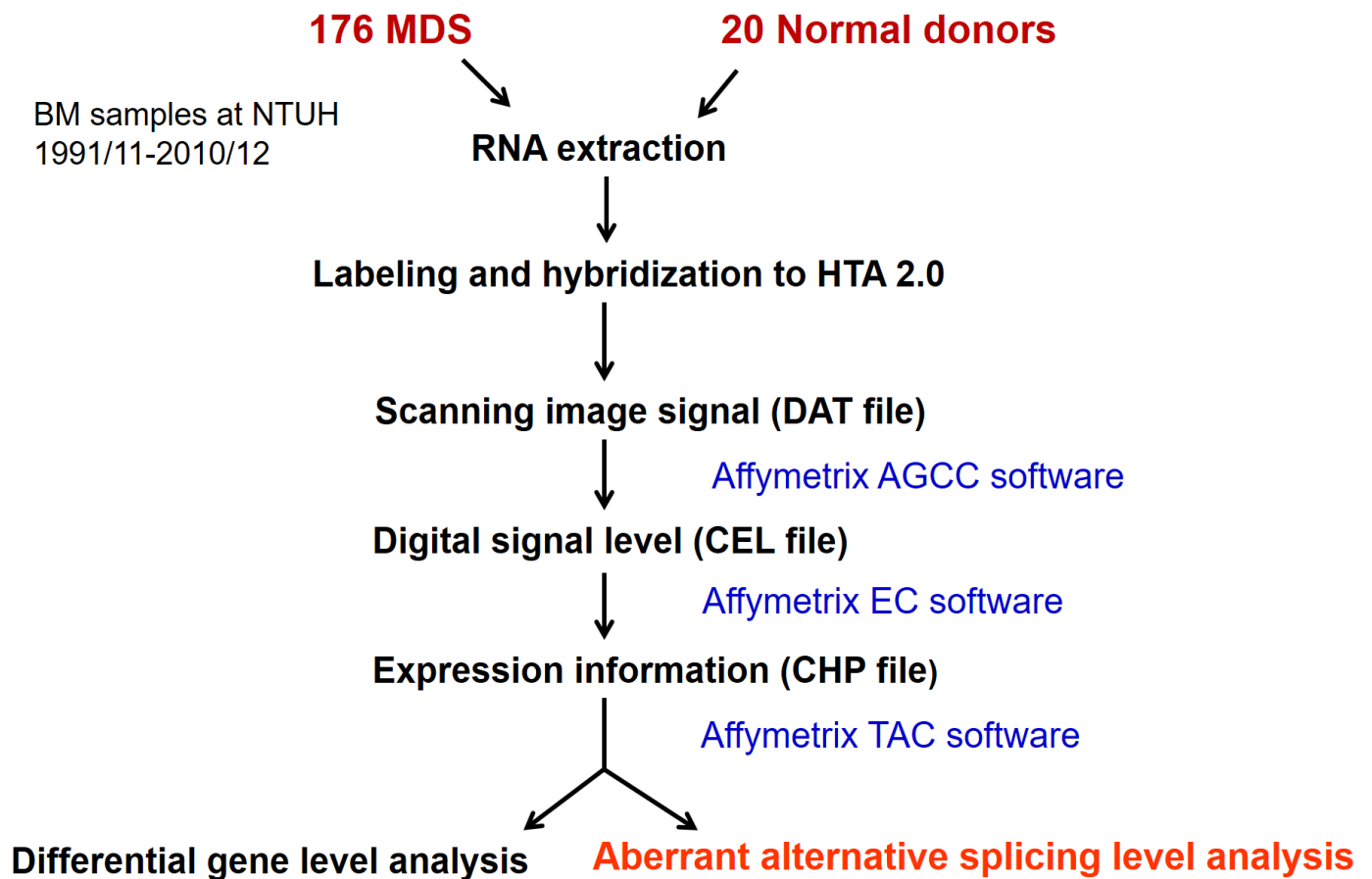
Source:

Datasheet of Affymetrix GeneChip® Human Transcriptome Array 2.0

圖 6：骨髓造血不良症候群(MDS)轉錄體中基因變異選擇性剪接事件
(aberrant alternative splicing events)分析的流程圖



Analysis workflow



Source:

Yang YT, et al. Blood Cancer J. 2018;8(8):78.

圖 7：在 microarray 平台上定義變異選擇性剪接(aberrant alternative splicing event)的示意圖

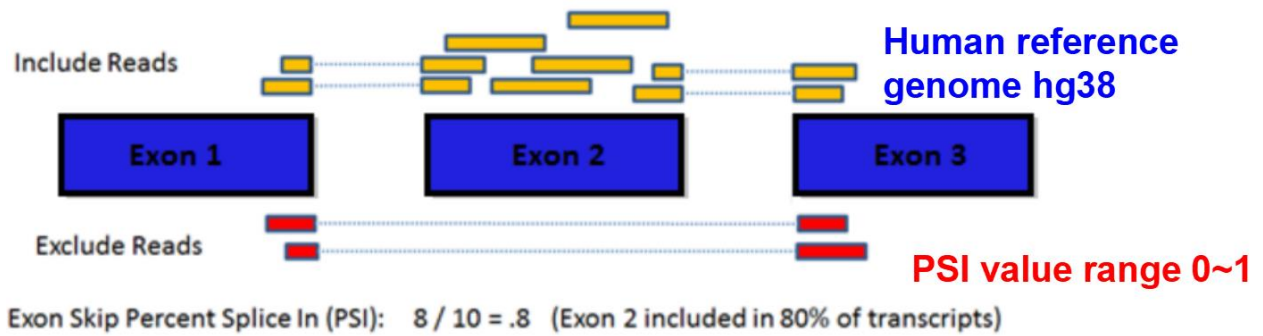


Aberrant alternative splicing event

	MDS			Normal		
Gene A	1	2	3	1	2	3
Splice index =	{ Signal exon 2 in MDS }			4	40	
	{ Signal Gene A in MDS }			2	2	
	{ Signal exon 2 in Normal }			2	2	
	{ Signal Gene A in Normal }			1	1	
				Splicing Happen?		
				No	Yes	

- aberrant AS event : $SI > 2$ or < -2
(+ : MDS > Normal ; - : MDS < Normal)
- aberrant AS gene : at least one aberrant AS event

圖 8：在 RNA-seq 平台上，定義選擇性剪接(alternative splicing event)的示意圖



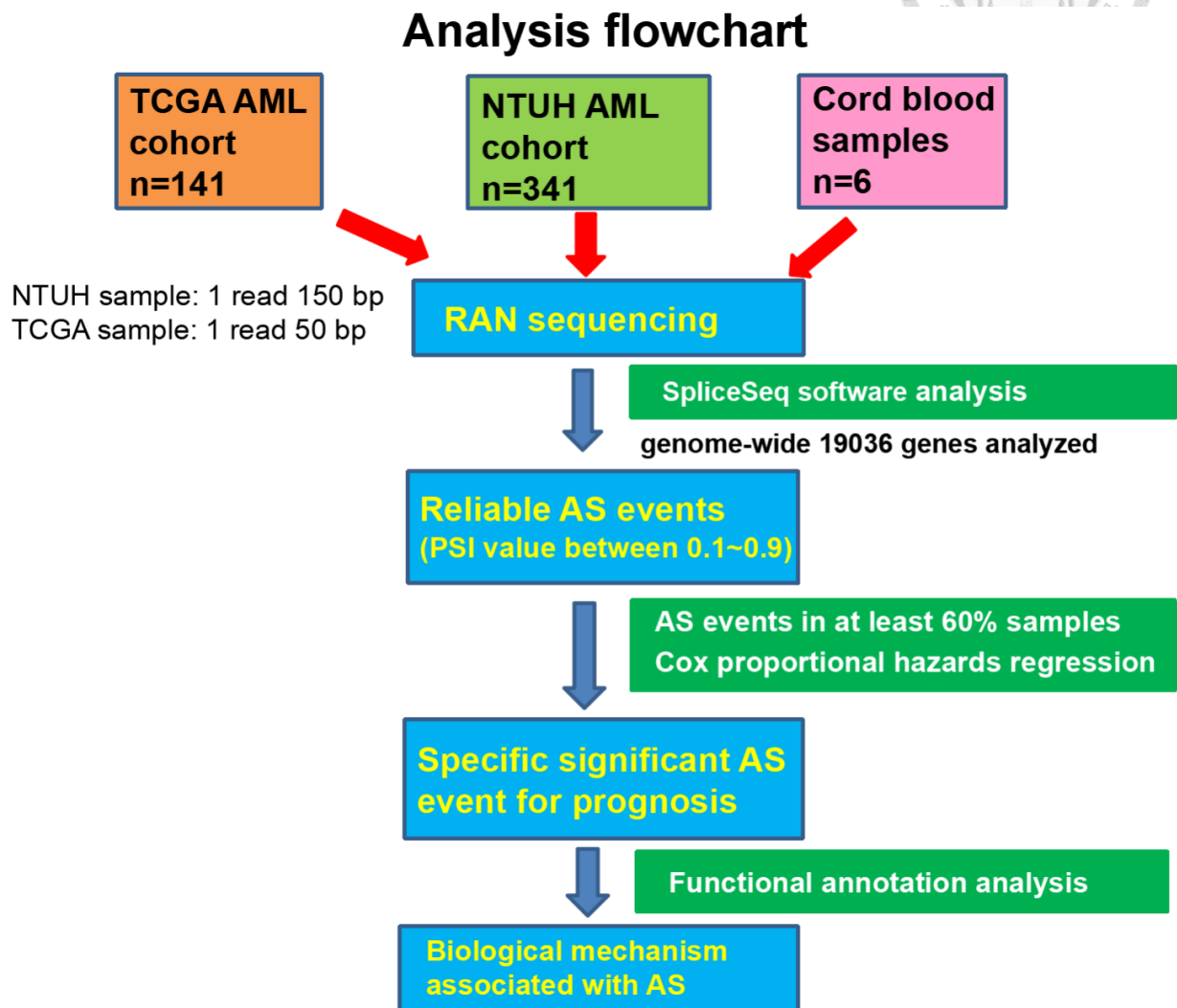
transcript isoform 1: E1 + E2 + E3

transcript isoform 2: E1 +E3

If **PSI of E2 higher** → **higher isoform 1**, but lower isoform 2

If **PSI of E2 lower** → lower isoform 1, but **higher isoform 2**

圖 9：急性骨髓性白血病(AML)轉錄體中基因選擇性剪接事件(alternative splicing events)分析的流程圖



Source:

Yang YT, et al. Am J Hematol. 2023;98(5):784-793.

圖 10：lenticral vector 示意图

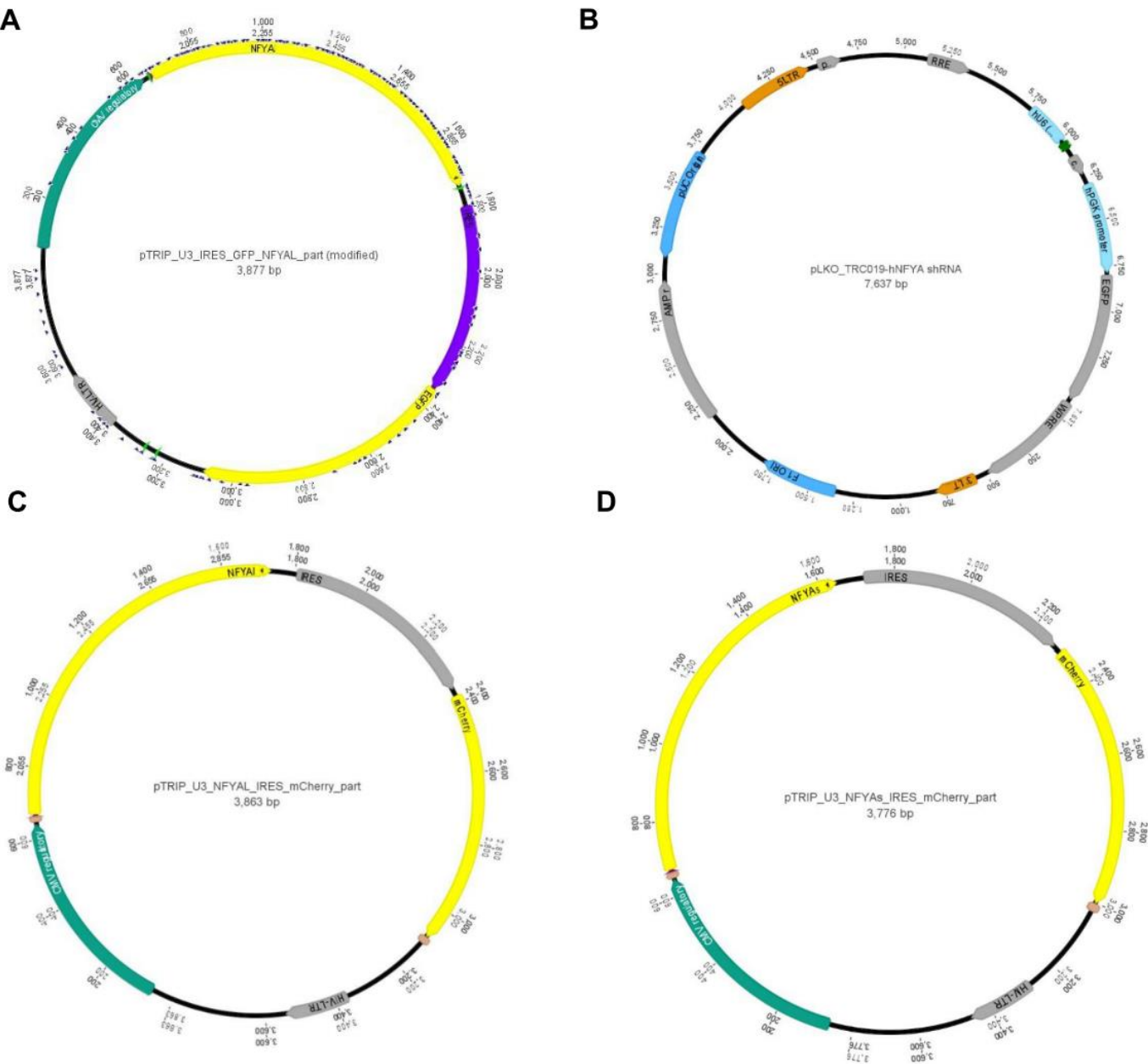
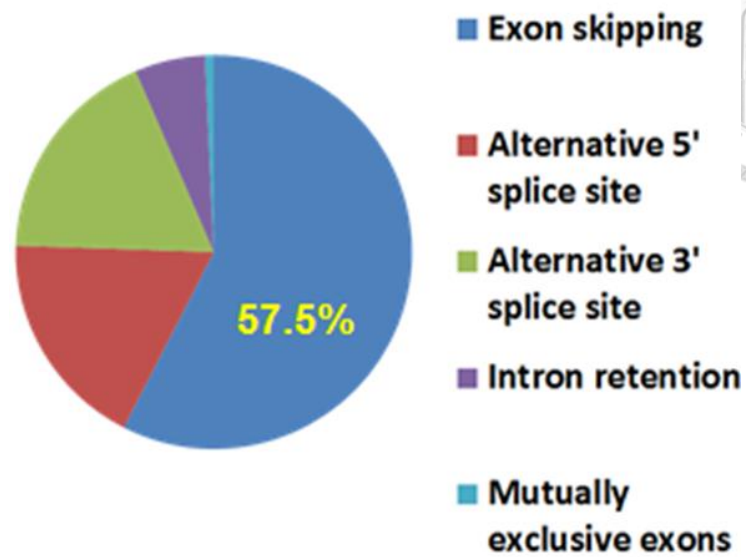


圖 10：lenticral vector 示意图：(A) pTRIP_dU3_IRES_GFP lenticral vector
(B) pLKO_TRC019_shRNA lenticral vector
(C,D) pTRIP_dU3_IRES_mCherry lenticral vector。

圖 11：分析 MDS 病患整個轉錄體變異選擇性剪接的類型(event type)



Source:

Yang YT et al. Blood Cancer J. 2018;8(8):78.

圖 12：Aberrant alternative splicing score

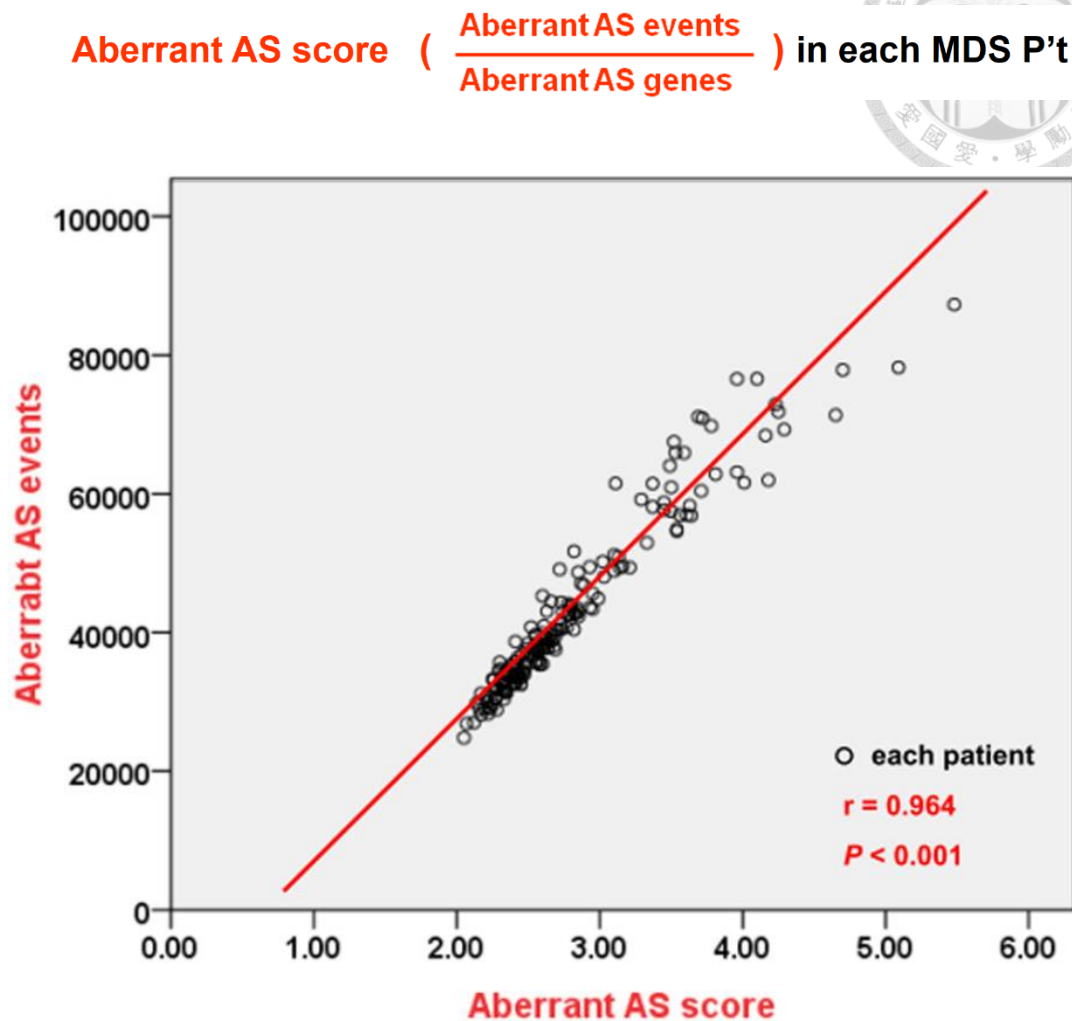


圖 12：Aberrant alternative splicing score 為轉錄體中所有 aberrant alternative splicing events 與所 involve 的基因(aberrant alternative splicing genes)相除。Aberrant alternative splicing score 與轉錄體中所有 aberrant alternative splicing events 呈現高度正相關，因此此數值高低可代表整個轉錄體整體 mRNA 變異選擇性剪接模式的程度(the degree of global aberrant alternative splicing pattern)。

Source:

Yang YT, et al. Blood Cancer J. 2018;8(8):78.

