出國報告(出國類別：進修)

血液腫瘤疾病之相關易感基因研究

服務機關：國立臺灣大學醫學院附設醫院新竹臺大分院

姓名職稱：李孟如主治醫師

派赴國家：美國

出國期間：111 年 9 月 1 日至 112 年 8 月 31 日

報告日期：112 年 9 月 5 日

摘要

　　研究是醫學進步之鑰。提升治療成績來自精確診斷、了解致病機轉、研發製藥和提供精準醫療服務。臨床照護和基礎醫學缺一不可且相輔相成。骨髓衰竭症候群（bone marrow failure syndrome）雖不屬癌症，然高異質性疾病種類和多樣性臨床表現特性使得診斷和照護益發困難。本次前往波士頓兒童醫院 Prof. Akiko Shimamura 實驗室進修，除了熟悉基因檢測判讀流程、操作細胞培養及增值實驗、功能驗證之實驗設計外，對於建立團隊所需人力及人員配置亦較有概念。接續將建立骨髓衰竭症候群研究團隊，建立跨院區跨體系之合作模式，期許提供台灣骨髓衰竭症候群病患及家庭與世界接軌之照護。

目次

本文

目的

探討骨髓衰竭症候群生殖細胞系基因體學（germline genomics），接軌臨床表現、強化診斷依據，結合造血系統病理機轉，提供未來藥物研發參酌方向。

過程

一、人員配置

波士頓兒童醫院從事小兒血液疾病研究的相關人員，從計畫主持人至助理，共約 280 人。骨髓衰竭症候群團隊定期舉辦個案討論會，與會人員包括主要照護該個案之臨床團隊、血液病理團隊、移植團隊、專科護理師及個管師，共計 20-30 人，針對診斷或處置進行專家討論及溝通。基因分析團隊定期舉辦討論，針對個案個別的候選基因變異點討論，亦同時涵蓋整個資料庫的分析。

實驗室成員除計畫主持人外，包括 1 位實驗室管理組長指導協助實驗設計及操作，1 名博士後研究員，2 名技術員及 1-3 位外國進修醫師。實驗室有許多合作研究團隊，包括跨部門與血液病理 Mark Fleming 團隊、幹細胞中心 Vijay Sankaran 團隊等及跟其他機構，如德國 Miriam Erlacher 團隊、以色列 Hannah Tamary 團隊合作。

二、變異點位分析流程

分析代表生殖細胞系基因之檢體，包括皮膚纖維母細胞（fibroblast）、骨髓纖維母細胞或周邊血液，進行全基因體序列（whole genome sequence）分析，與親源比對變異點，分析流程如圖一。針對基因編碼區（coding regions）及基因剪切位點（splicing site）進行變異點分析，首先根據人口資料庫 gnomAD 資料庫提供之等位基因頻率（AF, allele frequency）資訊篩選，以同基因合子（hom, homozygous）或複合異基因合子（CHZ, compound heterozygous）為例，等位基因頻率切點訂為

0.1。而針對異型合子變異（het, heterozygous），等位基因頻率切點訂為 0.01。篩選出的變異點需確認原始序列是可信的，如圖二所示，需確認變異點確實存在於兩個方向序列，且基因序列（alignment）無錯置（mismapping）。針對變異點個別變異點特性，如是否位於核甘酸或胺基酸之高度保存區（conserved region）、變異造成的效應，如點突變之胺基酸改變程度、是否形成不完整之蛋白質序列（truncated protein），同時使用電腦預測軟體（SIFT、Polyphen-2 等）預測變異點對蛋白質所造成的後果，根據美國醫學遺傳暨基因組學會（American College of Medical Genetics and Genomics; ACMG）與美國分子病理學會（Association for Molecular Pathology）共同制定的標準綜合判斷其致病性。同時針對有興趣之基因，根據基因功能及文獻，歸納出候選基因變異。

考量僅由單一個案或單一家族得出某個未報告過的基因為致病基因略嫌武斷，建議進一步搜尋世代研究個案資料庫，視是否有其他個案亦有同基因的變異點。倘若存在幾位臨床表現雷同的個案均同時有某特定候選基因，針對個別個案的變異點，分析其在不同物種間（orthologs）及同物種不同蛋白異構體間（paralog）保守序列（conserved sequences）程度，如圖三。個別變異點是否位於蛋白質功能區塊（protein domain）關係著此變異點是否可能與關鍵功能或三級結構形成相關，如圖四所示。亦可以 Pymol 或是 AlphaFold 軟體分析其蛋白質立體結構預測，如圖五所示。

三、實驗驗證設計

1. 以 Sanger 法驗證變異點確實存在。

2. 檢測 mRNA 及蛋白質表現量了解在轉錄（transcript）或轉譯階段受影響。

3. 以動物模型或細胞株，以基因剔除法（knockout, KO）、或常間回文重複序列叢集關聯蛋白替換部分基因序列（Clustered Regularly Interspaced Short Palindromic Repeat/ CRISPR associated protein 9, CRISPR/Cas9）或以細胞轉染

