

博士論文

論文題目

骨髄異形成症候群 944 例の標的シーケンスによる新たな分子メカニズムの理解と探索

氏 名 永田 安伸

骨髄異形成症候群 944 例の標的シーケンスによる新たな分子 メカニズムの理解と探索

東京大学大学院医学研究科 医学博士課程 生体物理医学 専攻

指導教員 宮川 清 教授

氏名 永田 安伸

目次

要旨	2
序文	3
方法	6
結果	19
考察	46
謝辞	50
参考文献	51

要旨

次世代シーケンス技術は骨髄異形成症候群（以下 MDS）の診断や予後予測に有用なことが示唆されており、我々はそれらに影響を及ぼす遺伝学的異常について調査した。944 例の MDS 患者、既知の 104 遺伝子を対象に標的ディープシーケンス技術を用い変異・欠失を解析した。845 例（89.5%）に遺伝子変異を認め、47 遺伝子が有意に変異していた。単変量解析で 25 遺伝子が有意となり、この内 14 遺伝子と臨床的な予後予測因子を組み合わせることで今までにない新たな予後予測モデルが構築可能となり、患者は生存が異なる 4 群に分類された。これは検証コホートでも確認され、MDS の予後予測に非常に有用性が高いことが示唆された。

序文

骨髄異形成症候群（以下 MDS）は白血球減少、貧血、血小板減少などの汎血球減少と高率の急性骨髄性白血病（以下 AML）への移行を特徴とする骨髄系造血器腫瘍の一つである(1, 2)。生物学的にも臨床学的にも多様性が認められ、正確な診断と個々の患者のリスク分類に基づいた最適な治療選択が現在の重要な治療戦略となっている(3-6)。実際に染色体情報や細胞形態学、末梢血情報などを元に国際予後分類（以下 IPSS）や類似の予後予測モデル(7, 8)が策定され、これらは予後を予測することが可能であった。近年さらに改変された国際予後分類（以下 IPSS-R）が発表され、優れた予後予測能を示した(9)。IPSS や IPSS-R は診断時に評価され、症状緩和治療のみが行われた症例において正確な予後分類が可能であった。予後不良 MDS で最も特徴的な臨床指標は骨髄の芽球細胞数であり、骨髄中の芽球細胞が増加することで AML への移行の頻度が高くなりその結果、予後不良となることが知られている。

また近年の次世代シーケンス技術の革新的な進歩により様々な癌腫でがんゲノミクスが解明され、今まで知られていなかった様々な機能を有する遺伝子、パスウェイの変異が同定され、それらがどのように病態に関わるのか明らかとなってきた(10, 11)。MDS においても今まで発見されていなかった様々な主要標的遺伝子が同定され、ここ 10 年で分子メカニズムの詳細は劇的に解明された(12)。これらの変異遺伝子は転

写因子やシグナル伝達など従来よく知られたパスウェイだけでなく(13-15)、エピゲノム (*TET2*(16-20)、*IDH1/2*(21-24)、*DNMT3A*(25, 26)、*ASXL1*(27-29)、*EZH2*(30, 31)) や RNA スプライシング機構など今までに報告のなかった新たなパスウェイ(32-35)にも関与しており、病態に関わる遺伝子はさらに解明が進んでいる(36-40)。最近、MDS またはその類縁の骨髄系腫瘍では 78%の患者ががん化に直接かかわる遺伝子変異を有する、という報告がされている(41)。AML で特徴的な *FLT3* や *NPM1*, *CEBPA* といった遺伝子異常は MDS で低頻度であり、*TET2* や *RUNX1* などは AML でも MDS でも同程度の変異頻度であった。再生不良性貧血でも MDS で認められるような遺伝子異常が報告されている。しかし、遺伝子異常を有する再生不良性貧血全例が、すべからく MDS のように末梢血に芽球が出現するようなことは観察されにくく、汎血球減少などの症状がないままそれらのクローンが存在し続けている症例もある。すなわち、遺伝子異常の有無のみで MDS と再生不良性貧血の境界を決定することは困難と考えられる。このように遺伝子変異の知識を蓄積することで我々の MDS 病態への理解が深まっているが、実際に臨床決断にどのように寄与するのか現時点では不明な点も多く、解決すべき危急の課題となっている。Bejar らは複数の標的遺伝子変異の有無により MDS 患者の予後を予測可能であると報告した(42, 43)。欧州白血病グループにおいても候補遺伝子変異の有無を調査することは MDS 患者の最終的な病型診断と信頼のお

ける予後予測に有用である、と推奨されている(44)。

本研究の目的は遺伝子解析の革新的な技術を用いる(45, 46)ことで判明する様々な遺伝子異常の有無が、臨床決断にどのように生かされ、MDS の病態をどのように深く解明するのか検証することである。我々は様々な病型を含む MDS 患者 944 症例を対象に既知もしくは想定される 104 個の遺伝子について、並列高速シーケンス技術を用いて遺伝子変異とコピー数異常を解析した。これらの遺伝子異常の情報により現存している予後予測モデルの性能を上回る新たな予後予測モデルの構築が可能かどうか、従来手法である二つの独立したコホートにおいてそれぞれ検証がなされた。また遺伝子変異の分布や相関、腫瘍アレル頻度や腫瘍内異質性、変異の序列や臨床学的指標との関連についても解析された。