圖 13：MDS 病患與 normal donor 骨髓細胞轉錄體整體變異選擇性剪接模式(pattern)

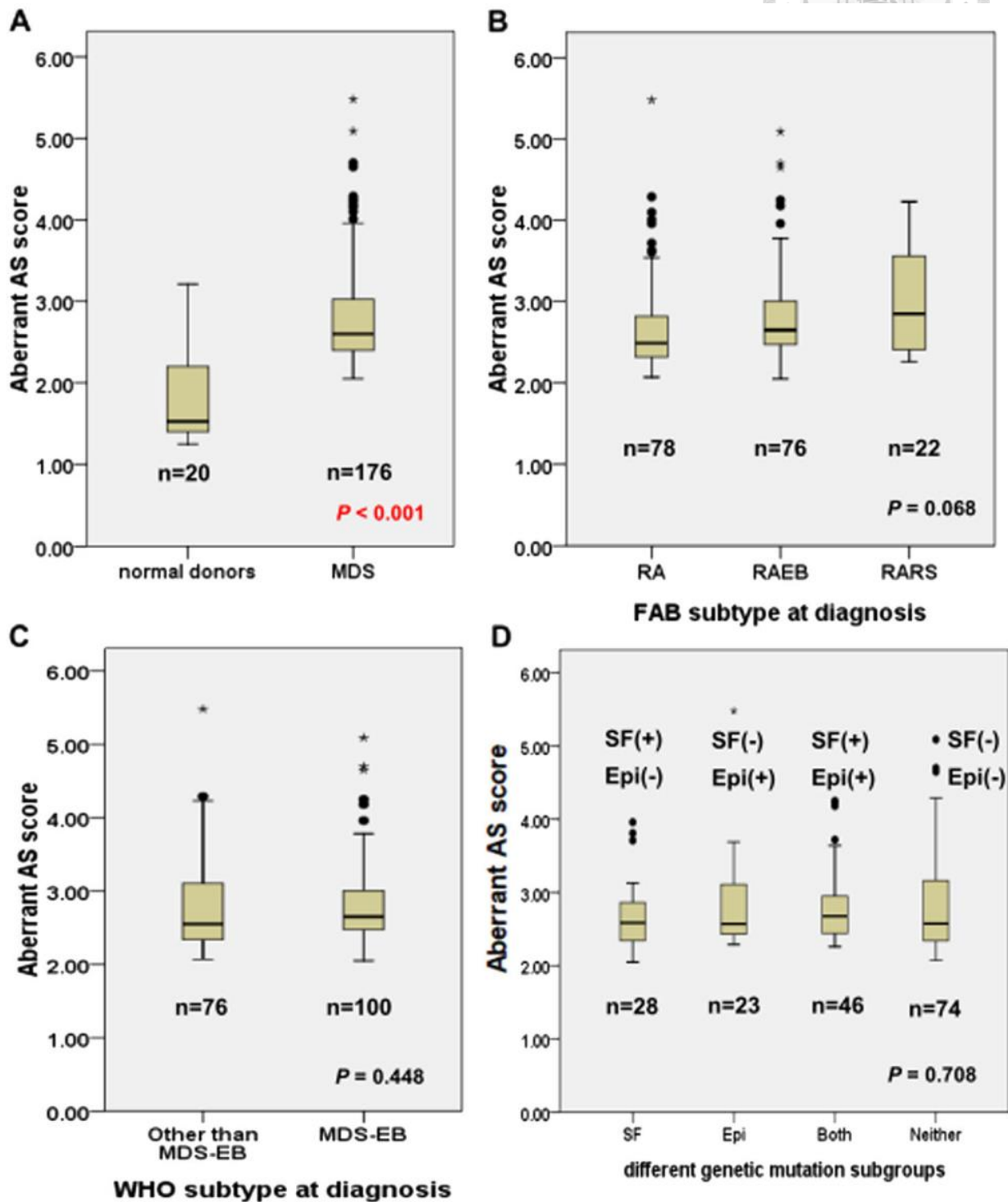


圖 13：分析轉錄體整體變異選擇性剪接模式(pattern)在 MDS 病患與正常捐髓者骨髓細胞間的差異(A)，及各個 MDS subtype 間的差異(B-D)。Abbreviations:

SF indicates splicing factor mutations, Epi indicates epigenetic mutations.

Source: Yang YT, et al. Blood Cancer J. 2018;8(8):78.

圖 14：比較 MDS 病患在變異選擇性剪接模式高變異度及低變異度臨床預後的差異

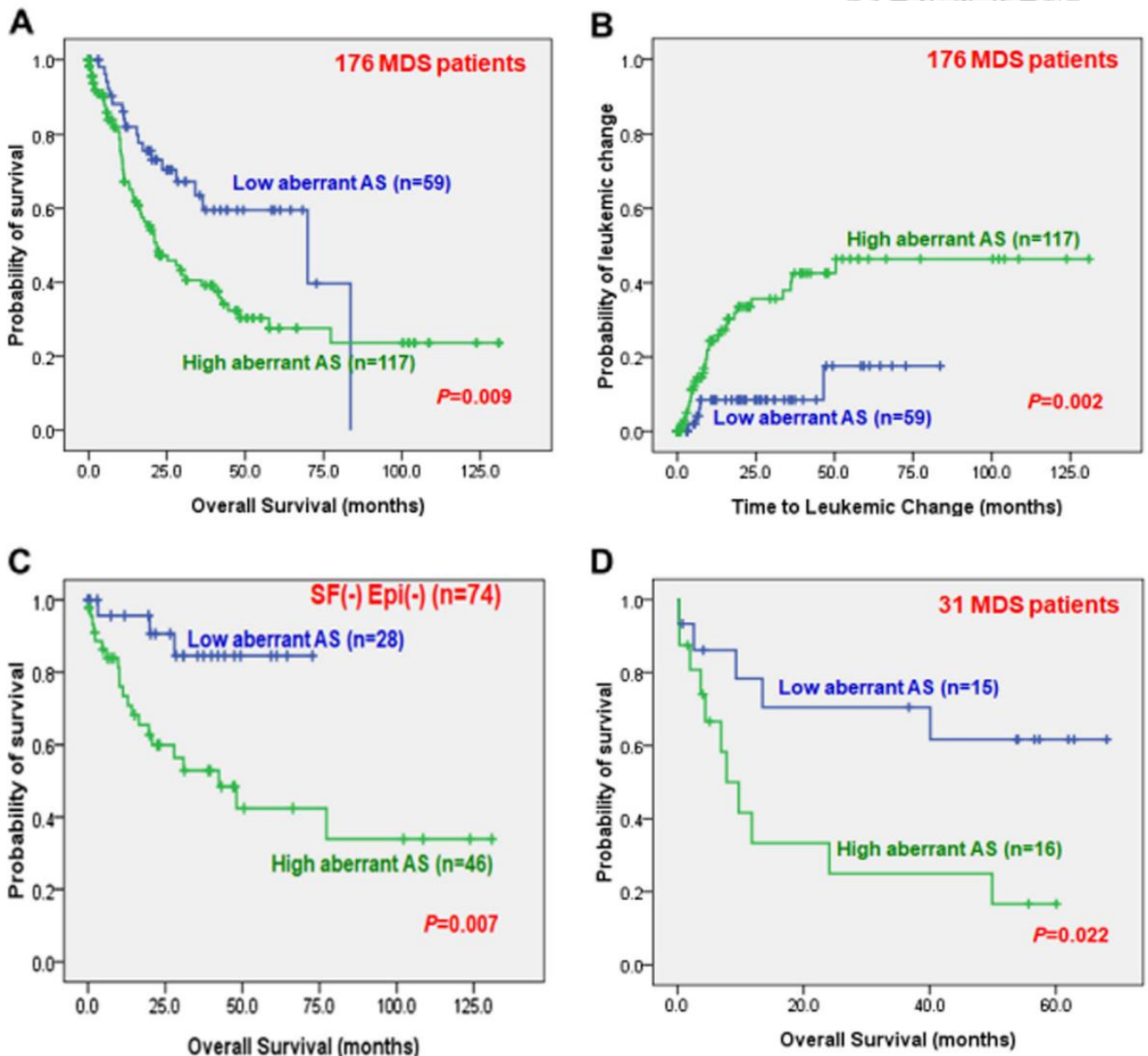


圖 14：比較 MDS 病患整個 study cohort (A,B)，無任何 spliceosome gene mutations 及 epigenetic modification gene mutations (C)，validation cohort (D)，在變異選擇性剪接模式高變異度及低變異度臨床預後的差異。

Source:

Yang YT, et al. Blood Cancer J. 2018;8(8):78.

圖 15：利用 PCR / Sanger sequencing 分析 MDS 病患 *AUP1* 基因(A-D)，*PVR2*，*GRIK5* 基因(E,F)的變異選擇性剪接，以驗證在 microarray 平台上所預測的變異選擇性剪接

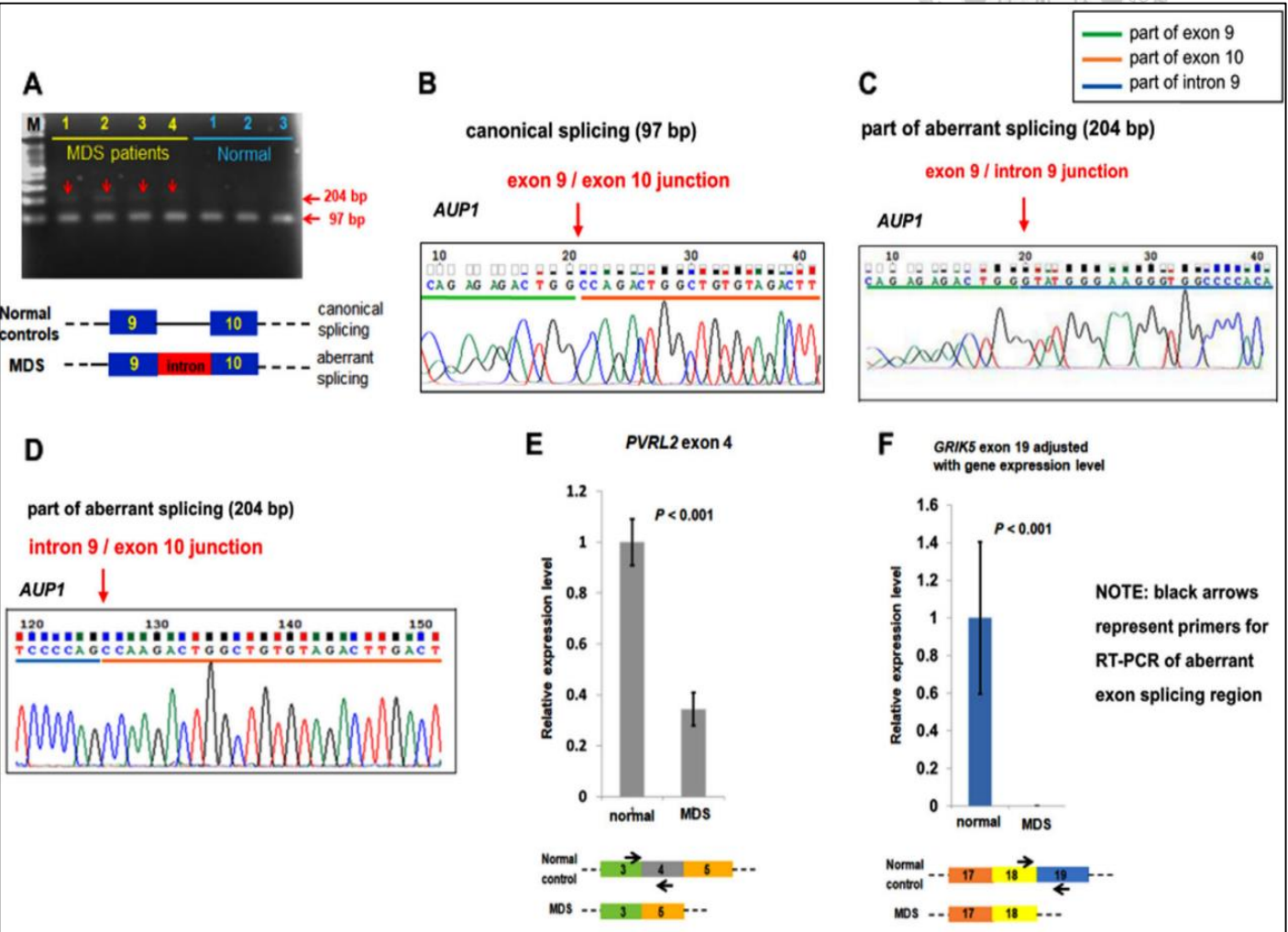


圖 16：藉由 Geneious 軟體預測 MDS 患者 *AUP1* 的變異選擇性剪接，將導致蛋白質截斷(protein truncation)

Source:
Yang YT et al. Am J Hematol.
2023;98(5):784-793

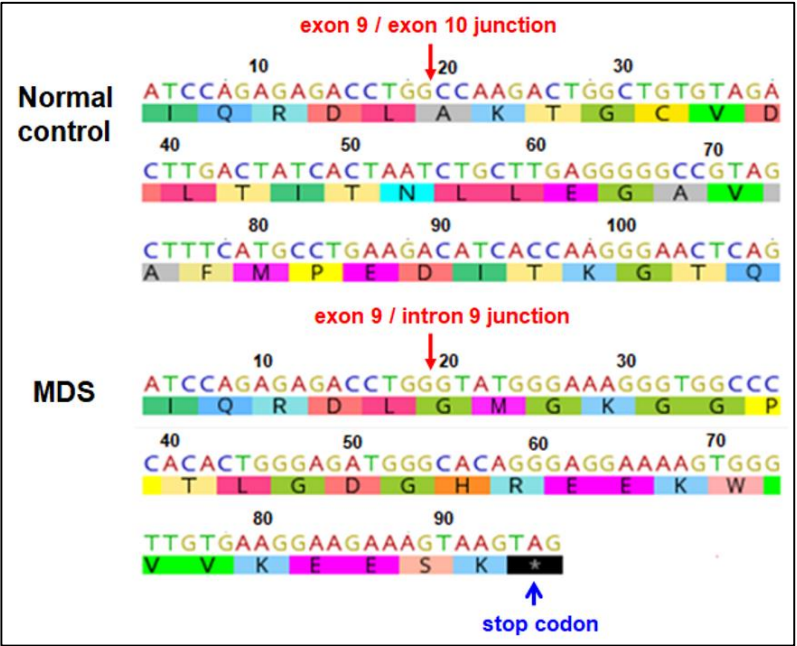
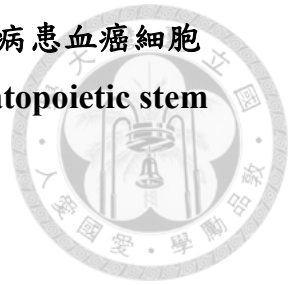
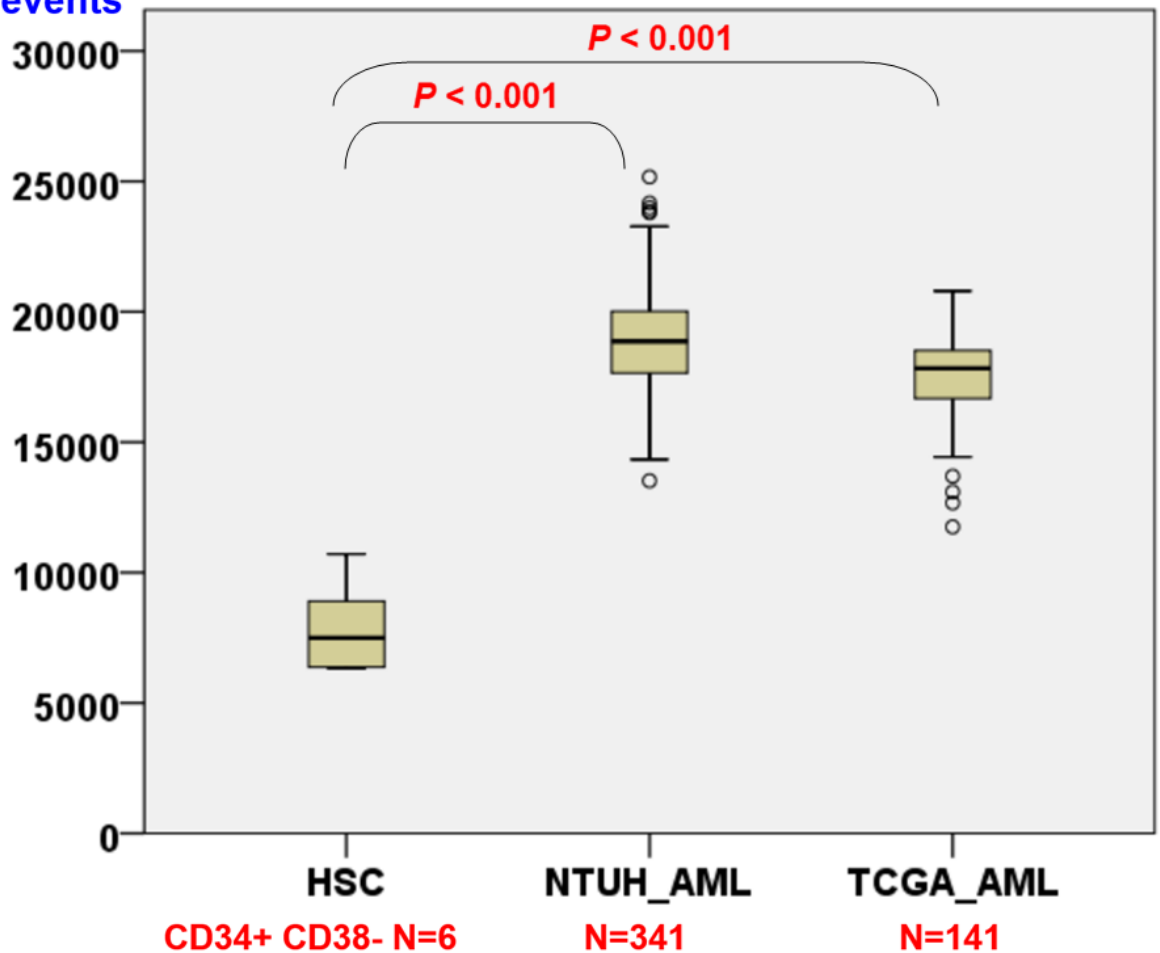


圖 17：比較轉錄體整體選擇性剪接模式(pattern)在 AML 病患血癌細胞 (NTUH and TCGA cohorts)與臍帶血造血細胞(HSC: hematopoietic stem cells)間的差異



Global AS events

(PSI 0.1~ 0.9)



Source:

Yang YT, et al. Am J Hematol. 2023;98(5):784-793.