（Transfection）法使變異基因過度表現（overexpression），了解功能性影響為功能喪失型（loss-of-function）或顯性抑制效應（Dominant negative）。隨著細胞分化，了解該基因於造血系統之角色。

```
                    ┌──────────────────────────┐
                    │  coding / splicing variants │
                    └──────────────────────────┘
                      ╱                        ╲
        ┌─────────────────────┐      ┌─────────────────────┐
        │  AF <0.01           │      │  AF <0.001          │
        │  Hom or CHZ         │      │  Het                │
        └─────────────────────┘      └─────────────────────┘
```

Adequate read depth, good quality

Variants	**Gene**
Pathogenicity Effects	**Disease, pathway, inheritance**
• **Functional effect:** frameshift, missense, nonsense, splicing	• **GeneCards**
• **Conservation** of nucleotide & amino acid	• **OMIM**
• **Changed amino acid** degree	• **HPO**
• **Bioinformatic tools**: Mutation Taster/ Polyphen2/ SIFT/ CADD/ Splicing Predictions	• **Pubmed**

Interpretations

Candidate Variants fit potential pathogenicity and gene character

圖一、全基因體序列（whole genome sequence）分析流程。基因變異點分析針對基因編碼區（coding regions）及基因剪切位點（splicing site），根據人口資料庫 gnomAD 資料庫提供之等位基因頻率（AF, allele frequency）資訊，篩選出原始序列可信的變異點，經由一系列變異點和基因分析，挑選出具潛在致病性的候選基因。縮寫：HPO, Human Phenotype Ontology; OMIM, Online Mendelian Inheritance in Man; Hom, homozygous; CHZ, compound heterozygous; Het, heterozygous; AF, gnomAD allele frequency; AD, autosomal dominant; AR, autosomal recessive
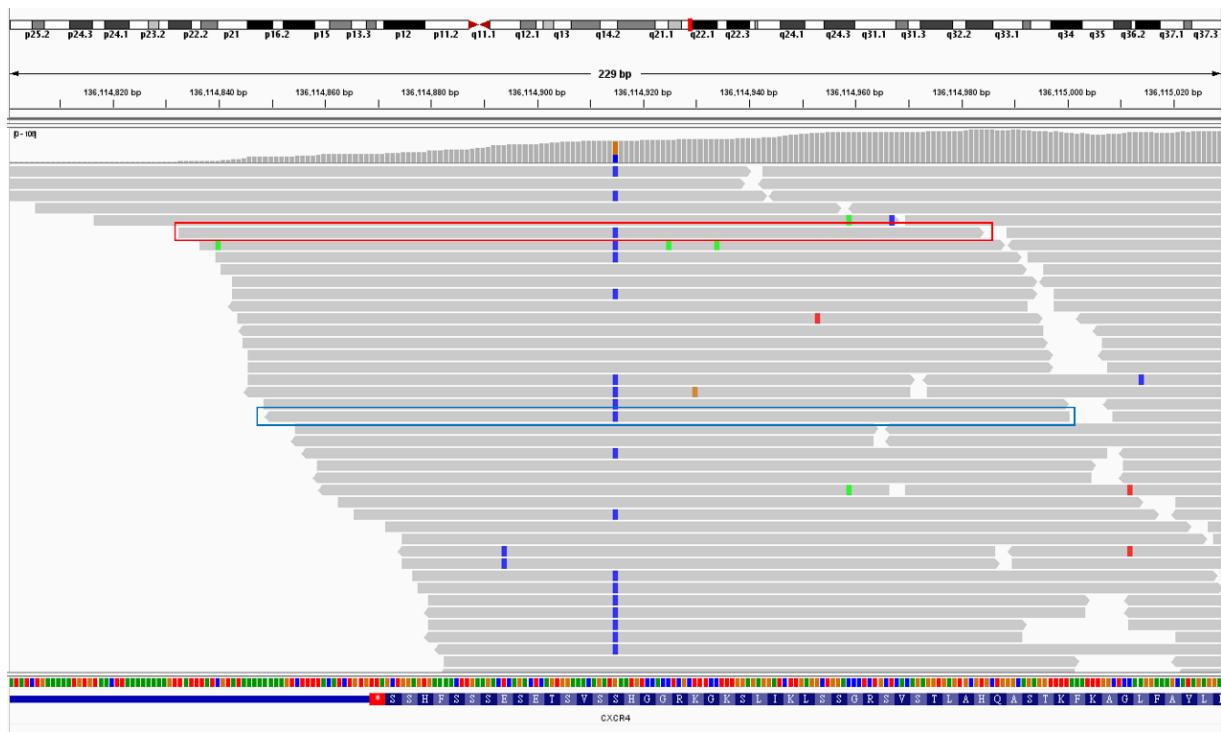
圖二、原始序列檢測。藉由基因檢視軟體 Integrative Genomics Viewer （IGV）確認變異點並存於兩個方向序列，紅色及藍色框分別代表兩個方向的序列。除了目標變異點外，序列裡無其他變異點代表基因序列對應（alignment）良好且無錯置（mismapping）。