圖 18：轉錄體整體選擇性剪接模式(pattern)在不同 AML subtype 間的差異

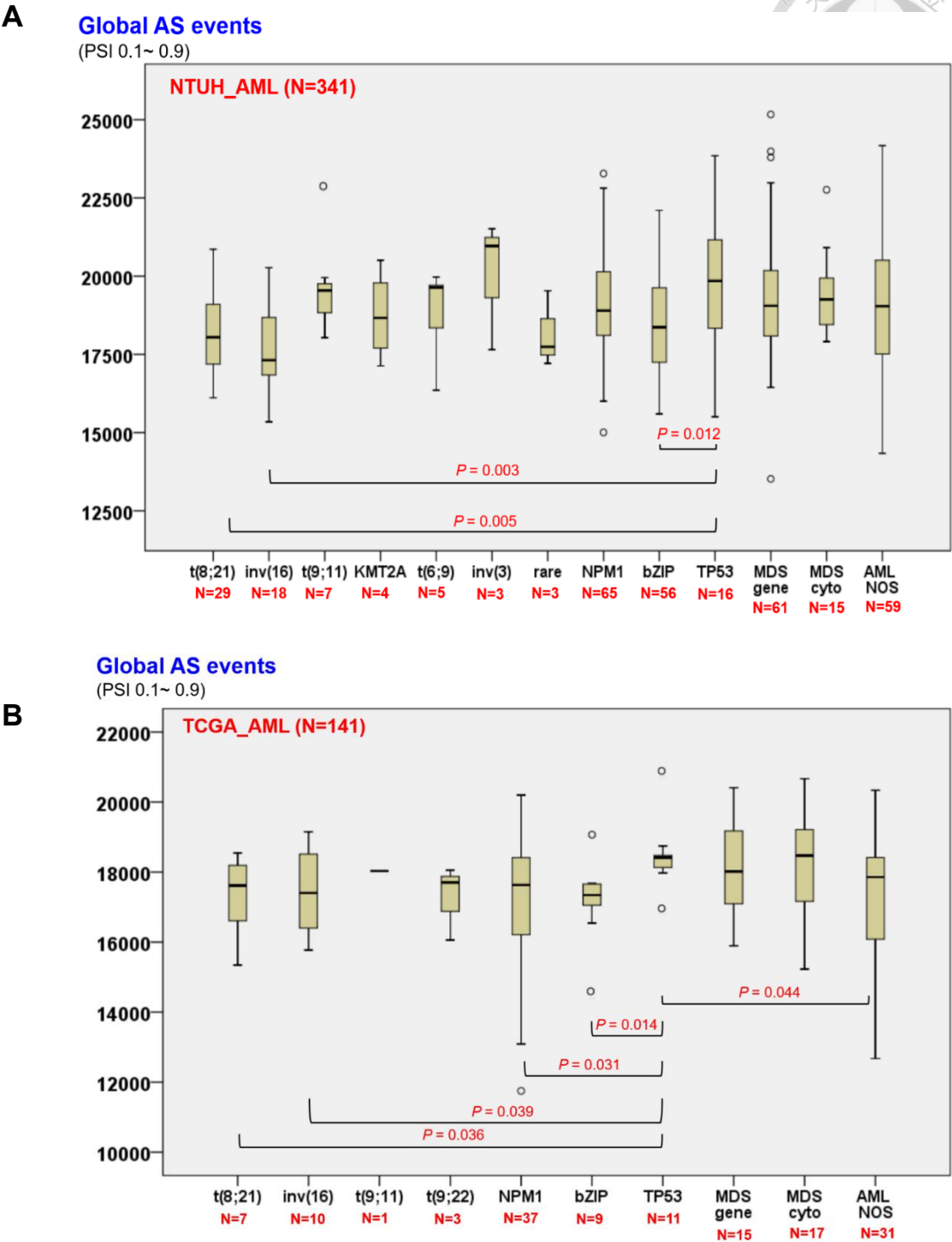
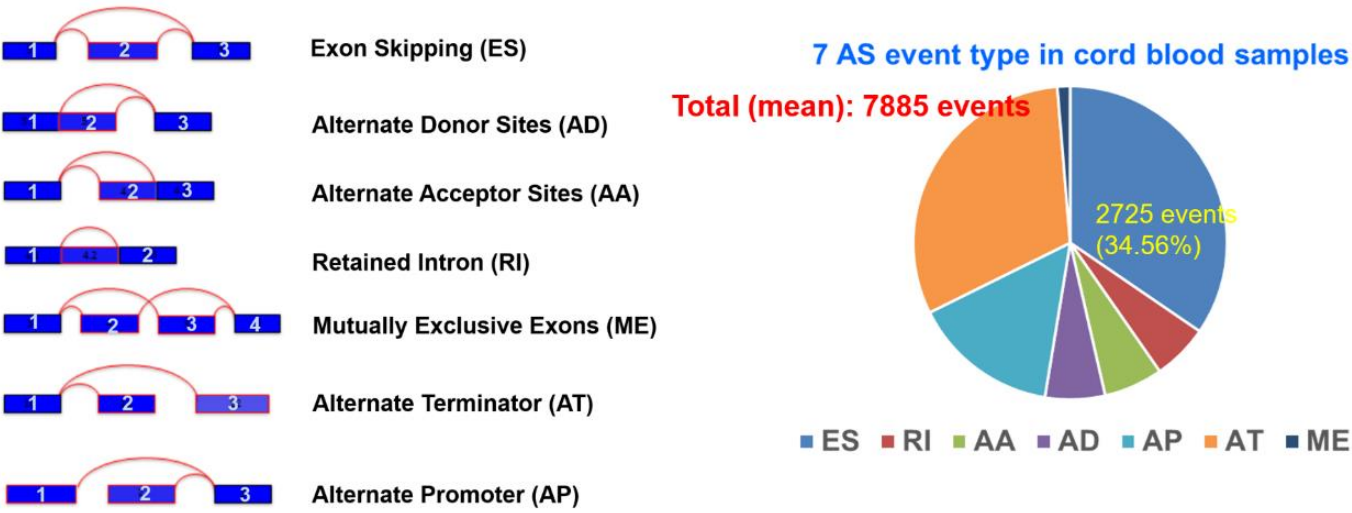


圖 18：比較轉錄體整體選擇性剪接模式(pattern)在不同 AML subtype(ICC classification)病患間的差異 NTUH cohort (A) TCGA cohort (B)。

Source:

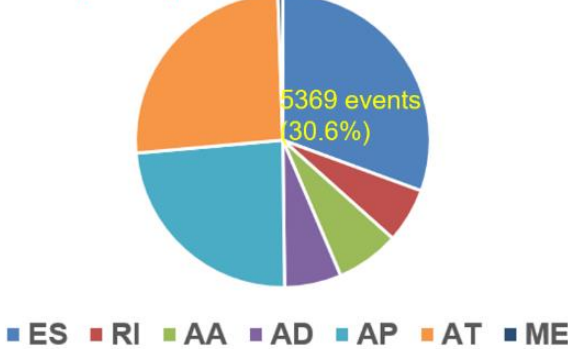
Yang YT et al. Am J Hematol. 2023;98(5):784-793

圖 19：分析 AML 病患(NTUH and TCGA cohorts)及 cord blood samples
整個轉錄體選擇性剪接的類型(event type)



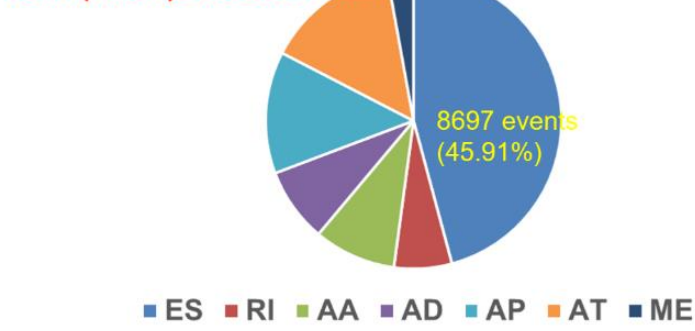
7 AS event type in TCGA_AML

Total (mean): 17551 events



7 AS event type in NTUH_AML

Total (mean): 18943 events



Source:
Yang YT, et al. Am J Hematol. 2023;98(5):784-793.

圖 20：AML 細胞基因體的基因突變數量似乎與轉錄體內所有基因選擇性剪接事件的數目沒有相關性(NTUH_AML cohort)

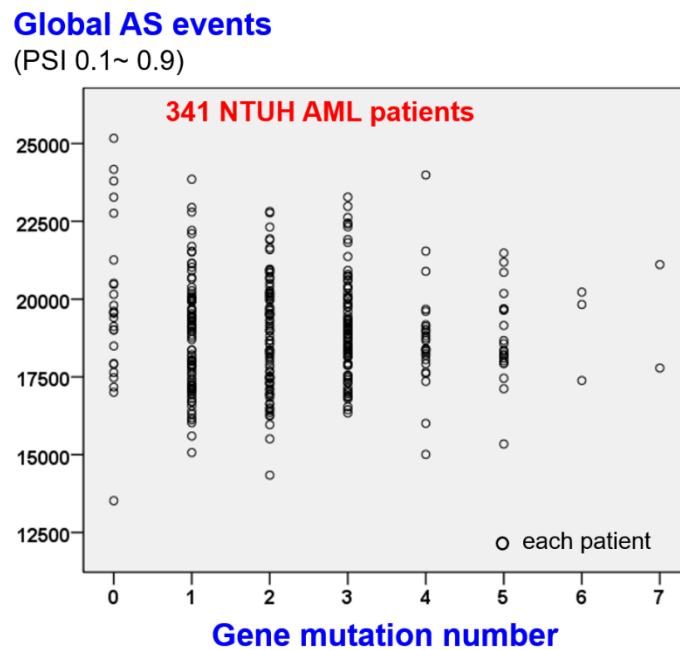
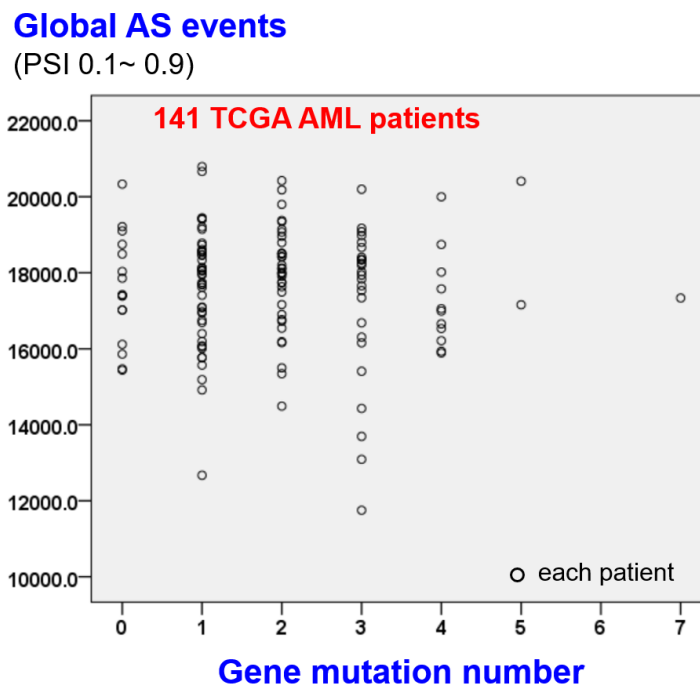


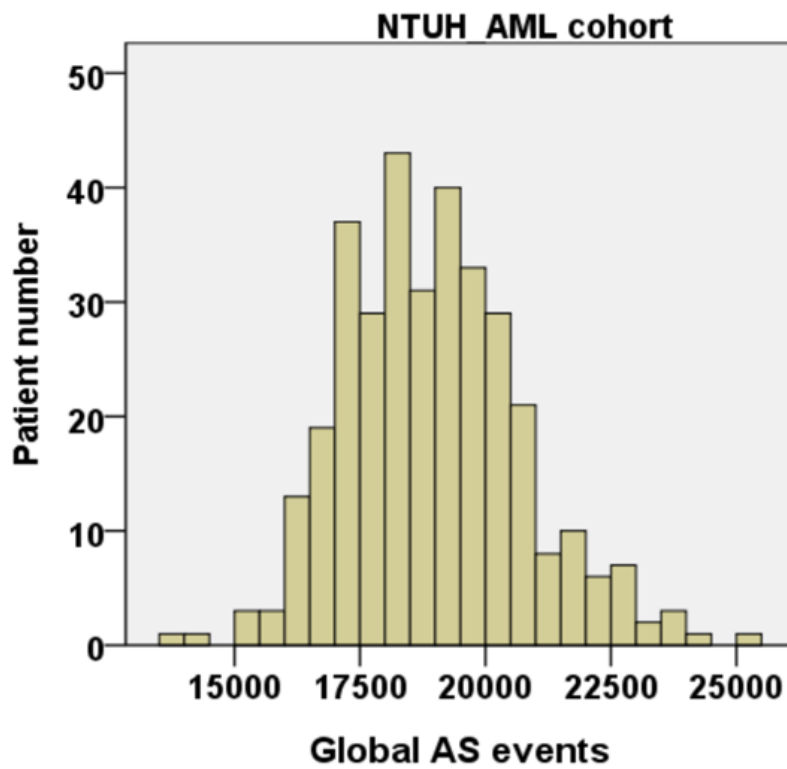
圖 21：AML 細胞基因體的基因突變數量似乎與轉錄體內所有基因選擇性剪接事件的數目沒有相關性(TCGA_AML cohort)



Source:

Yang YT et al. Am J Hematol. 2023;98(5):784-793.

圖 22：分析 AML 病患轉錄體中所有基因選擇性剪接事件的加總數目，
結果呈現鐘型曲線，近乎常態分佈



Source:

Yang YT, et al. Am J Hematol. 2023;98(5):784-793.

圖 23：不同的 global splicing pattern 對 NTUH_AML survival 的影響

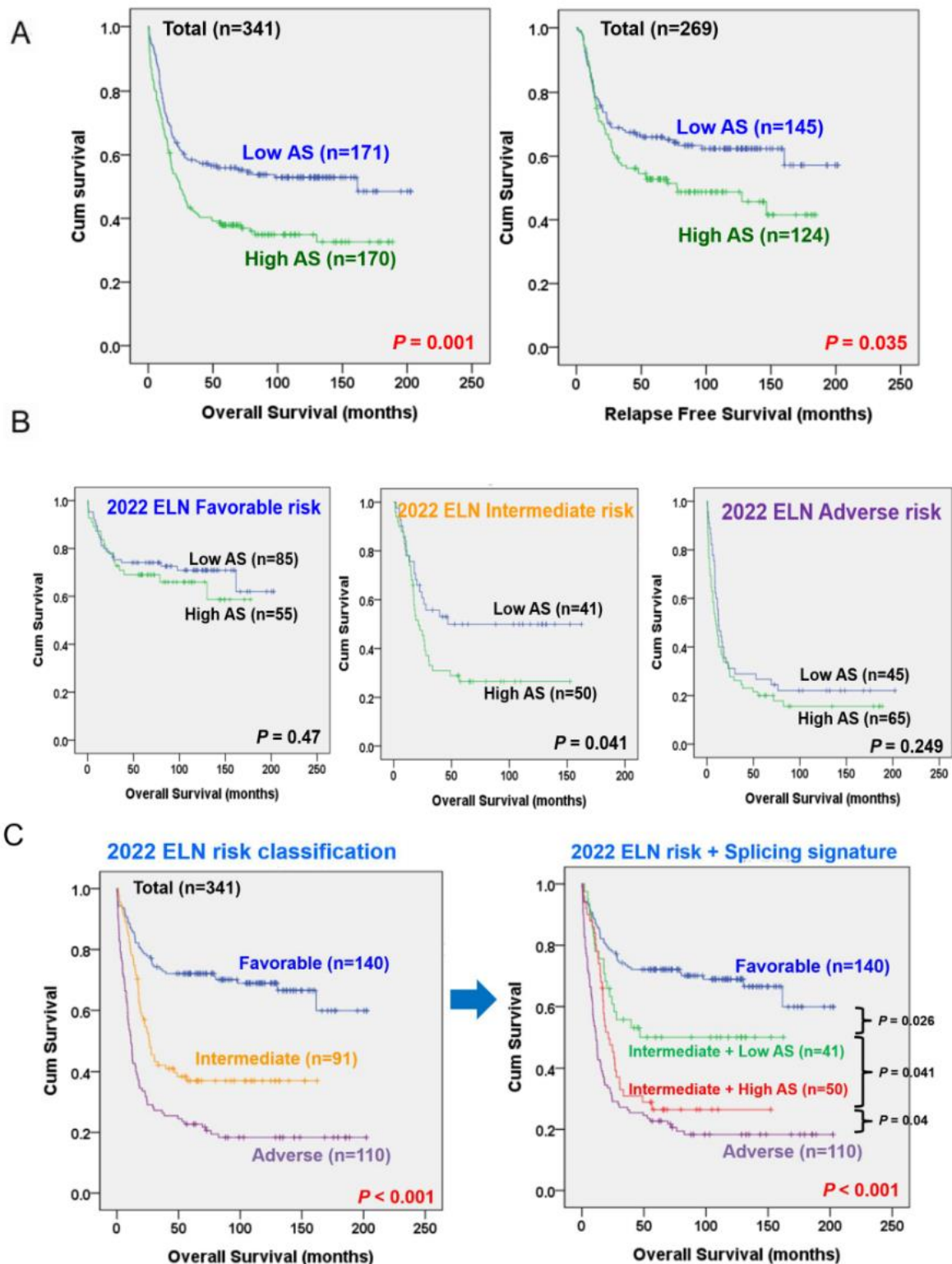
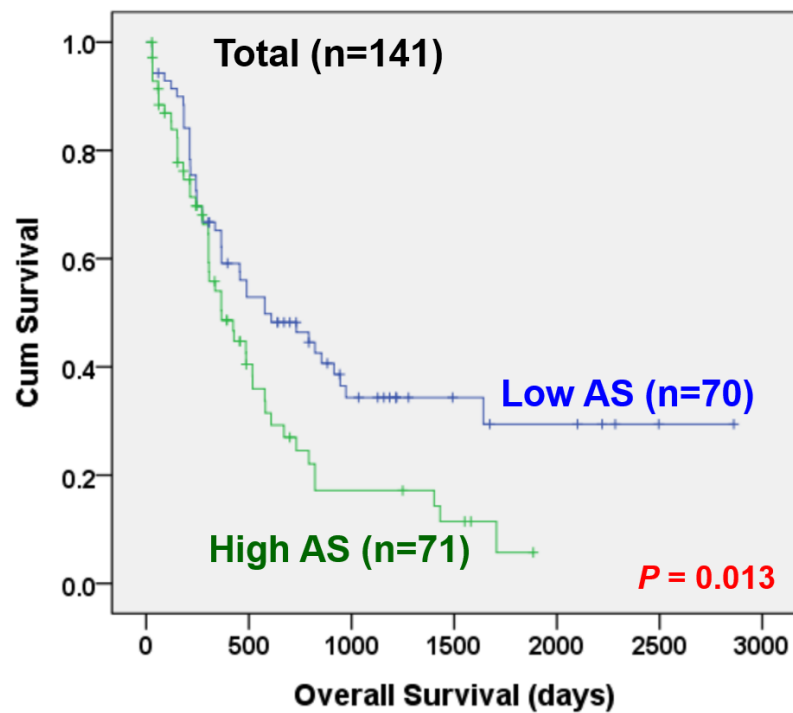


圖 23：比較 NTUH_AML 病患整個 cohort (A)及不同的 ELN risk group (B)，不同的 global splicing pattern 對 survival 的影響，若把 global splicing pattern 整併入 ELN risk classification，可 refine ELN risk model (C)。

Source: Yang YT et al. Am J Hematol. 2023;98(5):784-793.

圖 24：不同的 global splicing pattern 對 TCGA_AML 病患 survival 的影響



Source: Yang YT et al. Am J Hematol. 2023;98(5):784-793.

圖 25：分析 *NFYA* 基因在 AML 上 mRNA 及 protein 的表達

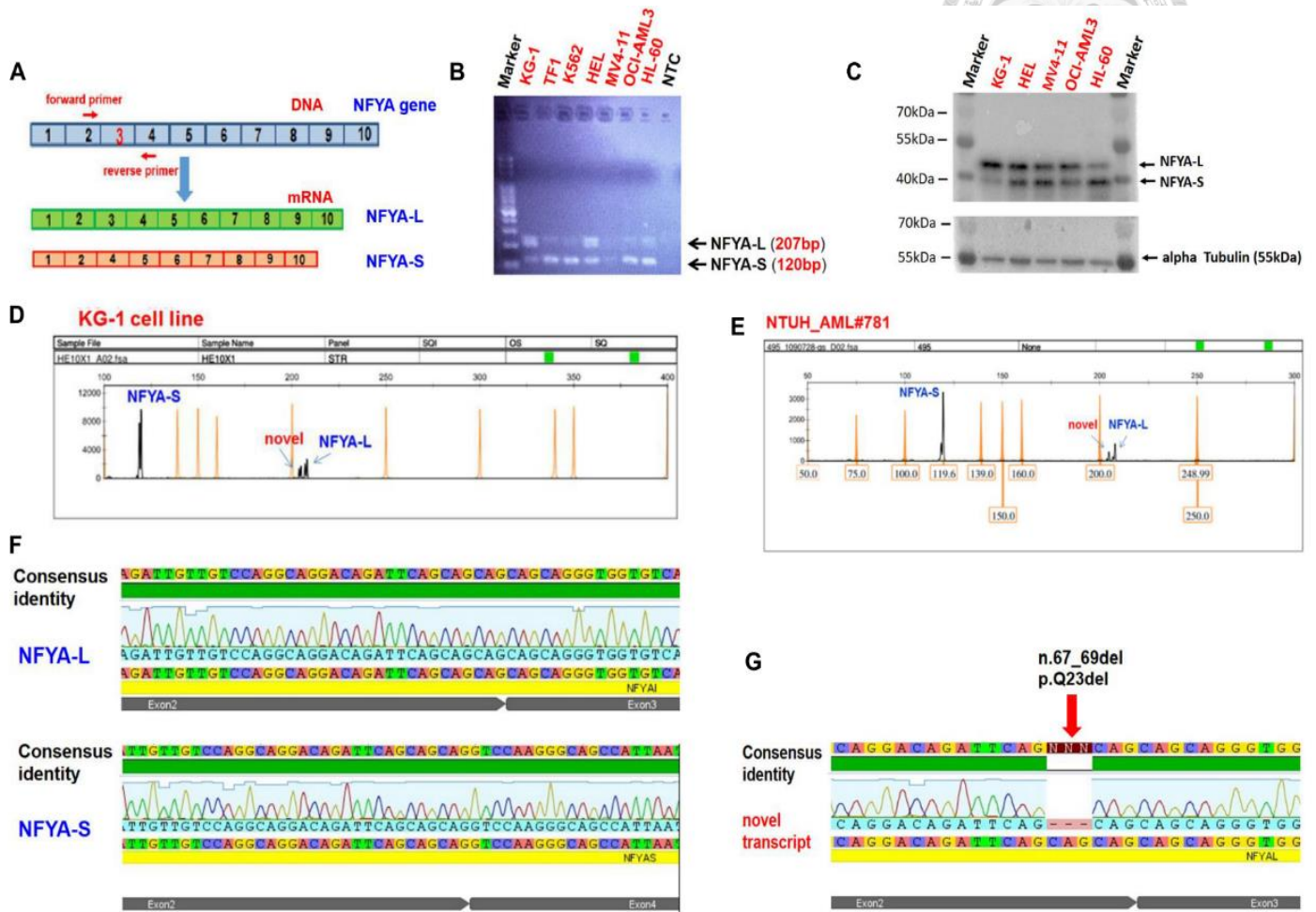


圖 25：分析 *NFYA* 基因在 AML 上 mRNA (A,B)及 protein (C)的表達，有兩個主要的異構體 *NFYA-L* 及 *NFYA-S*，另外意外發現了一個 novel transcript，經過 Sanger sequencing 定序，此 novel transcript 很靠近 *NFYA-L*，但比 *NFYA-L* 少了一個 codon：CAG (D-G)。

Note:

the forward primer, 5'-CAGCAATAGTTCGACAGAGCA-3', located in exon 2; and the reverse primer, 5'-GGTTGGCCAGTTGATGTGAT-3', positioned in exon 4. The reverse primer was labeled with the LIZ™ (orange) fluorophore at the 5'end.

The primary antibodies used for western blotting were anti-NFYA (G-2) (#sc-17753, Santa Cruz Biotechnology) diluted at 1:200 and alpha Tubulin antibody (#GTX628802, GeneTex, Inc.) diluted at 1:5000.

圖 26：NFYA 基因在 AML 血癌細胞與臍帶血的造血幹細胞(HSC) mRNA 表達量及 splicing pattern 的差異

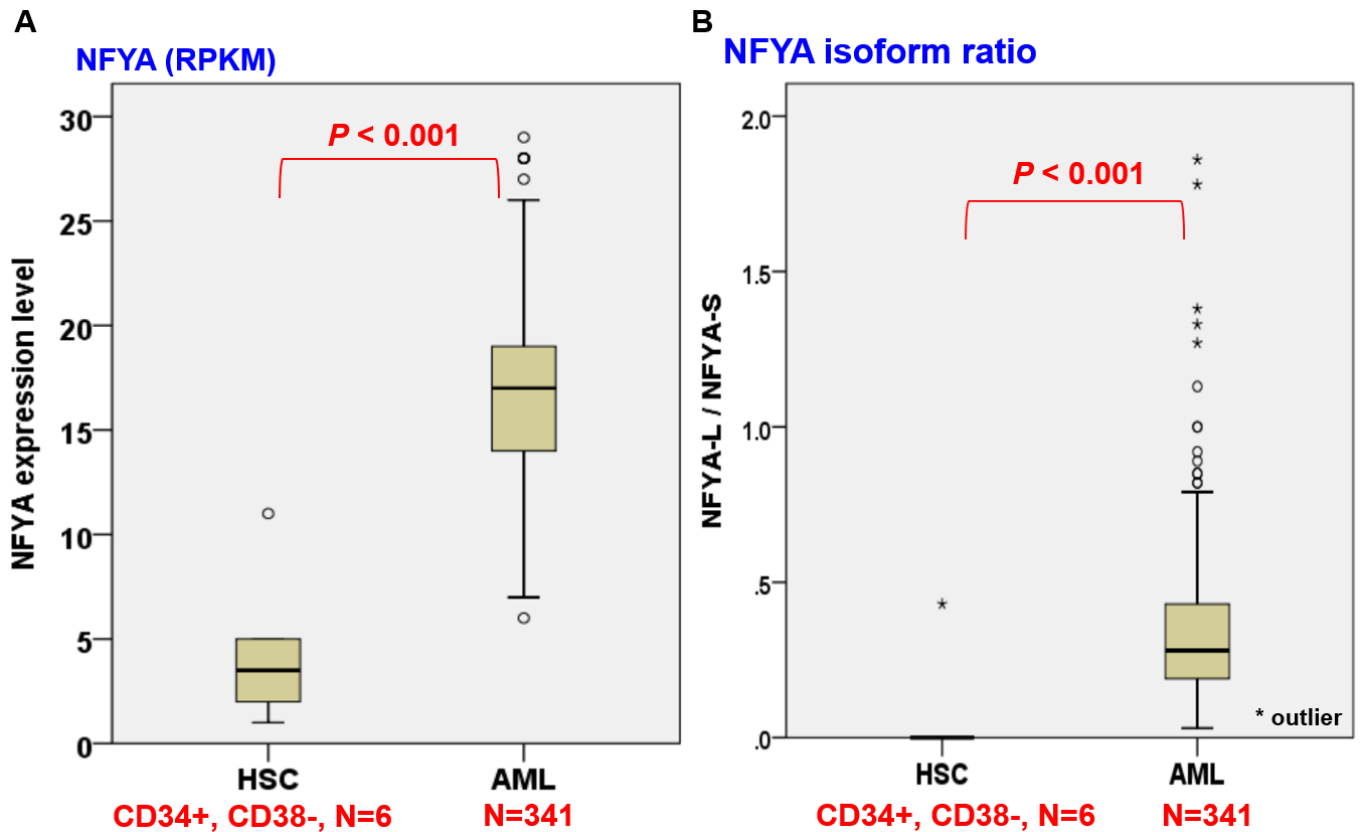
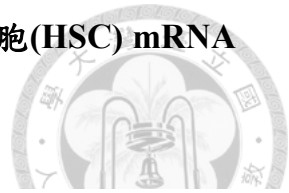


圖 26：比較 NFYA 基因在 AML 病患的血癌細胞與臍帶血的造血幹細胞 (HSC) 在 mRNA 表達量(A)及 splicing pattern(B)上的差異。

Note: RPKM stands for Reads Per Kilobase per Million. The ratios of NFYA-L to NFYA-S are used to represent the NFYA alternative splicing pattern.

圖 27：不同的 *NFYA* 及 *NFYA* 轉錄異構體表達量對 NTUH_AML survival 的影響

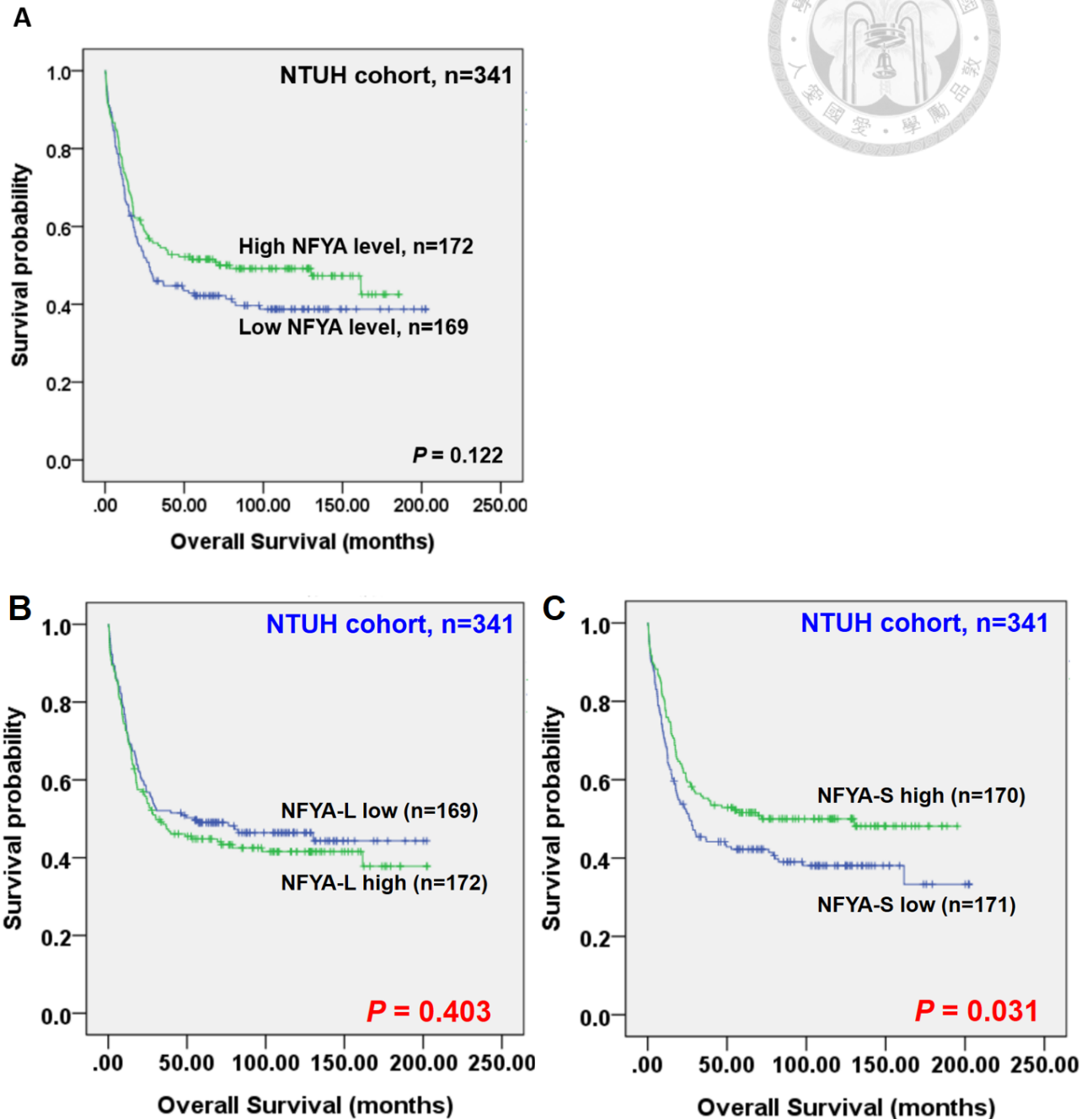


圖 27：比較 NTUH_AML 病患，不同的 *NFYA* 表達量(A)及不同的 *NFYA* 轉錄異構體表達量(B,C)對 survival 的影響。

Note: We employed the median gene expression level of *NFYA* as a threshold to classify patients into *NFYA* low- and high-expression groups. We used the median gene expression levels of *NFYA* isoforms as cutoff points to classify patients into *NFYA*-L or *NFYA*-S low- and high-expression groups.

圖 28：不同的 *NFYA* 轉錄異構體表達型式對 AML survival 的影響

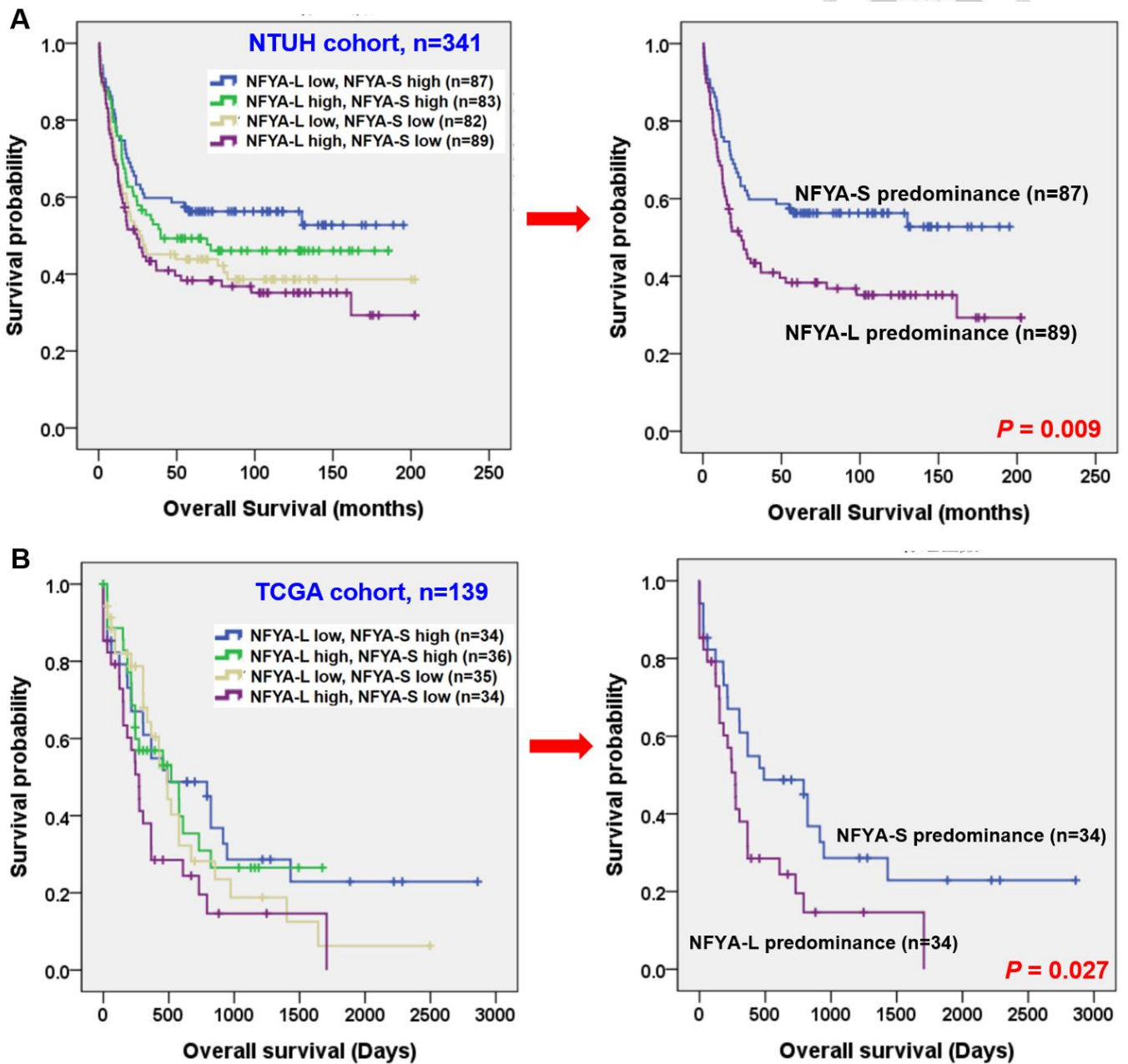


圖 28：比較 NTUH_AML 病患(A)，TCGA_AML 病患(B)，不同的 *NFYA* 轉錄異構體表達型式(splicing pattern)對 survival 的影響。

Note: We used the median gene expression levels of *NFYA* isoforms as cutoff points to classify patients into *NFYA*-L or *NFYA*-S low- and high-expression groups. *NFYA*-S predominance indicated AML patients with the *NFYA*-L low/*NFYA*-S high expression pattern, while *NFYA*-L predominance indicated AML patients with the *NFYA*-L high/*NFYA*-S low expression pattern.