Between isoforms

```
TUBB    MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLDRI
TUBB1   MREIV H IQIGQCGNQIGAKFWEMIGEEHGIDLAGSDRGASALQLERI
TUBB2A  MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGSYHGDSDLQLERI
TUBB2B  MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGSYHGDSDLQLERI
TUBB3   MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPSGNYVGDSDLQLERI
TUBB4A  MREIV H LQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLERI
TUBB4B  MREIV H LQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLERI
TUBB6   MREIV H IQAGQCGNQIGTKFWEVISDEHGIDPAGGYVGDSALQLERI
TUBB8   MREIV L TQIGQCGNQIGAKFWEVISDEHAIDSAGTYHGDSHLQLERI
```

Between species

```
Human       MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLDRI
Mouse       MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLDRI
Rat         MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLDRI
Macaca      MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLDRI
Chimpanzee  MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLDRI
Bovine      MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLDRI
Pig         MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLDRI
Hamster     MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLDRI
Chicken     MREIV H IQAGQCGNQIGTKFWEVISDEHGIDPAGGYVGDSALQLERI
Salmon      MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLDRI
Frog        MREIV H IQAGQCGNQIGAKFWEVISDEHGIDPTGTYHGDSDLQLDRI
```

圖三、基因變異點保守序列（**conserved sequences**）程度分析。比較在同物種不同蛋白異構體間（between isoforms, paralog）和不同物種間（between species, orthologs）基因序列保守程度，此例顯示之目標變異點在不同蛋白異構體間和不同物種間均具有高度保留狀況。

| Variant ID	HGVS	AA change	Exon	Domain
9-133440478-G-A	c.1921G>A	p.Glu641Lys	16	S domain
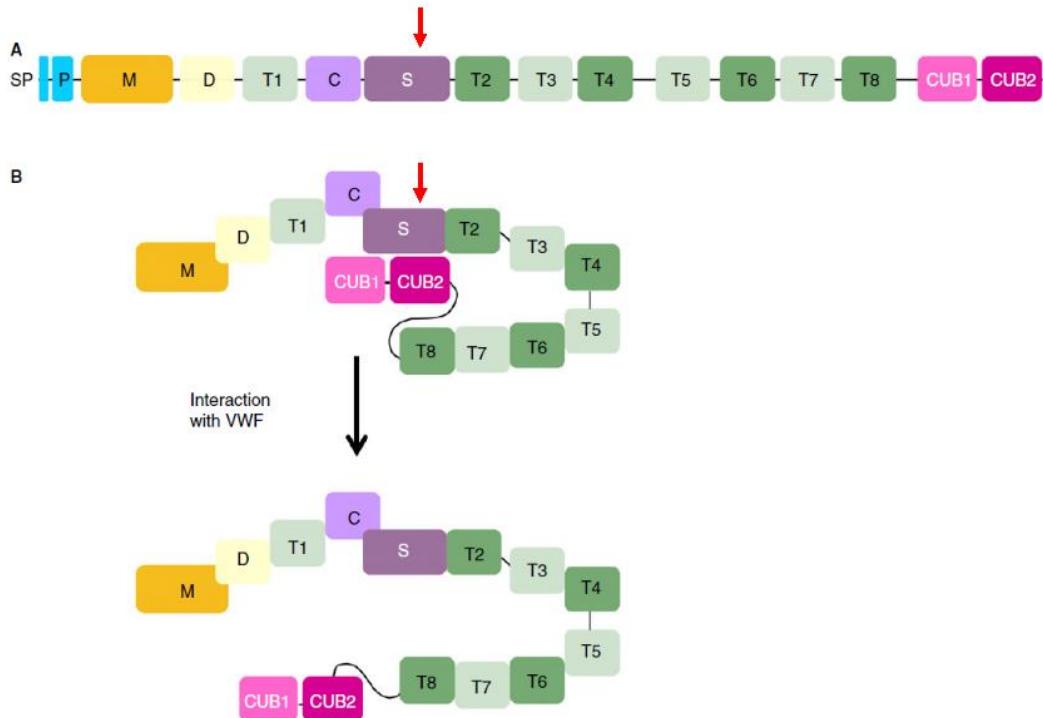
圖四、基因變異點於蛋白質功能區塊（protein domain）之位置。本例顯示此變異點位於 S 功能區塊，在未活化狀態下，S 功能區塊會與 CUB1 和 CUB2 功能區塊結合，於活化狀態時則會分離。次變異點恰恰位於 S 功能區塊，意味著此變異點可能三級結構形成相關。摘自 Journal of Thrombosis and Haemostasis, 16: 618–629.
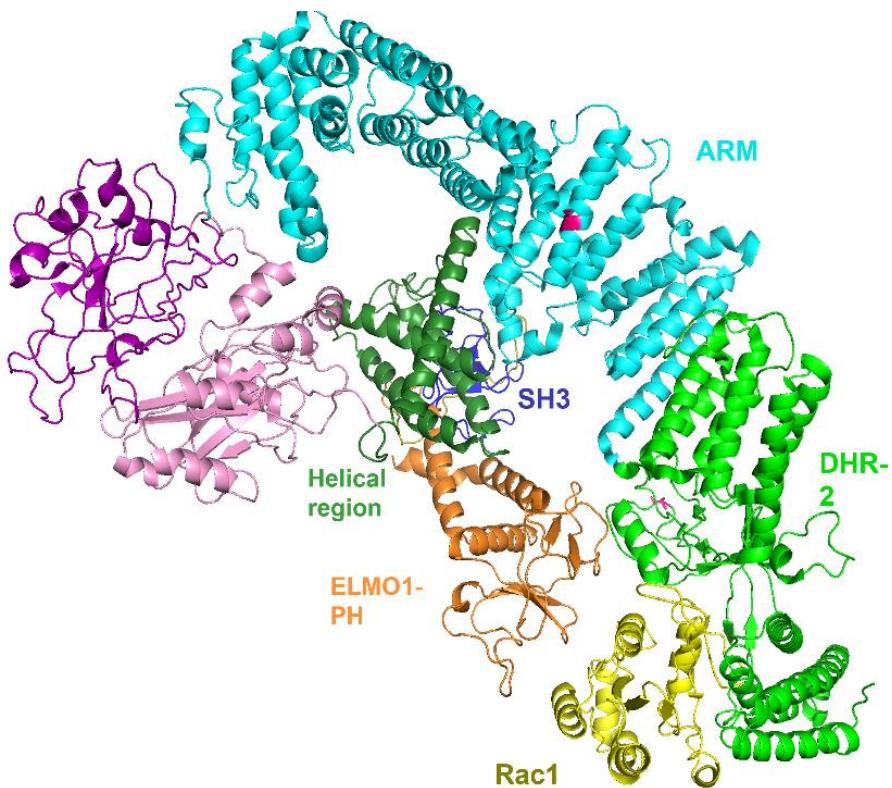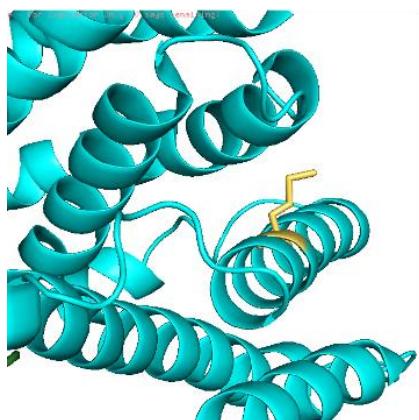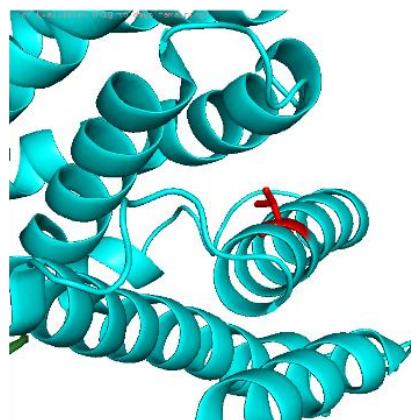
圖五、**Pymol** 模擬蛋白質結構。（上）標示突變點於蛋白質立體結構之相對應位置；（下）突變胺基酸與周圍結構示意圖。