圖 29：AML 病患不同的 *NFYA* 轉錄異構體表達型式對整體基因的表達差異

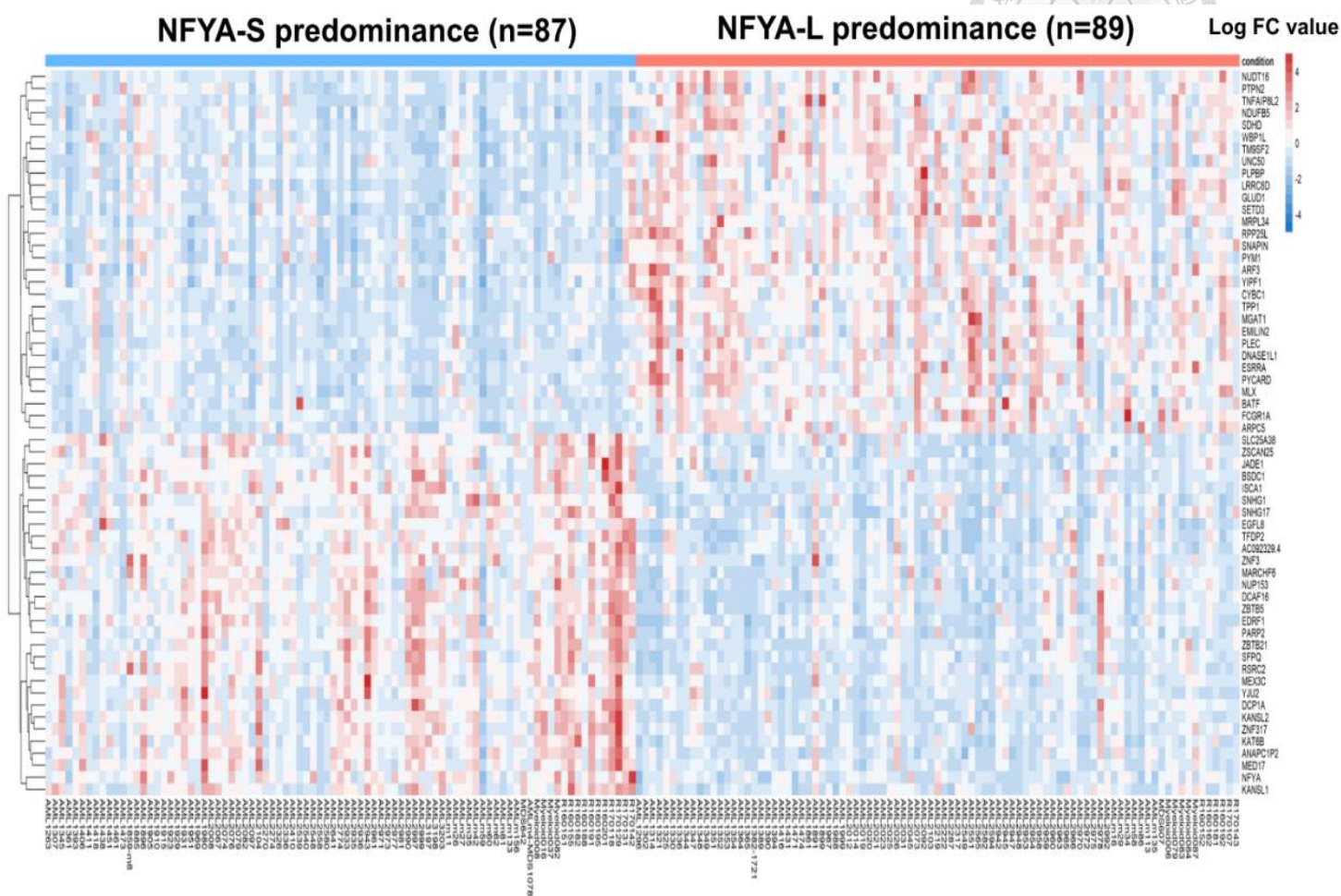


圖 29：比較 NTUH_AML 病患不同的 *NFYA* 轉錄異構體表達型式 (splicing pattern) 對整體基因的表達差異。

圖 30：AML 病患不同的 *NFYA* 轉錄異構體表達型式潛在的生物學效應

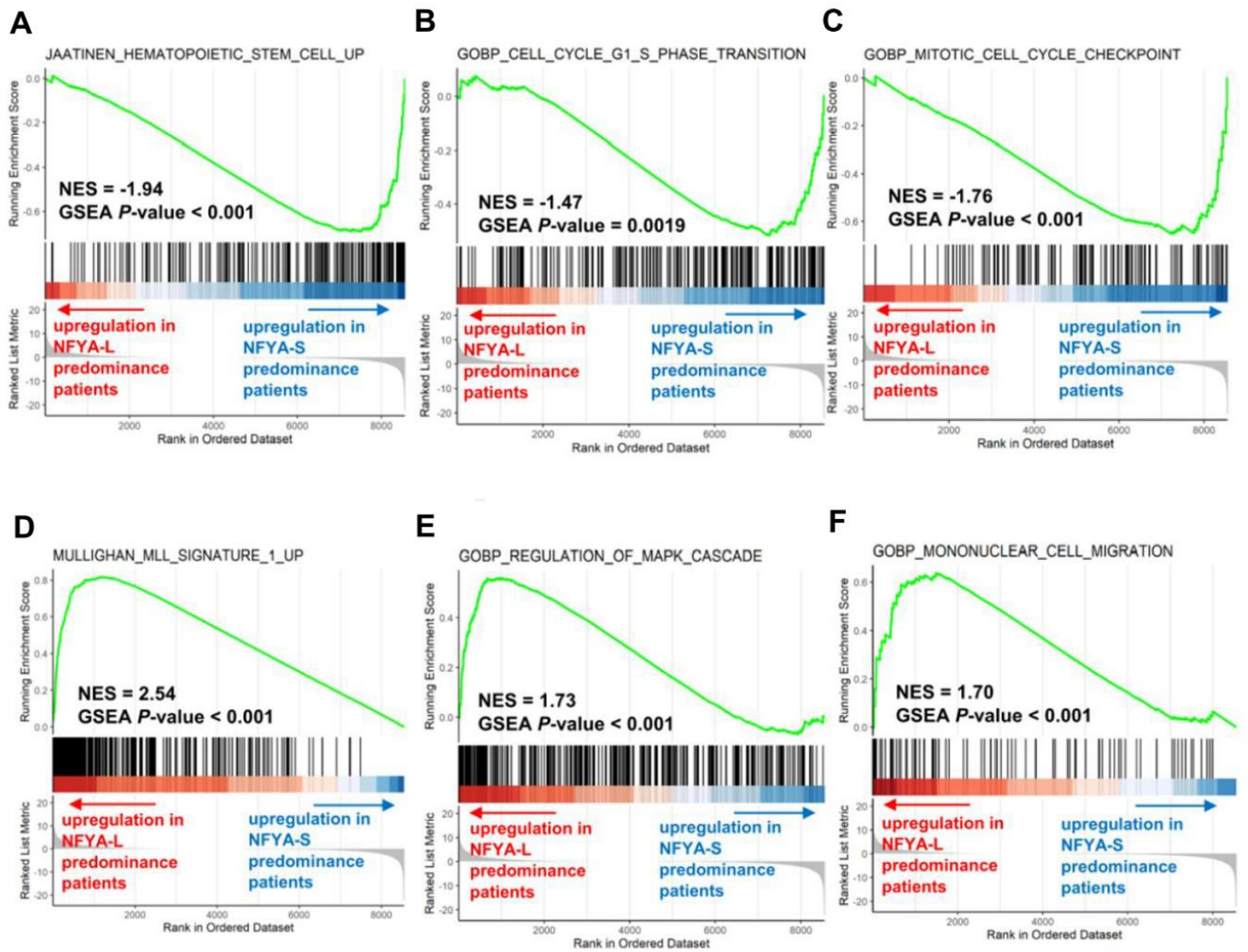


圖 30：比較 NTUH_AML 病患不同的 *NFYA* 轉錄異構體表達型式 (splicing pattern) 對潛在的生物學效應 (A-F)。

圖 31：OCI-AML3 細胞實驗的示意圖

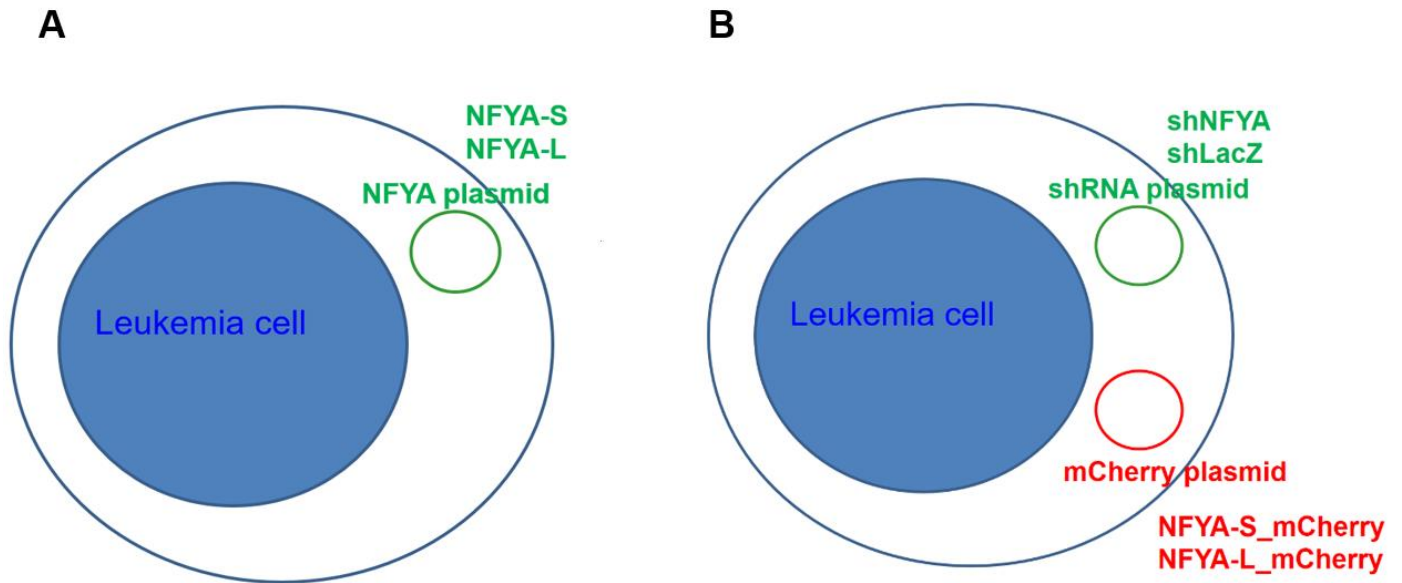


圖 31：OCI-AML3 細胞實驗的示意圖(A)分別把不同的 *NFYA* 轉錄異構體 cDNA 轉殖入 OCI-AML3 cell line (B)先把 shRNA 轉殖入 OCI-AML3 用以抑制內源性 *NFYA* 基因的表達，再分別把不同的 *NFYA* 轉錄異構體 cDNA 轉殖入 OCI-AML3 cell line。

Note:

The sequence of shNFYA was 5'-CTTTCAGTCAGGACCTATTTC-3'.

An shRNA vector targeting the LacZ gene (shLacZ) was used as a control. The sequence of shLacZ was 5'-CGCGATCGTAATCACCCGAGT-3'.

圖 32：NFYA 基因選擇性剪接在 OCI-AML3 白血病細胞株中的生物效應

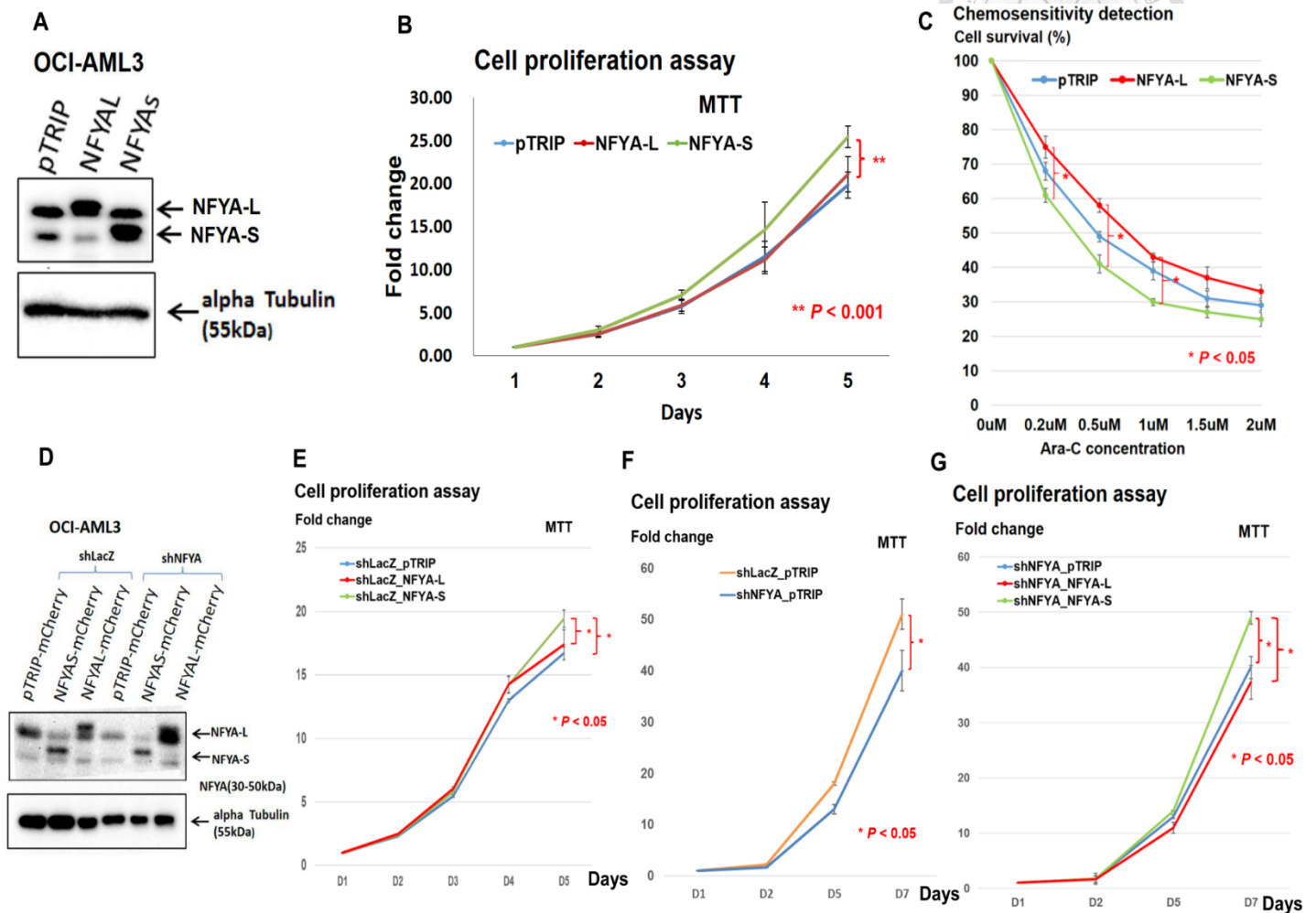


圖 32：(A)利用接受不同 *NFYA* 剪接異構體 lentiviral vector 處理後的 OCI-AML3 細胞株來進一步做實驗。(B)在 OCI-AML3 細胞中過度表達 *NFYA-S* 將導致細胞增殖顯著增加，而過度表達 *NFYA-L* 對細胞增殖沒有明顯影響。(C)評估這些不同 OCI-AML3 細胞株對 Ara-C 化療的敏感性。在 48 小時不同濃度的 Ara-C 處理後，過度表達 *NFYA-S* 的 OCI-AML3 細胞表現出比對照組和 *NFYA-L* 過度表達細胞有更大的化療敏感性。(D)為了消除 *NFYA* 原本內源性異構體的干擾，我們先對 *NFYA* 基因進行敲除(knockdown)。在敲除後，我們再利用上述方法表達 *NFYA-S* 或 *NFYA-L*。(E)當我們只對 *LacZ* control 基因敲除，結果顯示如上述。(F)當 *NFYA* 基因被敲除時，OCI-AML3 細胞的增殖明顯減弱。(G)在表達 *NFYA-S* 後，OCI-AML3 細胞的增殖顯示出恢復的跡象，而在表達 *NFYA-L* 後卻未觀察到這種情況。

Note: All data points represent results from three independent experiments performed in triplicate and are presented as mean $\pm$ SD. The expression of lentiviral vectors with the FLAG-tag led to slight differences in the translated proteins compared to their endogenous counterparts.

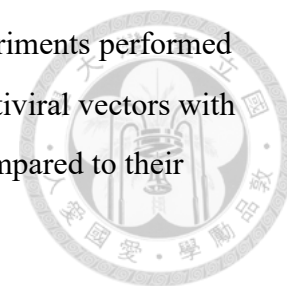
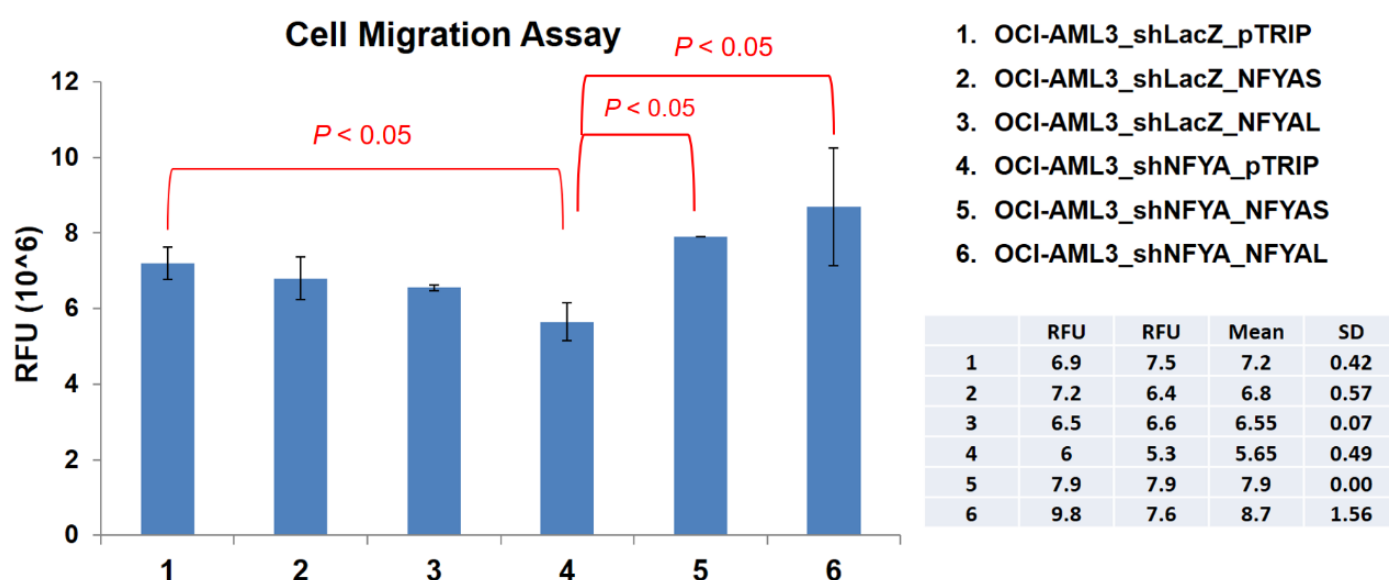


圖 33：OCI-AML3 不同的 *NFYA* 基因選擇性剪接表達型式(splicing pattern)對細胞的遷移能力(migration ability)的影響



Abbreviations: Relative fluorescence units (RFU)

圖 33：當 *NFYA* 基因被敲除(knockdown)時，OCI-AML3 細胞的遷移能力(migration ability)明顯減弱。然而 *NFYA-S* 和 *NFYA-L* 重新表達後均能使 OCI-AML3 細胞的遷移能力恢復。

Note: All results are derived from three independent experiments conducted in triplicate, and data are presented as mean $\pm$ SD.