心得

1. **過於繁重的臨床業務擠壓研究能量**：即便本院與波士頓兒童醫院臨床照護及研究規模差異甚大，目前臨床業務要求仍佔掉大多數的時間。參酌在美進修時的觀察，小兒血液科醫師臨床業務為每週一節門診，平均 6-8 週輪值一週的血液科住院病患照護。本人目前為本院唯一一位小兒血液腫瘤科醫師，不分晝夜平日假日均須及時反應處理次專科病患狀況，且血液科病患狀態不全如預期，有時狀況急轉直下，亦須即時到場處置。另一方面，除次專科病患照護外，亦須兼顧一般兒科照護。往往需擠壓私人休息時間和犧牲與家人相處時間才能以維持差強人意的研究表現。充滿荊棘的研究之途在在考驗意志力，令人萌生強大的挫折感。

2. **活細胞檢體庫之重要性**：隨著時代劇進，許多之前無解之案例有機會獲得解答。然而，關鍵的實驗驗證仍仰賴原始指標個案的檢體。標準確認生殖細胞系基因變異的檢體是皮膚衍生之纖維母細胞、再則是骨髓檢體衍生之纖維母細胞。口腔黏膜細胞及毛髮囊細胞因受汙染機會大及曾萃取之 DNA 量過少且純度不佳，皆不是理想的檢體。而週邊血液存在受體細胞基因變異影響之疑慮，並不是最建議的檢體。波士頓兒童醫院具有中央實驗室，保存週邊血液、週邊血液單個核細胞（peripheral blood mononuclear cell，PB-MNC）、骨髓檢體、骨髓單個核細胞（bone marrow mononuclear cell，BM-MNC）、骨髓衍生之纖維母細胞（bone marrow fibroblast）和皮膚組織衍生之纖維母細胞（skin fibroblast）。最後三種檢體均可經由細胞培養達到增值，在可接受代數之內均可進行相關功能性驗證的實驗操作。

3. **建立合作模式**：台灣醫療環境普遍與波士頓兒童醫院臨床照護及研究規模差異甚大，然為持續提升醫療照護品質，研究必不可缺。雖然實際操作上僅能針對現有資源做最大的利用，但至少做到資源不分散，建立不限於本院內部之合作模式團隊是唯一可行且刻不容緩之事。初步構想是結合不同院區基因

醫學及不同專科專長，同時跨體系合作，共同完成骨髓衰竭症候群檢測及基因檢查，提供更全面的精準診斷，如此一來，才能進入是否需調整治療決策的階段。也才有機會在未來得以鑽研造血系統致病機轉和研發製藥以提供病患更多治療選擇。

建議事項

1. **有能確實輪替職代該次專科的第二位人員**：臨床業務是影響研究的至關原因，尤其是病患狀況為高度不可預期和耗神費力之次專，如小兒血液腫瘤科。

2. **建立永續檢體庫**：中央實驗室協助處理及保存檢體對無獨立實驗室之研究人員是極有力的資源，在確保檢體品質狀態下，亦達到減少成本的優勢。考量基因研究蓬勃發展，擁有足夠及適當的檢體，對於研究如虎添翼。

3. **共享研究人事行政資源**：並非每位計畫主持人都隨時有足夠經費得以聘請全時研究人員，且研究人事費多用於實驗操作人員，然個案收案及相關行政流程繁瑣亦消耗計畫主持人研究心力，計畫主持人須同時兼顧臨床、研究操作、研究收案、申請倫理委員會核可、經費編列及執行，往往無法瞻前顧後，分神之餘即是延宕研究進度。建議由院方主導，讓計畫主持人得負擔之經費範圍提出申請，補助聘僱可協助執行行政事務的人員

Fibroblast culture from bone marrow

Materials:

- 10cm plates

- ChangD media (Irvine Scientific Cat. No. T105-500ml)

- Penicillin/Streptomycin (Gibco Cat. No. 15140-122) 用於未長出 colony 前，加進 ChangD

- Trypsin (0.05% EDTA)

- 1X PBS

- Freeze media (90% FBS + 10% DMSO)

- Grow in a hypoxic incubator for best results (5% O_2, 5% CO_2, 37 C). It can also be cultured in a regular normoxic incubator, but it will grow a bit slower

Procedure:

In ward: Collect 100-200µl BM in 5-10ml N/S (PBS) in sterile container (urine cup)

In lab

1. pipette above to 15ml conical tube, wash urine cup once with 3ml PBS.

2. centrifuge in RT (room temperature), 1200-1500rpm, 5min.

3. suction off supernatant, 可留一點點 N/S, 先用手敲打散 pellet，resuspend pellet with

(ChangD + 1% Pen/Strep) 8-10ml, and plate on a 10cm tissue culture plate

4. 24-36 hrs later, **change media**: suction off media and replace with (ChangD + 1% Pen/Strep) 10ml

5. check colonies grow status QD or QOD, **change media** once to twice a week

6. if you have a fair amount of colonies (70-80% confluent), <u>split cells</u>: trypsinize and replate

---

Estimation

D4-6: individual spindle-shaped fibroblast

D7-10: small clusters

*wait until D14 for fibroblasts grow, if still no growth: failed

W3-5: 70-80% confluent

Splitting Primary Fibroblasts

When it's around 70-80% confluent, you can split cells

1. Aspirate media (plate 斜置) and wash cells with 5 mL of PBS (沿左側邊緣).

2. Aspirate PBS (plate 斜置、於右側抽吸), add 1-2 mL of trypsin (用滴的), and incubate in the cell incubator for at least 1 min (讓細胞游離).

3. Add each 9 ml of Chang D media to 2-6 new plates (分盤的數量取決於細胞長的速度)

4. Check that cells are off the plate (細胞已游離，呈圓形) and add 3-4 mL of Chang D media to plate. **Break cells apart** (反覆抽吸細胞，確定黏在 plate 的細胞都被洗下來).

5. Collect all cells and transfer to a 15ml conical tube. Spin down at 1500rpm for 5 minutes

6. Suction off supernatant, flick the tube to loosen up cell pellet and resuspend with 1 mL of Chang D media.

7. Do a 1:8 split if cells are growing really fast by adding 125 ul of the resuspended pellet to the 8 plates. If cells are growing slowly, do a 1:2 plate into both plates by adding 500 ul to both plates. Give cells a shake (前後左右，勿轉圈) and incubate.
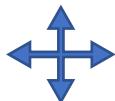
Freeze fibroblasts

When it's around 70-80% confluent, you can freeze down to store.

Freeze at least **5 vials** each time.

1. Aspirate media (plate 斜置) and wash cells with 5 mL of PBS (沿左側邊緣).

2. Aspirate PBS (plate 斜置、於右側抽吸), add 1-2 mL of trypsin (用滴的), and incubate in the cell incubator for at least 1 min (讓細胞游離).