圖 34：OCI-AML3 不同的 *NFYA* 基因選擇性剪接表達型式(splicing pattern)對細胞週期分佈的影響

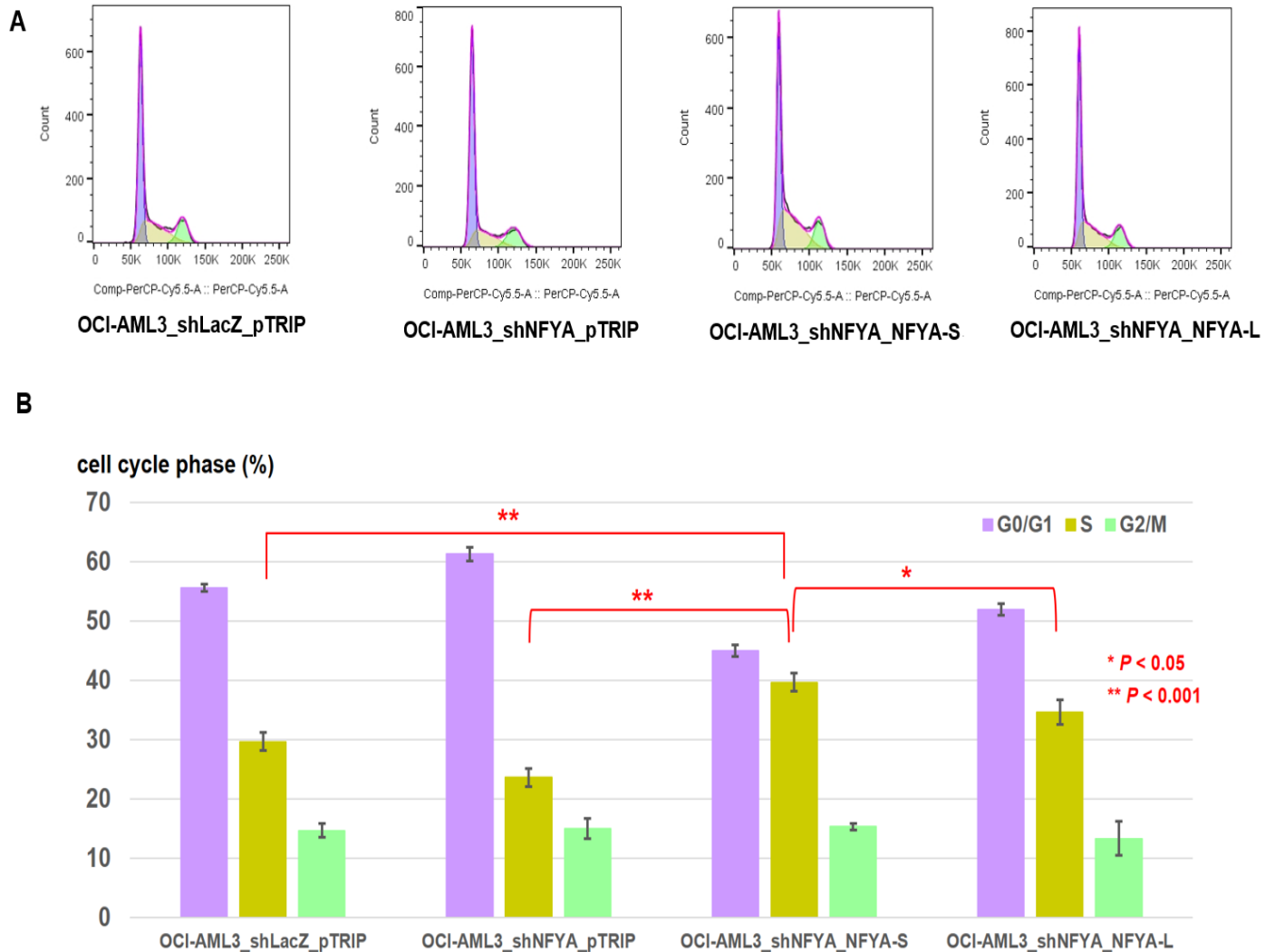
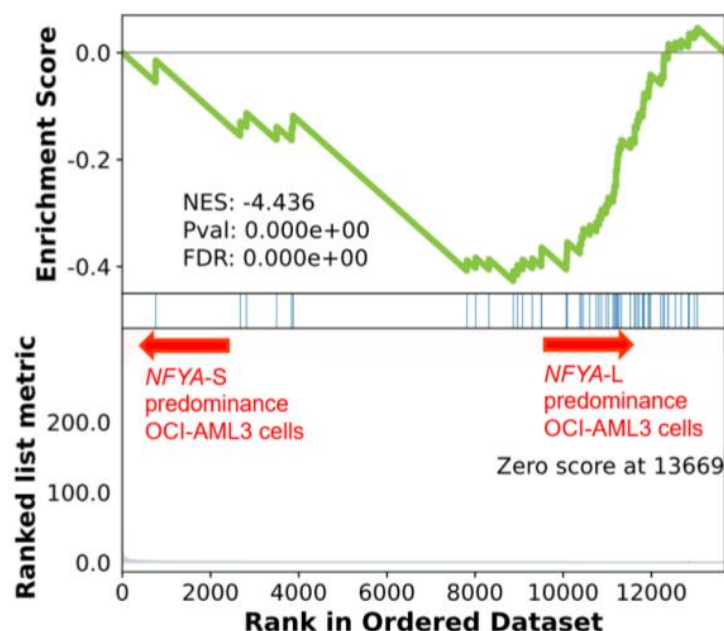


圖 34：細胞週期實驗分析中，當 *NFYA* 基因的敲除(knockdown)會導致 OCI-AML3 細胞 S 期細胞比例稍微減少。然而，在我們將 *NFYA-S* 重新表達後，S 期細胞的比例顯著增加。但重新表達 *NFYA-L* 後，S 期細胞的比例增加幅度較不明顯。(A)透過流式細胞儀檢測，OCI-AML3 細胞群組中 G0/G1、S 和 G2/M 期的細胞數(圖示為某次實驗的結果)。(B)直方圖顯示在三個獨立實驗中，不同 OCI-AML3 細胞群組中 G0/G1、S 和 G2/M 期細胞的百分比。數據以平均值 ± 標準差呈現。

圖 35：OCI-AML3 細胞不同的 *NFYA* 轉錄異構體表達型式潛在的生物學效應

A negative regulation of G2/M transition of mitotic cell cycle (GO:0010972)



B Oxidative Phosphorylation (GO:0006119)

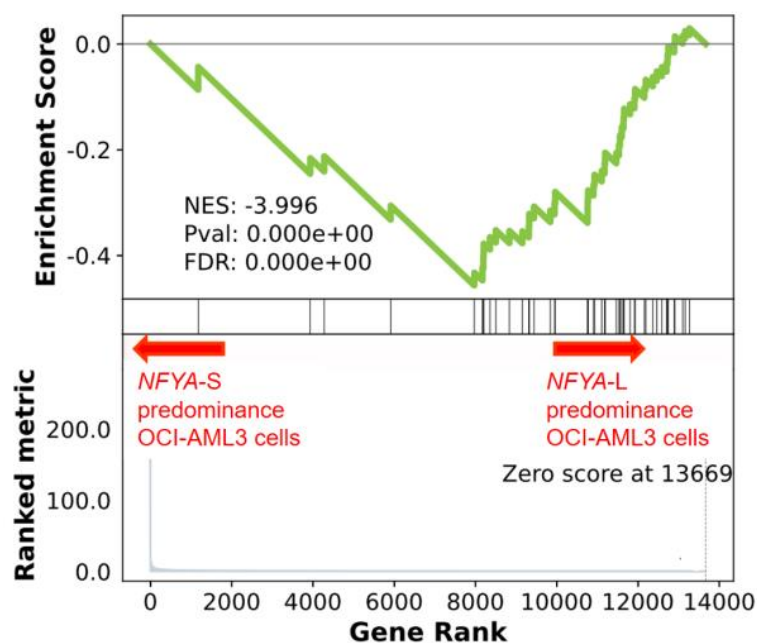


圖 35：基因集富集分析(GSEA)顯示主要表達 *NFYA-L* 的 OCI-AML3 細胞會抑制細胞週期中進展至 G2/M 調控的相關基因表達(A)，但會活化細胞代謝路徑 oxidative phosphorylation 的相關基因表達(B)。

圖 36： NFY 轉錄因子對目標基因及細胞生理路徑的影響示意圖；*NFYA* 基因兩個主要的剪接異構體(alternatively spliced isoforms)示意圖

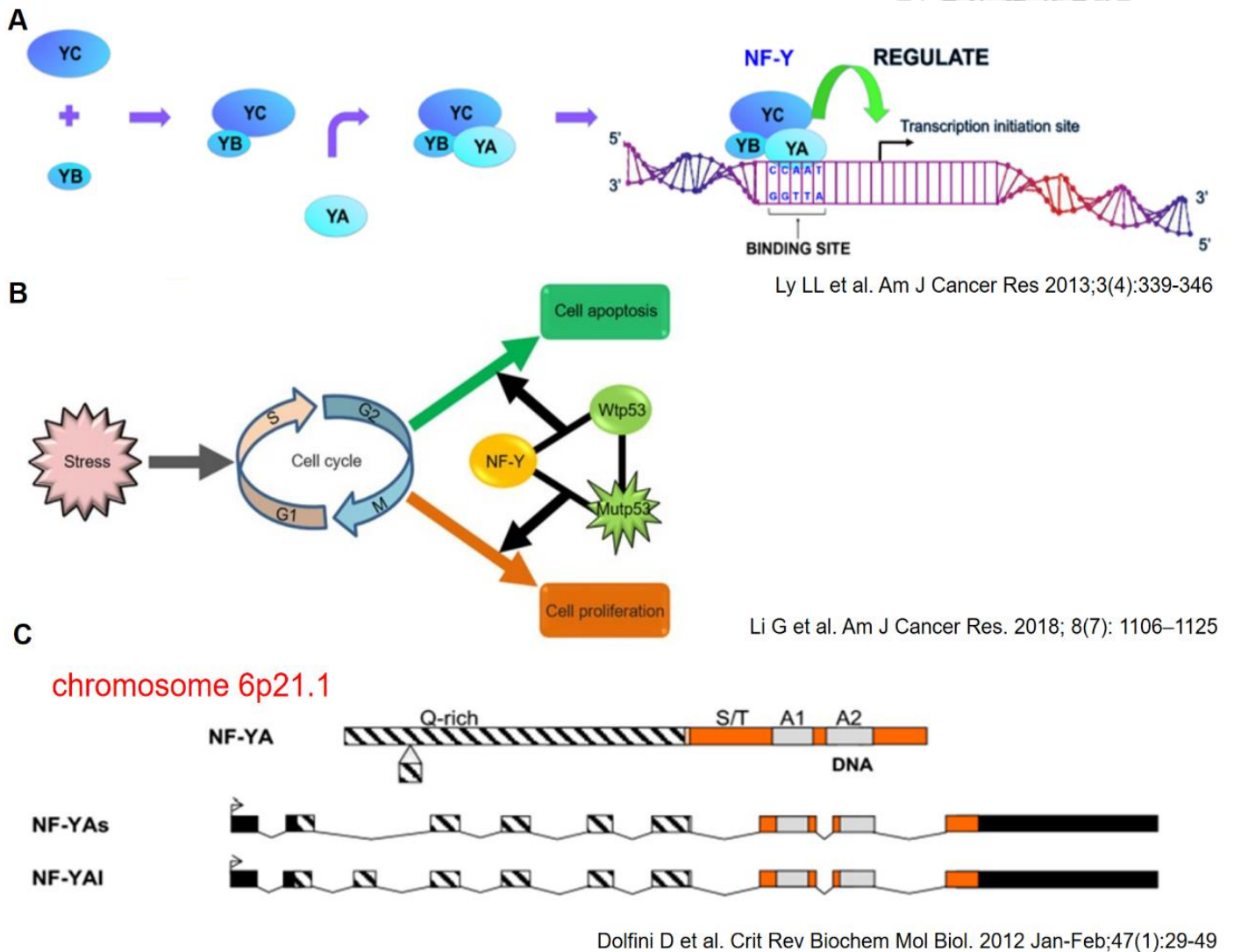


圖 36：(A) Nuclear Transcription Factor Y (NFY)是一種普遍存在的異源三聚體轉錄因子(heterotrimeric transcription factor)，它會結合到其目標基因的啟動子序列 CCAAT box，讓目標基因表達。(B) NFY 在細胞週期調控以及細胞凋亡和增殖的平衡中發揮重要作用。(C) *NFYA* 基因有兩個主要的剪接異構體(alternatively spliced isoforms)。 *NFYA* 長型(*NFYA-L*)包含完整的 10 個 exons，而 *NFYA* 短型(*NFYA-S*)在 pre-mRNA 剪接時 exon 3 skipping，導致 N 端的轉錄啟動區域(transactivation domain)缺少了 29 個胺基酸。

圖 37：不同 *NFYA*-L 及 *NFYA*-S 表達的白血病患族群中特定蛋白質酪胺酸激酶基因表達量差異

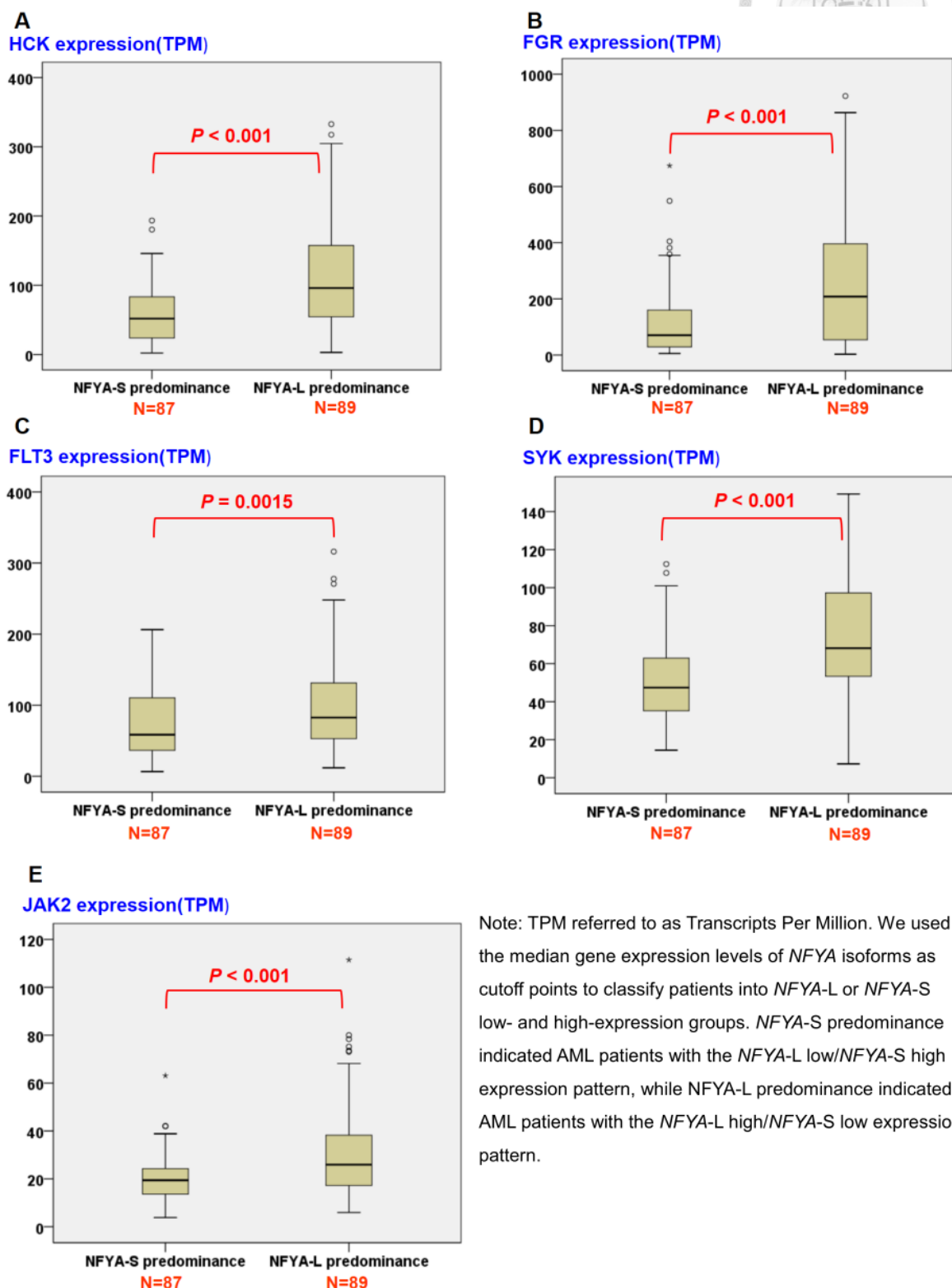


圖 37 使用 RNA-seq 資料分析 NTUH 兩群不同 *NFYA*-L 及 *NFYA*-S 表達的白血病患 *HCK*(A)、*FGR*(B)、*FLT3*(C)、*SYK*(D)和 *JAK2*(E)的基因表達量差異。

表 1：MDS 患者的臨床基本資料(clinical characteristics of the patients)

Characteristics	Total number
Age (ranged from 19 to 94 years, median 69 years)	
≥ 65 years	102 (58%)
< 65 years	74 (42%)
Gender	
Male	121 (69%)
Female	55 (31%)
FAB	
RA	78 (44%)
RARS	22 (13%)
RAEB	76 (43%)
WHO 2016	
MDS-SLD	40 (23%)
MDS-MLD	38 (22%)
MDS-RS-SLD	13 (7%)
MDS-RS-MLD	9 (5%)
MDS-EB1	32 (18%)
MDS-EB2	44 (25%)
IPSS-R (only 164 patients could be evaluated)	
Very low	6 (4%)
Low	54 (33%)
Intermediate	41 (25%)
High	35 (21%)
Very high	28 (17%)

Abbreviations:

FAB, French-American-British classification;

WHO, World Health Organization;

IPSS-R, revised international prognostic scoring system;

MDS, myelodysplastic syndrome;

SLD, single lineage dysplasia; MLD, multilineage dysplasia;

RS, ring sideroblasts; EB, excess blasts.

Source:

Yang YT, et al. Blood Cancer J. 2018;8(8):78.



表 2：分析 MDS 病患整個轉錄體變異選擇性剪接在基因體上所影響的位置

Analysis of aberrantly spliced transcripts among 176 MDS patients

MDS patient cohort	Splicing events mapped to the coding regions of the genome	Splicing events mapped to the non-coding regions of the genome
RCMD	16213 of 18282 (88.7%)	2069 of 18282 (11.3%)
RA	15727 of 17735 (88.7%)	2008 of 17735 (11.3%)
RCMDRS	16734 of 19018 (88.0%)	2284 of 19018 (12.0%)
RARS	16320 of 18864 (86.5%)	2544 of 18864 (13.5%)
RAEB-1	15893 of 17980 (88.4%)	2087 of 17980 (11.6%)
RAEB-2	15643 of 17657 (88.6%)	2014 of 17657 (11.4%)

Splicing events were obtained as compared with normal donor samples ($P < 0.05$)

表 3：
比較 MDS 病患在變
異選擇性剪接模式
高變異度及低變異
度的臨床表徵差異

Variables	Total (n)	High aberrant AS score (n)	Low aberrant AS score (n)	P value
<i>Sex^a</i>				0.024*
Male	121	87	34	
Female	55	30	25	
<i>Age (year)^b</i>	69 (19–94)	66 (19–89)	73 (26–94)	0.121
≥65	102	63	39	
<65	74	54	20	
<i>Laboratory data^b</i>				
WBC (/μL)	3825 (490–20440)	3780 (490–20440)	4660 (1710–11690)	0.426
Hb (g/dL)	8.1 (3.5–14.6)	8.1 (3.5–13.6)	8.1 (3.7–14.6)	0.792
Platelet (×1000 /μL)	85 (3–721)	82 (9–721)	106 (3–460)	0.08
<i>Cytogenetics^a</i>				
Favorable ^c	115	71	44	0.09
Intermediate ^d	23	17	6	0.382
Poor ^e	26	20	6	0.197
<i>Genetic alterations^a</i>				
<i>SF3B1</i>	29	17	12	0.342
<i>U2AF1</i>	14	13	1	0.028*
<i>SRSF2</i>	24	15	9	0.675
<i>ZRSR2</i>	15	12	3	0.254
<i>TET2</i>	22	17	5	0.246
<i>ASXL1</i>	36	26	10	0.386
<i>DNMT3A</i>	25	18	7	0.517
<i>EZH2</i>	10	7	3	0.809
<i>IDH1/IDH2</i>	4	4	0	0.151
<i>TP53</i>	12	10	2	0.193
<i>RUNX1</i>	25	18	7	0.517

*Statistically significant ($P < 0.05$)

^aNumber of patients

^bMedian (range)

^cFavorable cytogenetics: -Y, del(11q), Normal, del(5q), del(12p), del(20q), double including del(5q)

^dIntermediate-risk cytogenetics: del(7q), +8, +19, i(17q), any other single or double independent clones

^ePoor-risk cytogenetics: -7, inv(3)/t(3q)/del(3q), double including -7/de(7q), complex: 3 abnormalities, complex: >3 abnormalities

Source:

Yang YT, et al. Blood Cancer J. 2018;8(8):78.

表 4：利用多變數分析(multivariate analysis) MDS 病患的預後因子，顯示變異選擇性剪接模式是一個獨立的預後因子



Variables	Overall survival				Time to leukemic change			
	HR	95% CI		<i>P</i>	HR	95% CI		<i>P</i>
		Lower	Upper			Lower	Upper	
Age ^a	1.019	1.003	1.035	0.017*	0.984	0.965	1.005	0.133
IPSS-R ^b	3.114	1.839	5.270	<0.001*	3.694	1.710	7.982	0.001*
<i>SRSF2</i>	1.092	0.501	2.379	0.825	0.913	0.294	2.837	0.875
<i>ZRSR2</i>	1.400	0.665	2.950	0.376	0.690	0.212	2.242	0.537
<i>ASXL1</i>	1.628	0.784	3.381	0.191	4.788	1.933	11.857	0.001*
<i>EZH2</i>	1.227	0.456	3.300	0.685	0.701	0.171	2.881	0.622
<i>TP53</i>	4.563	1.936	10.753	0.001*	5.009	1.250	20.079	0.023*
Aberrant AS score ^c	1.801	1.019	3.181	0.043*	2.736	1.036	7.222	0.042*

HR hazard ratio, CI confidence interval

*Statistically significant ($P < 0.05$)

^aAge as a continuous variable

^bIPSS-R risk score > 4.5 relative to IPSS-R ≤ 4.5 (the reference)

^cHigh global aberrant AS score relative to low global aberrant AS score (the reference)

Source:

Yang YT, et al. Blood Cancer J. 2018;8(8):78.