3. Check that cells are off the plate (細胞已游離，呈圓形) and add 3-4 mL of Chang D media to plate. **Break cells apart** (反覆抽吸細胞，確定黏在 plate 的細胞都被洗下來).

4. Collect all cells and transfer to a 15ml conical tube. Spin down at 1500rpm for 5 minutes

5. Suction off supernatant, flick the tube to loosen up cell pellet and resuspend with 1 mL (FBS + 10% DMSO). Aliquot in 1ml in cryotubes

6. Place cryotubes in cryo contrainer and store at -80C overnight. Transfer tubes to liquid nitrogen storage

11

Label

Sample No.	AML-001
BM-F P0	BM-F P0
1x10cm plate	0.5x10cm plate
Processed date, Initials	2022/10/31 ML

Primary fibroblast culture from BM: P0 (passage 0)

Split with trypsin: P1, P2…

Generation of Fibroblasts from BM-MNC (Bone Marrow Mononuclear Cells)

Protocol:

1. Prepare before taking cells out of LN2 (liquid nitrogen): equilibrate Chang D media and 50 mL conical of FBS to RT (okay if cold). Take BM MNC cryovial from liquid nitrogen and place on dry ice.
2. Put the cryovial in bead or water bath and let thaw until there is only a small chunk of ice remaining. 幾乎融但未完全融化，進 hood 前先酒精消毒外面
3. Put 5mL of FBS into 15mL conical for resuspension/ centrifugation.
4. Using a 2mL serological pipette, take up 1mL FBS and gently resuspend cells in cryovial
   a. **Gently and quickly** (max 2 tubes) repipette 2 or 3 times until the ice is fully melted.
5. Take up 2mL of cells from the cryovial and gently pipette dropwise on top of FBS in 15mL conical. 用滴的
6. Spin down for 5 mins at 1200 rpm at RT (Wash #1).
7. Pipette Chang D media 8-10 mL into labeled 10cm plate.
   a. Label with Sample No., BM-F with passage number, thaw date, and initials.
8. When cells are done spinning, aspirate media.
   a. Don't get too close to the pellet because it will be loosely packed.
9. Resuspend cells in 1 mL Chang D. Plate 300 uL of the resuspended cells on 10 cm plate to grow fibroblasts. About remaining 700 uL:
   a. separate into 2-3 more plates for generating fibroblast
   b. Add 1mL PBS to the remaining 700uL of BM MNC cells in the conical.
   i. Count total cells and store back or
   ii. Proceed to DNA extraction (somatic mutation)
10. Change Chang D media the following day then twice a week until ready (70-80% confluent) for split. Eventually, 1 plate to extract DNA and 5 to freeze back.

Fibroblasts growth from BM-F

Protocol: steps 1-8 is same as above

9. Resuspend cells in 1 mL Chang D. Plate all resuspended cells on 10 cm plate to grow fibroblasts.
10. Change Chang D media the following day then maybe ready (70-80% confluent) for split 3 days later (usually within 1 week).

DNA Extraction

[Different steps]
BM-MNC or BM-Fibroblast samples
DNeasy Blood & Tissue Kit (Ref # 69506) (Collection tube: DNeasy Mini Spin Column)
1. Label Epp tubes (1, 2, 3…) and have corresponding subject ID # recorded in lab notebook. **Note**: Maximum number of cells to load a column is 5x10^6. In general, it is okay to process one 10cm plate of patient primary fibroblasts per column.
2. Resuspend (thawed) cell pellet in 200ul PBS.
3. Add 20ul **proteinase K** (in the kit box at RT).

Blood samples
QI Amp DNA Blood Mini Kit (Ref # 51106)
1. Label Epp tubes (1, 2, 3…) and have corresponding subject ID # recorded in lab notebook.   Process 1 column per sample at a time.
2. Add 20ul **QIAGEN Protease** into the bottom of an Epp tube. Quiagen Protease kept in 4C.
3. Add 200ul blood per tube (can add PBS to bring volume up to 200ul) (keep remainder of blood in vial on ice since will only process 200ul at a time. Flash freeze vial after confirm extracted enough DNA. Otherwise, bring vial to RT again and process another 200ul aliquot).

[Same steps]
4. Add 200ul AL buffer, shake 15 seconds, then vortex for 30s, and quick spin in Mini Mouse.
5. Place Epp tubes in heatblock set to 56C for 10 minutes.
6. Spin down Epp tubes in Mini Mouse.
7. Add 200ul ETOH (100%, 200 proof), shake 15 seconds, vortex 30s to mix thoroughly, and quick spin.
8. Transfer to spin column in 2ml collection tube, close the cap.
9. Centrifuge at 6,000g for 1 min, then transfer column to a fresh collection tube.
10. Add 500ul buffer AW1, centrifuge at 6,000g for 1min, then transfer column to a fresh collection tube.
11. Add 500ul buffer AW2, centrifuge at 17,000g for 3 min, then transfer column to a fresh collection tube.
12. Dry spin, 17,000g for 1 min.
13. Transfer column to a clean Epp tube (labeled on the side with Sample ID # and today's date).
14. Add AE (PB: 80uL; BM-F or BM-MNC: 100ul) to the column, incubate for 5 min, then centrifuge at 6,000g for 1 min.
15. Pipette the eluted DNA back on the column to re-elute, incubate for 5 min, then centrifuge at 6,000g for 1 min.
16. Nanodrop (blanked with AE), record DNA concentration (ng/ul), 260/280, and 260/230, and total eluted volume.
17. Print labels for tubes, store in -20 freezer and document on the -20 C log. Top = ID #. Side Labels:

Subject ID #
BL / BM-F / BM-MNC DNA
Date collected (collt'd ____)
DNA ng/ul
Date processed & initials

<Sanger sequence PCR primer design>

chr17:74737110A>C, RAB37, NM_001163989.1:c.101A>C (p.Asn34Thr), exon 1
NC_000017.11 (74671131..74747335)

Ensembl

https://useast.ensembl.org/index.html

choose yellow one but **with interested position**

Choose exon

tgagctcgggccgggtgacttaacagatgaggaagtgtctcggggccatcggcggaggcgcagcccaggggtccccagctccccgcct
cgccacctggggacagcccacggcccggggct[cgggcgccgcctgctgtcgcggtgcgcagc
GACTACGGGAACTCTTCCGCAGCAGACGGGGTCCCCGCGGCCCGCTCCCCCAGGGGCAAGCAAGCGAC
CACAGGGGACCGGTCCCGGGGCTGGATGTGGCTCATGTCCGAAGCGCACGGAGCCGAGCCGGTGTTGC
TCAGGGAGGCTGCCCGCCCCTTCACGCAGACCCTGCGGCTCTGCG<mark>TGCCCTCAGGGAA</mark>CAGCAAG
gtccgagccggtgtcgtcgaggggggcgacg]ggacggagggaggagcctgaggggtcccggtcgagggagggggaggagtgggcggg
gcggggtggggccgttcctgcgctctccttcgcctgcgggccggcactgctcacctctcgtcca
count 30 nucleotides of both intron sides [   ]

Primer3

https://primer3.ut.ee/

adjust GC, temperature to find primers

```
OLIGO                start   len       tm       gc%   any_th  3'_th  hairpin  seq
LEFT PRIMER           33      18     59.01    61.11     4.39    0.00     0.00  aagtgtctcggggccatc
RIGHT PRIMER          431     18     60.04    72.22     0.00    0.00     0.00  ccactcctcccctccctc
PRODUCT SIZE: 399, PAIR ANY_TH COMPL: 0.00, PAIR 3'_TH COMPL: 0.00

    1 tgagctcgggccgggtgacttaacagatgaggaagtgtctcggggccatcggcggaggcg
                                      >>>>>>>>>>>>>>>>>>
   61 cagcccaggggtccccagctccccgcctcgccacctggggacagcccacggcccggggct

  121 cgggcgccgcctgctgtcgcggtgcgcagcGACTACGGGAACTCTTCCGCAGCAGACGGG
      ************************************************************
  181 GTCCCCGCGGCCCGCTCCCCCAGGGGCAAGCAAGCGACCACAGGGGACCGGTCCCGGGGC
      ************************************************************
  241 TGGATGTGGCTCATGTCCGAAGCGCACGGAGCCGAGCCGGTGTTGCTCAGGGAGGCTGCC
      ************************************************************
  301 CGCCCCTTCACGCAGACCCTGCGGCTCTGCGTGCCCTCAGGGAACAGCAAGgtccgagcc
      ************************************************************
  361 ggtgtcgtcgaggggggcgacgggacggagggaggagcctgaggggtcccggtcgagggag
      *******************                             <<<<<<<
  421 gggaggagtgggcggggcggggtggggccgttcctgcgctctccttcgcctgcgggccg
      <<<<<<<<<<<<
  481 gcactgctcacctctcgtcca
```

● **Double check primers**

Primer BLAST

https://www.ncbi.nlm.nih.gov/tools/primer-blast/

"Refseq representative genomes" that may contain these ALT_REF_LOCI records; "Genomes for selected eukaryotic organisms (primary assemblies only)" should not contain these. Including those records is a more comprehensive search, but since alignments to those records is almost always non-informative, you may decide to avoid them for convenience.

- **Double check primers**

BLASTn: IDT protocol

https://blast.ncbi.nlm.nih.gov/Blast.cgi

https://www.idtdna.com/pages/education/decoded/article/tips-for-using-blast-to-locate-pcr-primers

Use these tips to refine Primer-BLAST results:

1. Concatenate the two primer sequences into one sequence separated by 5–10 Ns and enter into BLAST sequence box.

TGCTGTCTCCCATTTCCAGTNNNNNNNNNNNCACTACCTCTACCCTCCGTT

2. Before submitting, narrow the search by selecting the species, if known; otherwise, choose **Nucleotide Collection (nr/nt)**. If you're looking for RT-PCR primers, select the reference **mRNA sequences (refseq_mRNA)** database.