表 5：分析 MDS 病患整個轉錄體變異選擇性剪接對細胞生理路徑的影響

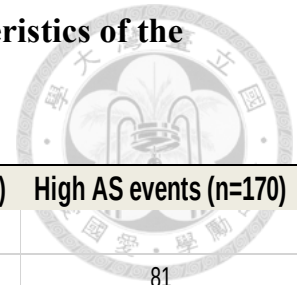
Term-pathways	Count	P-value	FDR	Fold enrichment
Porphyrin biosynthesis	8	1.14E-07	1.65E-04	13.05
Protein refolding	11	3.75E-08	7.08E-05	8.64
Heme biosynthesis	8	3.23E-05	0.047	7.46
Hemoglobin's Chaperone	10	5.74E-06	0.008	5.66
Chromosome segregation	19	9.65E-06	0.018	3.29
DNA replication	22	5.61E-06	0.008	3.09
Mitosis	51	8.65E-10	1.26E-06	2.58
Lysosome	44	6.86E-07	9.95E-04	2.26
Cell division	66	9.43E-10	1.37E-06	2.25
Cell cycle	100	1.91E-11	2.77E-08	2.02
Immunity	70	3.43E-08	4.98E-05	2.01
Cell-cell adherens junction	51	9.44E-06	0.014	1.92
Cytoskeleton	57	1.08E-05	0.016	1.84
Focal adhesion	59	9.89E-06	0.015	1.82
Protein kinase binding	58	1.91E-05	0.032	1.79
Ubl conjugation	229	1.27E-15	1.78E-12	1.68
ATP-binding	178	5.79E-12	8.40E-09	1.67
Acetylation	423	1.72E-26	2.50E-23	1.61
Protein binding	681	3.63E-18	5.99E-15	1.28
Splice variant	752	1.53E-16	2.11E-13	1.26

Abbreviation: FDR, false discovery rate.

Source:

Yang YT, et al. Blood Cancer J. 2018;8(8):78.

表 6：NTUH_AML 患者的臨床基本資料(clinical characteristics of the patients)



	Total (n=341)	Low AS events (n=171)	High AS events (n=170)	P value
sex				0.483
Male	169	88	81	
Female	172	83	89	
Age[†]	47 (16-84)	47 (18-84)	48 (16-77)	0.706
Lab data[†]				
WBC (×10 ³ /μL)	2.73 (0.49-405.65)	41.89 (0.52-405.65)	17.24 (0.49-352.83)	0.001
Hb (g/dL)	8.2 (2.4-16)	8.1 (2.7-13.6)	8.3(2.4-16)	0.305
Platelett (×10 ³ /μL)	45 (3-751)	43 (3-751)	49 (3-712)	0.418
LDH (U/L)	761 (129-8693)	919 (150-8693)	560 (129-6977)	< 0.001
FAB				
M0	10	3	7	0.197
M1	72	38	34	0.616
M2	120	57	63	0.473
M4	104	55	49	0.504
M5	19	12	7	0.244
M6	14	5	9	0.271
M7	0	0	0	
unclassified	2	1	1	
5th WHO edition				
AML with defining genetic abnormalities	244	126	118	0.383
AML, defined by differentiation	97	45	52	0.383
ICC				
AML with t(8;21)	29	20	9	0.034
AML with inv(16)	18	14	4	0.016
AML with t(9;11)	7	2	5	0.25
AML with other <i>KMT2A</i> rearrangements	4	2	2	0.995
AML with t(6;9)	5	2	3	0.649
AML with inv(3)	3	1	2	0.56
AML with other rare translocations	3	2	1	0.567
AML with mutated <i>NPM1</i>	65	32	33	0.87
AML with in-frame bZIP <i>CEBPA</i> mutations	56	34	22	0.084
AML with mutated <i>TP53</i>	16	4	12	0.039
AML with myelodysplasia-related gene mutations	61	27	34	0.312
AML with myelodysplasia-related cytogenetic abnormalities	15	5	10	0.184
AML NOS	59	26	33	0.306

表 6 (continued)



	Total (n=341)	Low AS events (n=171)	High AS events (n=170)	P value
Cytogenetics (MRC AML10 risk category)				
favorable	47	34	13	0.001
intermediate	245	120	125	0.493
poor	43	12	31	0.002
no data (no mitosis)	6	5	1	
Induction chemotherapy				
I3A7	281	144	137	0.381
I2A5	45	23	22	0.89
Others	15	4	11	0.063
Hematopoietic stem cell transplantation	151	75	76	0.875
Induction response				
CR	257	138	119	0.022
PR/Refractory	63	27	36	0.201
Induction death	21	6	15	0.041
Relapse	121	62	59	0.458

n: number of patients

†median (range)

P values in red color indicate statistically significant ($P < 0.05$)

Source:

Yang YT, et al. Am J Hematol. 2023;98(5):784-793.

表 7：TCGA_AML 患者的臨床基本資料(clinical characteristics of the patients)

	Total (n=141)	Low AS events (n=70)	High AS events (n=71)	P value
sex				0.562
Male	76	36	40	
Female	65	34	31	
Age[†]	58 (18-88)	55 (18-81)	62 (21-88)	0.035
Lab data[†]				
WBC ($\times 10^3/\mu\text{L}$)	18 (1-224)	30 (1-203)	10 (1-224)	0.009
Hb (g/dL)	9 (6-14)	9 (6-13)	9 (6-14)	0.867
Platelett ($\times 10^3/\mu\text{L}$)	50 (8-351)	51 (9-351)	48 (8-232)	0.675
FAB				
M0	13	7	6	0.753
M1	37	13	24	0.04
M2	35	16	19	0.595
M4	33	21	12	0.067
M5	17	12	5	0.066
M6	2	1	1	0.992
M7	3	0	3	0.083
undclassified	1	0	1	
5th WHO edition				
AML with defining genetic abnormalities	110	55	55	0.875
AML, defined by differentiation	31	15	16	0.875
ICC				
AML with t(8;21)	7	4	3	0.687
AML with inv(16)	10	5	5	0.982
AML with t(9;11)	1	1	0	0.316
AML with t(9;22)	3	2	1	0.554
AML with mutated <i>NPM1</i>	37	20	17	0.536
AML with in-frame bZIP <i>CEBPA</i> mutations	9	6	3	0.295
AML with mutated <i>TP53</i>	11	2	9	0.03
AML with myelodysplasia-related gene mutations	15	7	8	0.809
AML with myelodysplasia-related cytogenetic abnormalities	17	8	9	0.822
AML NOS	31	15	16	0.875
Cytogenetics				
favorable	17	9	8	0.974
intermediate	90	47	43	0.413
poor	32	13	19	0.248
no data	2			

Source:

Yang YT et al. Am J Hematol. 2023;98(5):784-793.

表 8：比較 NTUH_AML cohort 不同的 global splicing pattern 間基因突變的差異

variables	number of tested	No. of patients with alteration			P value
		mutated patients	Low AS events (n=171)	High AS events (n=170)	
Nucleophosmin					
<i>NPM1</i>	341	66	33	33	0.979
Signaling and kinase pathway					
<i>FLT3-ITD</i>	341	77	42	35	0.398
<i>FLT3-TKD</i>	340	34	21	13	0.159
<i>KRAS</i>	341	16	8	8	0.99
<i>NRAS</i>	341	59	35	24	0.122
<i>KIT</i>	340	27	21	6	0.003
<i>PTPN11</i>	340	29	10	10	0.978
<i>JAK-2</i>	291	3	2	1	0.504
<i>CSF3R</i>	151	2	1	1	0.948
Epigenetic modifiers					
<i>DNMT3A</i>	340	68	32	36	0.552
<i>IDH1</i>	340	17	6	11	0.206
<i>IDH2</i>	340	42	18	24	0.305
<i>TET2</i>	340	41	23	18	0.43
<i>ASXL1</i>	340	35	15	20	0.354
<i>EZH2</i>	338	8	3	5	0.465
<i>PHF6</i>	337	9	4	5	0.73
<i>MLL</i>	340	13	7	6	0.778
<i>SETBP1</i>	274	4	0	4	0.052
Transcription factors					
<i>CEBPA_{mono-mutation}</i>	341	27	12	15	0.538
<i>CEBPA_{double-mutations}</i>	341	48	32	16	0.013
<i>CEBPA_{bZIP}</i>	341	59	36	23	0.067
<i>RUNX1</i>	339	33	17	16	0.869
<i>GATA2</i>	340	39	19	20	0.835
<i>WT1</i>	340	32	16	16	0.972
<i>BCOR</i>	338	8	3	5	0.465
Tumor suppressors					
<i>TP53</i>	340	16	4	12	0.038
Spliceosome complex					
<i>SF3B1</i>	336	6	3	3	1
<i>U2AF1</i>	336	11	9	2	0.03
<i>SRSF2</i>	337	11	7	4	0.364
<i>ZRSR2</i>	314	2	1	1	0.986
Cohesine complex					
<i>RAD21</i>	339	9	4	5	0.716
<i>STAG2</i>	339	10	5	5	0.977
<i>SMC3</i>	338	5	5	0	0.025
<i>SMC1A</i>	339	7	3	4	0.686

n: number of patients

P values in red color indicate statistically significant ($P < 0.05$)

Source: Yang YT et al. Am J Hematol. 2023;98(5):784-793.

表 9：比較 TCGA_AML cohort 不同的 global splicing pattern 間基因突變的差異

variables	number of tested	No. of patients with alteration			P value
		mutated patients	Low AS events (n=70)	High AS events (n=71)	
Nucleophosmin					
<i>NPM1</i>	141	38	21	17	0.421
Signaling and kinase pathway					
<i>FLT3-ITD</i>	141	23	9	14	0.273
<i>FLT3-TKD</i>	141	16	9	7	0.578
<i>KRAS</i>	141	6	4	2	0.398
<i>NRAS</i>	141	11	8	3	0.112
<i>KIT</i>	141	7	3	4	0.715
<i>PTPN11</i>	141	8	3	5	0.483
Epigenetic modifiers					
<i>DNMT3A</i>	141	35	15	20	0.603
<i>IDH1</i>	141	14	5	9	0.433
<i>IDH2</i>	141	16	8	8	0.976
<i>TET2</i>	141	16	6	10	0.305
<i>ASXL1</i>	141	3	0	3	0.083
<i>PHF6</i>	141	5	5	0	0.022
<i>EZH2</i>	141	2	0	2	0.16
Transcription factors					
<i>CEBPA</i>	141	12	7	5	0.533
<i>CEBPA_{bZIP}</i>	141	9	6	3	0.295
<i>RUNX1</i>	141	13	6	7	0.793
<i>WT1</i>	141	9	4	5	0.749
Tumor suppressors					
<i>TP53</i>	141	11	2	9	0.03
Spliceosome complex					
<i>U2AF1</i>	141	6	2	4	0.418
Cohesine complex					
<i>RAD21</i>	141	4	2	2	0.989
<i>STAG2</i>	141	5	3	2	0.64
<i>SMC3</i>	141	6	3	3	0.986
<i>SMC1A</i>	141	6	4	2	0.24

n: number of patients

P values in red color indicate statistically significant ($P < 0.05$)

Source: Yang YT et al. Am J Hematol. 2023;98(5):784-793.

表 10：利用單變數及多變數分析(univariate and multivariate analysis)
NTUH_AML 病患的預後因子，顯示轉錄體整體的選擇性剪接模式是一個獨立的預後因子



Variables	Overall Survival (univariate)				Overall Survival (multivariate)				Overall Survival (multivariate)			
	HR	95% CI		<i>P</i>	HR	95% CI		<i>P</i>	HR	95% CI		<i>P</i>
		Lower	Upper			Lower	Upper			Lower	Upper	
Age [†]	1.033	1.023	1.044	< 0.001*	1.029	1.019	1.039	< 0.001*	1.029	1.019	1.039	< 0.001*
WBC [§]	1.002	1	1.004	0.034*	1.004	1.002	1.006	< 0.001*	1.004	1.002	1.006	< 0.001*
2022 ELN risk [‡]	3.426	2.438	4.814	< 0.001*	2.943	2.082	4.162	< 0.001*	3.007	2.131	4.243	< 0.001*
AS event count [※]	1.127	1.042	1.218	0.003*					1.137	1.044	1.238	0.003*
AS event pattern [◎]	1.619	1.212	2.162	0.001*	1.692	1.242	2.305	0.001*				

Abbreviations: HR, hazards ratio; CI, confidence interval; AS, alternative splicing.

*Statistically significant ($P < 0.05$).

[†]Age as a continuous variable.

[§]WBC as a continuous variable.

[‡]Adverse and intermediate risk relative to favorable risk (the reference).

[※]AS event count as a continuous variable.

[◎]High global AS pattern relative to low global AS pattern (the reference).

表 11：分析 NTUH_AML 病患整個轉錄體選擇性剪接對細胞生理路徑的影響



Term-pathways	Count	P value	FDR	Fold enrichment
Bromodomain	16	1.37E-04	6.31E-04	2.93
regulation of small GTPase mediated signal transduction	38	1.02E-07	3.32E-04	2.51
cytosolic ribosome	24	5.79E-05	2.67E-03	2.44
Rho GTPase activation protein	30	1.73E-05	1.06E-02	2.30
SH2 domain	29	4.69E-04	1.80E-03	1.97
GTPase activation	54	4.05E-06	1.17E-04	1.88
Guanine-nucleotide releasing factor	43	1.57E-04	1.95E-03	1.79
SH3 domain	54	3.54E-05	2.71E-04	1.76
Pleckstrin homology-like domain	103	7.17E-08	7.32E-05	1.69
protein phosphorylation	110	3.04E-08	1.97E-04	1.69
Glycyl lysine isopeptide (Lys-Gly) (interchain with G-Cter in SUMO2)	261	1.47E-14	3.95E-11	1.58
ATP binding	305	1.61E-12	1.01E-09	1.46
Proton acceptor	160	2.15E-06	1.58E-03	1.43
Cell cycle	139	2.57E-05	2.04E-03	1.39
Kinase	150	1.62E-05	3.24E-04	1.38
Isopeptide bond	361	1.91E-09	1.48E-08	1.32
Ubl conjugation	512	1.79E-13	1.85E-12	1.32
Zinc-finger	303	8.54E-05	4.91E-04	1.22
metal ion binding	434	4.06E-06	5.45E-04	1.21
Nucleotide-binding	342	4.50E-05	1.01E-03	1.19

Source: Yang YT et al. Am J Hematol. 2023;98(5):784-793.

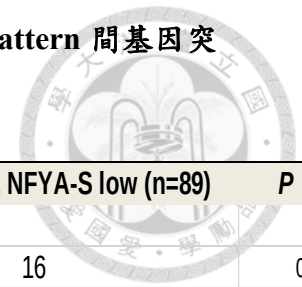
表 12：比較 NTUH_AML cohort 不同的 *NFYA* 基因 splicing pattern 間臨床表徵的差異

	NFYA-L low, NFYA-S high (n=87)	NFYA-L high, NFYA-S low (n=89)	P value
sex			0.368
Male	40	47	
Female	47	42	
Age[†]	43 (19-77)	49 (18-84)	0.193
Lab data[†]			
WBC (×10 ³ /μL)	13.92 (0.49-405.65)	56.32 (0.51-361.30)	0.002*
Hb (g/dL)	8.3 (4.4-16)	7.9 (2.4-13.2)	0.2
Platelet (×10 ³ /μL)	43 (3-751)	49 (3-712)	0.584
LDH (U/L)	545 (145-7400)	954 (201-8693)	0.007*
FAB			
M0	3	3	0.978
M1	23	19	0.431
M2	38	18	0.001*
M4	16	33	0.005*
M5	1	14	< 0.001*
M6	5	2	0.237
M7	0	0	
unclassified	1	0	
5th WHO edition			
AML with defining genetic abnormalities	63	63	0.812
AML, defined by differentiation	24	26	0.812
ICC			
AML with t(8;21)	6	3	0.291
AML with inv(16)	5	5	0.971
AML with t(9;11)	0	3	0.085
AML with other <i>KMT2A</i> rearrangements	0	2	0.161
AML with t(6;9)	2	1	0.55
AML with inv(3)	2	2	0.982
AML with other rare translocations	0	2	0.161
AML with mutated <i>NPM1</i>	15	15	0.946
AML with in-frame bZIP <i>CEBPA</i> mutations	22	12	0.048*
AML with mutated <i>TP53</i>	1	4	0.184
AML with myelodysplasia-related gene mutations	13	20	0.203
AML with myelodysplasia-related cytogenetic abnormalities	3	6	0.324
AML NOS	18	14	0.397
Cytogenetics (MRC AML10 risk category)			
favorable	11	8	0.438
intermediate	65	61	0.367
poor	10	17	0.163
no data (no mitosis)	1	3	
Induction response			
CR	71	60	0.031*
PR/Refractory	12	24	0.03*
Induction death	4	5	0.76
Relapse	30	37	0.122

n: number of patients, †median (range), *Statistically significant ($P < 0.05$)

Note: We used the median gene expression levels of *NFYA* isoforms as cutoff points to classify patients into *NFYA*-L or *NFYA*-S low- and high-expression groups.

表 13：比較 NTUH_AML cohort 不同的 *NFYA* 基因 splicing pattern 間基因突變的差異



Mutations	NFYA-L low, NFYA-S high (n=87)	NFYA-L high, NFYA-S low (n=89)	P value
Nucleophosmin			
<i>NPM1</i>	15	16	0.899
Signaling and kinase pathway			
<i>FLT3-ITD</i>	18	16	0.677
<i>FLT3-TKD</i>	4	15	0.008
<i>KRAS</i>	4	8	0.25
<i>NRAS</i>	11	22	0.04
<i>KIT</i>	8	4	0.226
<i>PTPN11</i>	3	4	0.713
<i>JAK-2</i>	1	0	0.325
Epigenetic modifiers			
<i>DNMT3A</i>	13	18	0.342
<i>IDH1</i>	5	6	0.772
<i>IDH2</i>	10	9	0.789
<i>TET2</i>	9	14	0.279
<i>ASXL1</i>	8	8	0.981
<i>EZH2</i>	4	1	0.167
<i>PHF6</i>	2	4	0.425
<i>MLL</i>	3	4	0.724
Transcription factors			
<i>CEBPA_{mono-mutation}</i>	5	7	0.58
<i>CEBPA_{double-mutations}</i>	20	10	0.038
<i>CEBPA_{bZIP}</i>	23	13	0.052
<i>RUNX1</i>	8	10	0.621
<i>GATA2</i>	17	4	0.002
<i>WT1</i>	10	7	0.432
<i>BCOR</i>	1	3	0.326
Tumor suppressors			
<i>TP53</i>	2	4	0.417
Spliceosome complex			
<i>SF3B1</i>	0	3	0.081
<i>U2AF1</i>	1	7	0.029
<i>SRSF2</i>	3	2	0.652
Cohesine complex			
<i>RAD21</i>	3	2	0.634
<i>STAG2</i>	3	5	0.493

n: number of patients, *P* value in red color indicate statistically significant ($P < 0.05$)

表 14：比較 NTUH_AML 病患，不同的 *NFYA-S* 表達量對誘導性化學治療(induction chemotherapy)的影響



	Total (n=341)	<i>NFYA-S</i> low expression (n=171)	<i>NFYA-S</i> high expression (n=170)	<i>P</i> value
Induction response				
CR	257	118	139	0.006*
PR/Refractory	63	44	19	< 0.001*
Induction death	21	9	12	0.492
Relapse	121	64	57	0.161

Abbreviations: CR, complete remission; PR, partial remission

*Statistically significant ($P < 0.05$)

表 15：利用單變數及多變數分析(univariate and multivariate analysis) NTUH_AML 病患的預後因子，顯示 *NFYA* 轉錄異構體表達型式(splicing pattern)是一個獨立的預後因子



Variables	Overall Survival (univariate)	Overall Survival (multivariate)			
			95% CI		
	<i>P</i>	HR	Lower	Upper	<i>P</i>
Age [†]	< 0.001*	1.029	1.019	1.039	< 0.001*
2022 ELN risk [‡]	< 0.001*	2.931	2.081	4.129	< 0.001*
<i>NFYA</i> splicing pattern[§]					
<i>NFYA</i> -L low / <i>NFYA</i> -S high (<i>NFYA</i>-S predominance)	Reference	Reference	Reference	Reference	Reference
<i>NFYA</i> -L high / <i>NFYA</i> -S high	0.299	1.225	0.794	1.89	0.359
<i>NFYA</i> -L low / <i>NFYA</i> -S low	0.044*	1.514	0.993	2.307	0.054
<i>NFYA</i> -L high / <i>NFYA</i> -S low (<i>NFYA</i>-L predominance)	0.009*	1.524	1.013	2.292	0.043*

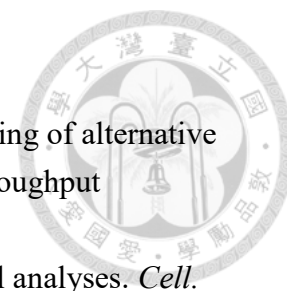
*Statistically significant ($P < 0.05$)

[†] Age as a continuous variable

[‡] 2022 ELN adverse and intermediate risk relative to favorable risk (the reference)

[§] The *NFYA* splicing pattern was divided into four groups: *NFYA*-L low/*NFYA*-S high, *NFYA*-L high/*NFYA*-S high, *NFYA*-L low/*NFYA*-S low, and *NFYA*-L high/*NFYA*-S low.

參考文獻(References)

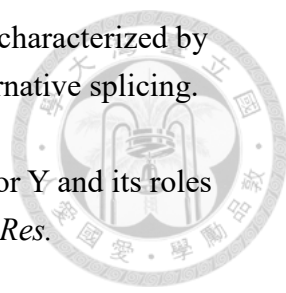


1. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet.* 2008;40(12):1413-1415.
2. Blencowe BJ. Alternative splicing: new insights from global analyses. *Cell.* 2006;126(1):37-47.
3. Srebrow A, Kornblihtt AR. The connection between splicing and cancer. *J Cell Sci.* 2006;119(Pt 13):2635-2641.
4. Malcovati L, Hellstrom-Lindberg E, Bowen D, et al. Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukemiaNet. *Blood.* 2013;122(17):2943-2964.
5. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med.* 2016;374(23):2209-2221.
6. Hou HA, Tien HF. Genomic landscape in acute myeloid leukemia and its implications in risk classification and targeted therapies. *J Biomed Sci.* 2020;27(1):81.
7. Tsai CH, Yao CY, Tien FM, et al. Incorporation of long non-coding RNA expression profile in the 2017 ELN risk classification can improve prognostic prediction of acute myeloid leukemia patients. *EBioMedicine.* 2019;40:240-250.
8. Huang HH, Chen FY, Chou WC, et al. Long non-coding RNA HOXB-AS3 promotes myeloid cell proliferation and its higher expression is an adverse prognostic marker in patients with acute myeloid leukemia and myelodysplastic syndrome. *BMC Cancer.* 2019;19(1):617.
9. Li G, Gao Y, Li K, Lin A, Jiang Z. Genomic analysis of biomarkers related to the prognosis of acute myeloid leukemia. *Oncol Lett.* 2020;20(2):1824-1834.
10. Haferlach T, Schmidts I. The power and potential of integrated diagnostics in acute myeloid leukaemia. *Br J Haematol.* 2020;188(1):36-48.
11. Osterroos A, Bjorklund M, Eriksson A, et al. Integrated transcriptomic and genomic analysis improves prediction of complete remission and survival in elderly patients with acute myeloid leukemia. *Blood Cancer J.* 2020;10(6):67.
12. Buhagiar A, Borg J, Ayers D. Overview of current microRNA biomarker signatures as potential diagnostic tools for leukaemic conditions. *Noncoding RNA Res.* 2020;5(1):22-26.
13. Wang YH, Lin CC, Hsu CL, et al. Distinct clinical and biological characteristics of acute myeloid leukemia with higher expression of long

- noncoding RNA KIAA0125. *Ann Hematol.* 2021;100(2):487-498.
14. Chen Z, Song J, Wang W, et al. A novel 4-mRNA signature predicts the overall survival in acute myeloid leukemia. *Am J Hematol.* 2021;96(11):1385-1395.
15. Oltean S, Bates DO. Hallmarks of alternative splicing in cancer. *Oncogene.* 2014;33(46):5311-5318.
16. Adamia S, Haibe-Kains B, Pilarski PM, et al. A genome-wide aberrant RNA splicing in patients with acute myeloid leukemia identifies novel potential disease markers and therapeutic targets. *Clin Cancer Res.* 2014;20(5):1135-1145.
17. Zhou J, Chng WJ. Aberrant RNA splicing and mutations in spliceosome complex in acute myeloid leukemia. *Stem Cell Investig.* 2017;4:6.
18. Jin P, Tan Y, Zhang W, Li J, Wang K. Prognostic alternative mRNA splicing signatures and associated splicing factors in acute myeloid leukemia. *Neoplasia.* 2020;22(9):447-457.
19. Xie ZC, Gao L, Chen G, et al. Prognostic alternative splicing regulatory network of splicing events in acute myeloid leukemia patients based on SpliceSeq data from 136 cases. *Neoplasma.* 2020;67(3):623-635.
20. Chen XX, Zhu JH, Li ZP, Xiao HT, Zhou H. Comprehensive Characterization of the Prognosis Value of Alternative Splicing Events in Acute Myeloid Leukemia. *DNA Cell Biol.* 2020;39(7):1243-1255.
21. Anande G, Deshpande NP, Mareschal S, et al. RNA Splicing Alterations Induce a Cellular Stress Response Associated with Poor Prognosis in Acute Myeloid Leukemia. *Clin Cancer Res.* 2020;26(14):3597-3607.
22. Zhang B, Yang L, Wang X, Fu D. Identification of survival-related alternative splicing signatures in acute myeloid leukemia. *Biosci Rep.* 2021;41(7).
23. Hershberger CE, Moyer DC, Adema V, et al. Complex landscape of alternative splicing in myeloid neoplasms. *Leukemia.* 2021;35(4):1108-1120.
24. El Marabti E, Abdel-Wahab O. Therapeutic Modulation of RNA Splicing in Malignant and Non-Malignant Disease. *Trends Mol Med.* 2021;27(7):643-659.
25. Adams MD, Kelley JM, Gocayne JD, et al. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science.* 1991;252(5013):1651-1656.
26. Lee FCY, Ule J. Advances in CLIP Technologies for Studies of Protein-RNA Interactions. *Mol Cell.* 2018;69(3):354-369.
27. Thorsen K, Sorensen KD, Brems-Eskildsen AS, et al. Alternative splicing in colon, bladder, and prostate cancer identified by exon array analysis. *Mol Cell Proteomics.* 2008;7(7):1214-1224.
28. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for

- transcriptomics. *Nat Rev Genet.* 2009;10(1):57-63.
29. Nazarov PV, Muller A, Kaoma T, et al. RNA sequencing and transcriptome arrays analyses show opposing results for alternative splicing in patient derived samples. *BMC Genomics.* 2017;18(1):443.
 30. Romero JP, Ortiz-Estevéz M, Muniategui A, et al. Comparison of RNA-seq and microarray platforms for splice event detection using a cross-platform algorithm. *BMC Genomics.* 2018;19(1):703.
 31. Collins CA, Guthrie C. The question remains: is the spliceosome a ribozyme? *Nat Struct Biol.* 2000;7(10):850-854.
 32. Kim HK, Pham MHC, Ko KS, Rhee BD, Han J. Alternative splicing isoforms in health and disease. *Pflugers Arch.* 2018;470(7):995-1016.
 33. Nikom D, Zheng S. Alternative splicing in neurodegenerative disease and the promise of RNA therapies. *Nat Rev Neurosci.* 2023;24(8):457-473.
 34. Sveen A, Kilpinen S, Ruusulehto A, Lothe RA, Skotheim RI. Aberrant RNA splicing in cancer; expression changes and driver mutations of splicing factor genes. *Oncogene.* 2016;35(19):2413-2427.
 35. Zhang Y, Qian J, Gu C, Yang Y. Alternative splicing and cancer: a systematic review. *Signal Transduct Target Ther.* 2021;6(1):78.
 36. Saez B, Walter MJ, Graubert TA. Splicing factor gene mutations in hematologic malignancies. *Blood.* 2017;129(10):1260-1269.
 37. Hou HA, Liu CY, Kuo YY, et al. Splicing factor mutations predict poor prognosis in patients with de novo acute myeloid leukemia. *Oncotarget.* 2016;7(8):9084-9101.
 38. Kelly LM, Gilliland DG. Genetics of myeloid leukemias. *Annu Rev Genomics Hum Genet.* 2002;3:179-198.
 39. Shih AH, Abdel-Wahab O, Patel JP, Levine RL. The role of mutations in epigenetic regulators in myeloid malignancies. *Nat Rev Cancer.* 2012;12(9):599-612.
 40. Yoshimi A, Lin KT, Wiseman DH, et al. Coordinated alterations in RNA splicing and epigenetic regulation drive leukaemogenesis. *Nature.* 2019;574(7777):273-277.
 41. Berta D, Girma M, Melku M, Adane T, Birke B, Yalew A. Role of RNA Splicing Mutations in Diffuse Large B Cell Lymphoma. *Int J Gen Med.* 2023;16:2469-2480.
 42. Ritchie MD, Holzinger ER, Li R, Pendergrass SA, Kim D. Methods of integrating data to uncover genotype-phenotype interactions. *Nat Rev Genet.* 2015;16(2):85-97.
 43. Bejar R, Steensma DP. Recent developments in myelodysplastic syndromes.

- Blood*. 2014;124(18):2793-2803.
44. Sveen A, Agesen TH, Nesbakken A, Rognum TO, Lothe RA, Skotheim RI. Transcriptome instability in colorectal cancer identified by exon microarray analyses: Associations with splicing factor expression levels and patient survival. *Genome Med*. 2011;3(5):32.
 45. Wang K, Wu D, Zhang H, et al. Comprehensive map of age-associated splicing changes across human tissues and their contributions to age-associated diseases. *Sci Rep*. 2018;8(1):10929.
 46. Wang C, Chen L, Chen Y, et al. Abnormal global alternative RNA splicing in COVID-19 patients. *PLoS Genet*. 2022;18(4):e1010137.
 47. Sun Y, Xiao H. Identification of alternative splicing events by RNA sequencing in early growth tomato fruits. *BMC Genomics*. 2015;16:948.
 48. Clark TA, Sugnet CW, Ares M, Jr. Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. *Science*. 2002;296(5569):907-910.
 49. Srinivasan K, Shiue L, Hayes JD, et al. Detection and measurement of alternative splicing using splicing-sensitive microarrays. *Methods*. 2005;37(4):345-359.
 50. Ryan MC, Cleland J, Kim R, Wong WC, Weinstein JN. SpliceSeq: a resource for analysis and visualization of RNA-Seq data on alternative splicing and its functional impacts. *Bioinformatics*. 2012;28(18):2385-2387.
 51. Ryan M, Wong WC, Brown R, et al. TCGASpliceSeq a compendium of alternative mRNA splicing in cancer. *Nucleic Acids Res*. 2016;44(D1):D1018-1022.
 52. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545-15550.
 53. Farge T, Saland E, de Toni F, et al. Chemotherapy-Resistant Human Acute Myeloid Leukemia Cells Are Not Enriched for Leukemic Stem Cells but Require Oxidative Metabolism. *Cancer Discov*. 2017;7(7):716-735.
 54. Yang YT, Chiu YC, Kao CJ, et al. The prognostic significance of global aberrant alternative splicing in patients with myelodysplastic syndrome. *Blood Cancer J*. 2018;8(8):78.
 55. Yang YT, Yao CY, Chiu PJ, et al. Evaluation of the clinical significance of global mRNA alternative splicing in patients with acute myeloid leukemia. *Am J Hematol*. 2023;98(5):784-793.
 56. Shiozawa Y, Malcovati L, Galli A, et al. Gene expression and risk of leukemic transformation in myelodysplasia. *Blood*. 2017;130(24):2642-2653.

- 
57. Harries LW, Hernandez D, Henley W, et al. Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. *Aging Cell*. 2011;10(5):868-878.
 58. Ly LL, Yoshida H, Yamaguchi M. Nuclear transcription factor Y and its roles in cellular processes related to human disease. *Am J Cancer Res*. 2013;3(4):339-346.
 59. Li G, Zhao H, Wang L, Wang Y, Guo X, Xu B. The animal nuclear factor Y: an enigmatic and important heterotrimeric transcription factor. *Am J Cancer Res*. 2018;8(7):1106-1125.
 60. Maity SN. NF-Y (CBF) regulation in specific cell types and mouse models. *Biochim Biophys Acta Gene Regul Mech*. 2017;1860(5):598-603.
 61. Ly LL, Suyari O, Yoshioka Y, Tue NT, Yoshida H, Yamaguchi M. dNF-YB plays dual roles in cell death and cell differentiation during *Drosophila* eye development. *Gene*. 2013;520(2):106-118.
 62. Gurtner A, Manni I, Piaggio G. NF-Y in cancer: Impact on cell transformation of a gene essential for proliferation. *Biochim Biophys Acta Gene Regul Mech*. 2017;1860(5):604-616.
 63. Bezzecchi E, Ronzio M, Dolfini D, Mantovani R. NF-YA Overexpression in Lung Cancer: LUSC. *Genes (Basel)*. 2019;10(11).
 64. Cappabianca L, Farina AR, Di Marcotullio L, et al. Discovery, characterization and potential roles of a novel NF-YAx splice variant in human neuroblastoma. *J Exp Clin Cancer Res*. 2019;38(1):482.
 65. Dolfini D, Andrioletti V, Mantovani R. Overexpression and alternative splicing of NF-YA in breast cancer. *Sci Rep*. 2019;9(1):12955.
 66. Bezzecchi E, Ronzio M, Semeghini V, Andrioletti V, Mantovani R, Dolfini D. NF-YA Overexpression in Lung Cancer: LUAD. *Genes (Basel)*. 2020;11(2).
 67. Li Y, Xiao X, Chen H, Chen Z, Hu K, Yin D. Transcription factor NFYA promotes G1/S cell cycle transition and cell proliferation by transactivating cyclin D1 and CDK4 in clear cell renal cell carcinoma. *Am J Cancer Res*. 2020;10(8):2446-2463.
 68. Yang WT, Feng Q, Ma HM, Lei D, Zheng PS. NF-YA promotes the cell proliferation and tumorigenic properties by transcriptional activation of SOX2 in cervical cancer. *J Cell Mol Med*. 2020;24(21):12464-12475.
 69. Bezzecchi E, Ronzio M, Mantovani R, Dolfini D. NF-Y Overexpression in Liver Hepatocellular Carcinoma (HCC). *Int J Mol Sci*. 2020;21(23).
 70. Dolfini D, Minuzzo M, Pavesi G, Mantovani R. The short isoform of NF-YA belongs to the embryonic stem cell transcription factor circuitry. *Stem Cells*. 2012;30(11):2450-2459.

- 
71. Zhu J, Zhang Y, Joe GJ, Pompetti R, Emerson SG. NF-Ya activates multiple hematopoietic stem cell (HSC) regulatory genes and promotes HSC self-renewal. *Proc Natl Acad Sci U S A*. 2005;102(33):11728-11733.
72. Bungartz G, Land H, Scadden DT, Emerson SG. NF-Y is necessary for hematopoietic stem cell proliferation and survival. *Blood*. 2012;119(6):1380-1389.
73. Dolfini D, Gatta R, Mantovani R. NF-Y and the transcriptional activation of CCAAT promoters. *Crit Rev Biochem Mol Biol*. 2012;47(1):29-49.
74. Basile V, Baruffaldi F, Dolfini D, et al. NF-YA splice variants have different roles on muscle differentiation. *Biochim Biophys Acta*. 2016;1859(4):627-638.
75. Libetti D, Bernardini A, Sertic S, Messina G, Dolfini D, Mantovani R. The Switch from NF-YA1 to NF-YAs Isoform Impairs Myotubes Formation. *Cells*. 2020;9(3).
76. O'Reilly E, Zeinabad HA, Szegezdi E. Hematopoietic versus leukemic stem cell quiescence: Challenges and therapeutic opportunities. *Blood Rev*. 2021;50:100850.
77. Esposito MT, Hagstrom-Andersson A, Stam RW, Bortoluzzi S. Editorial: Harnessing chemotherapy resistance and development of novel therapeutic strategies for acute leukemia with KMT2A (MLL)-gene rearrangements. *Front Pharmacol*. 2022;13:977741.
78. Uckun FM, Qazi S. Tyrosine kinases in KMT2A/MLL-rearranged acute leukemias as potential therapeutic targets to overcome cancer drug resistance. *Cancer Drug Resist*. 2022;5(4):902-916.
79. Fajardo-Orduna GR, Ledesma-Martinez E, Aguiniga-Sanchez I, Mora-Garcia ML, Weiss-Steider B, Santiago-Osorio E. Inhibitors of Chemoresistance Pathways in Combination with Ara-C to Overcome Multidrug Resistance in AML. A Mini Review. *Int J Mol Sci*. 2021;22(9).
80. Steensma DP, Wermke M, Klimek VM, et al. Phase I First-in-Human Dose Escalation Study of the oral SF3B1 modulator H3B-8800 in myeloid neoplasms. *Leukemia*. 2021;35(12):3542-3550.

附錄 個人在修業期間所發表之論文清冊

與本論文相關之第一作者論文

1. **Yang YT**, Chiu YC, Kao CJ, Hou HA, Lin CC, Tsai CH, Tseng MH, Chou WC, Tien HF. The prognostic significance of global aberrant alternative splicing in patients with myelodysplastic syndrome. *Blood Cancer Journal*. 2018 Aug 13;8(8):78. (5 year IF: 6.7)
2. **Yang YT**, Yao CY, Chiu PJ, Kao CJ, Hou HA, Lin CC, Chou WC, Tien HF. Evaluation of the clinical significance of global mRNA alternative splicing in patients with acute myeloid leukemia. *Am J Hematol*. 2023 May;98(5):784-793. (5 year IF: 10)
3. **Yang YT**, Yao CY, Kao CJ, Chiu PJ, Lin ME, Hou HA, Lin CC, Chou WC, Tien HF. Clinical relevance of NFYA splice variants in patients with acute myeloid leukemia undergoing intensive chemotherapy. (manuscript submitted)

其他修業期間所發表之論文

4. Lee WH, Lin CC, Tsai CH, Tien FM, Lo MY, Tseng MH, Kuo YY, Yu SC, Liu MC, Yuan CT, Yang YT, Chuang MK, Ko BS, Tang JL, Sun HI, Chuang YK, Tien HF, Chou WC. Comparison of the 2022 World Health Organization Classification and International Consensus Classification in Myelodysplastic Syndromes/Neoplasms. *Blood Cancer J*. 2024 Apr 9;14(1):57. (本人為共同作者)
5. Lee WH, Tsai MT, Tsai CH, Tien FM, Lo MY, Tseng MH, Kuo YY, Liu MC, Yang YT, Chen JC, Tang JL, Sun HI, Chuang YK, Lin LI, Chou WC, Lin CC, Hou HA, Tien HF. Validation of the molecular international prognostic scoring system in patients with myelodysplastic syndromes defined by international consensus classification. *Blood Cancer J*. 2023 Aug 9;13(1):120. (本人為共同作者)
6. Wang YH, Yu SC, Ko BS, Yang YT, Yao M, Tang JL, Huang TC. Correlative analysis of overall survival with clinical characteristics in 127 patients with mantle cell lymphoma: a multi-institutional cohort in Taiwan. *Int J Hematol*. 2020 Sep;112(3):385-394. (本人為共同作者)