3. Under Program Selection, select the Somewhat similar sequences (blastn) program.

4. Under Algorithm parameters, decrease word size to 7, increase expect threshold to 1000, and turn off the low complexity filter.

In the example below, the results with a line connecting the 2 boxes indicate the 2 primers are in the same sequence (Figure 1A). Clicking on these results shows location within the sequence (between bases 1005–1766 of the *MRE11* transcript in yeast) and indicate the expected PCR product should be 762 bp in length (Figure 1B).
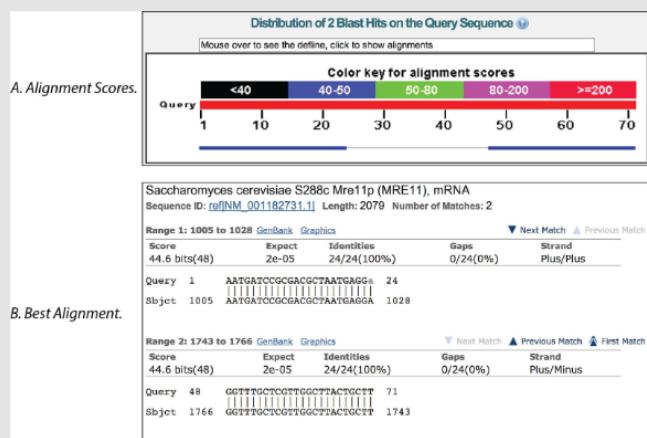
Figure 1. BLAST results show primer positioning and PCR fragment size. (A) Only the best alignment is shown (blue lines at the bottom). (B) Selecting the alignment provides the actual sequence with base by base alignment of the 2 portions of sequence showing 100% identity.

BLASTn: NCBI protocol

blast-help@ncbi.nlm.nih.gov

I started that search from the BLAST home page, selecting human in the "BLAST Genomes" section of the page. Then choose the database, RefSeq Genomic, because it includes those ALT_REF_LOCI records.

In the results, if you click on the alignment to the first **NT_**, then on the "Graphics" link, you can determine that the full primer alignments are to the TUBB gene.

點縮小，會看到這一段位於 TUBB 上

Homo sapiens chromosome 6 genomic scaffold, GRCh38.p14 alternate locus group ALT_REF_LOCI_7 HSCHR6_MHC_SSTO_CTG1

NCBI Reference Sequence: NT_167249.2
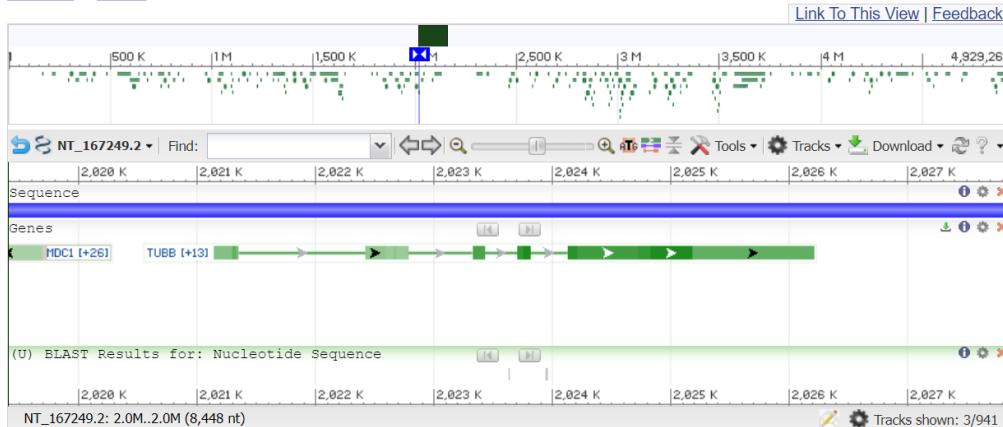
Read about these ALT_REF_LOCI records here:

https://www.ncbi.nlm.nih.gov/grc/help/faq/#difference-between-alternate-loci-and-novel-patch

Alternate loci and novel patches enable the reference assembly to represent allelic diversity. They are scaffold sequences that are given chromosome context through alignments to the corresponding chromosome regions. Alternate loci scaffolds and their alignments are included in major assembly releases, while novel patch scaffolds and their alignments are included in subsequent patch releases for that assembly. They can be considered functionally equivalent, as novel patches will be reassigned to the role of alternate loci scaffolds at the time of the next major assembly release. Assembly regions for which the GRC provides alternate loci or novel patch scaffolds are typically those with known alternate haplotypes (e.g. immune-associated regions), highly variable genomic regions (e.g. olfactory receptor regions) or those where there are structural variants having 5 Kb or more sequence not represented on the chromosome. Human alternate loci and all novel patch scaffolds also include one or more anchor sequence components to ensure their robust alignment to the chromosomes. Anchor sequences are component(s) that are also found in the corresponding chromosome. The sequence locations corresponding to anchor components are annotated on the GenBank records for all alternate loci and patch scaffolds. For more detail on patches please see Introductions to Patches.