方法

1 患者

成人 MDS 患者で様々な病型を含む全 944 症例が本研究に登録された。2005 年 8 月から 2011 年 8 月の間に採取された骨髄もしくは末梢血が共同研究機関であるミュンヘン白血病研究所から供与されそれらの DNA が解析対象となった。(表 1) これらは予後モデル作成のために探索コホート (730 症例) と検証コホート (214 症例) に分けられた。探索コホートにおいて追跡可能な症例は 670 例で、治療情報が判明したのは 648 例であった。(症状緩和治療のみ: 504 例 (77.8%)) 一方、検証コホートでは追跡可能な症例は 205 例で、治療情報が判明したのは 192 例であり、探索コホートとよく似た割合であった。(症状緩和治療のみ: 148 例 (77.1%)) また、53 例の正常人の非増幅ゲノム DNA や 16 例の追加 MDS 患者の腫瘍/正常ペア DNA 検体も我々のデータ解析の性能評価のために用いられた。すべての検体は細胞形態検査、染色体分析などの標準的な解析がされており(47)、一部は蛍光 *in situ* ハイブリダイゼーションにより特定遺伝子の解析が追加された。すべての患者は遺伝子解析について書面にてインフォームドコンセントが得られており、本研究はヘルシンキ宣言に則ってミュンヘン白血病研究所と東京大学いずれの倫理委員会においても承認を取得していた。

(東京大学 ヒトゲノム・遺伝子解析研究倫理審査委員会、審査番号: 948-(14)、研

究課題名：造血器腫瘍における遺伝子異常の網羅的解析)

表 1. 患者背景とその特徴

パラメーター	全コホート (n=944) 患者数 (比, 区間もしくはパーセンテージ)
全症例数	944
男性：女性 (比)	580:364 (1.6)
年齢中央値 (区間)	72.8 (23.3–90.8)
<50	46 (4.9%)
50- <60	81 (8.6%)
60- <70	240 (25.4%)
70- <80	423 (44.8%)
≥80	154 (16.3%)
追跡可能症例数	875
(追跡期間の中央月数)	(32.3)
(生存期間の中央月数)	(54.5)
MDS 病型 (WHO, 2008)	944
不応性貧血 (RA)	41 (4.3%)
環状鉄芽球を伴う不応性貧血 (RARS)	81 (8.6%)
血小板増加を伴った環状鉄芽球増加を伴う不応性貧血 (RARS-T)	28 (3.0%)
多血球系異型性を伴う不応性血球減少症 (RCMD)	195 (20.7%)
多血球系異形成と環状鉄芽球を伴う不応性血球減少症 (RCMD-RS)	183 (19.4%)
芽球増加を伴う不応性貧血 - 1 (RAEB-1)	191 (20.7%)
芽球増加を伴う不応性貧血 - 2 (RAEB-2)	188 (19.9%)
5q-症候群 (MDS with isolated 5q-deletion)	37 (3.9%)
染色体情報	944
正常核型	648 (68.6%)
異常核型	296 (31.4%)

IPSS リスク分類	848
Low	324 (38.2%)
Intermediate-1	319 (37.6%)
Intermediate-2	171 (20.2%)
High	34 (4.0%)
IPSS-R リスク分類	848
Very low	122 (14.4%)
Low	340 (40.1%)
Intermediate	203 (23.9%)
High	130 (15.3%)
Very high	53 (6.3%)

2 手法

MDS において報告のある、もしくは病態に関わることが想定される 104 個の遺伝子について、全 944 症例を対象に解析された（表 2A）。これら標的遺伝子は下記に示す 6 個の基準で選択決定しており、その根拠となっている論文についても追記している。（表 2B）

1) 2011 年 10 月 30 日時点で、MDS もしくは他の骨髄系腫瘍で変異報告のある遺伝子（51 個）

2) 2011 年 10 月 30 日時点で、我々の研究室の全エクソン解析もしくは標的シーケンス解析により nonsense もしくは frameshift 変異が同定されている遺伝子（13 個）

3) 遺伝性鉄芽球性貧血の原因遺伝子（2 個）

4) 既知の標的と同族体（4 個）

以下の候補は全体で 500Kb のカスタムアレイを作成する際の残りを埋めるためいくつか探索的に調査した。

5) 既知の変異遺伝子と同一もしくは類似の機能的パスウェイに認められる遺伝子 (スプライシング 3 個、コヒーシ関連 9 個、DNA 修復機構 13 個、エピゲノム 4 個、シグナル伝達 2 個)

6) ホモ欠失もしくは UPD などのコピー数異常が認められ標的候補と考えられる遺伝子(3 個)

アジレント社の Sureselct の技術を用いて標的遺伝子領域を濃縮し、次世代シーケンサーであるイルミナ社の HiSeq2000 により超並列大量シーケンスを行った。すべてのシーケンスデータは我々が独自に開発したアルゴリズムによって解析され(48, 49)、そのパイプラインはシーケンスエラーやマッピングエラーだけではなく、既存のデータベースに登録がある、もしくはアレル頻度から可能性が高いと想定される一塩基多型 (SNPs) を取り除くことが可能となり最終的に発がんに関わる体細胞性変異を高率に同定することが可能となった。(50) (図 1) 遺伝子領域のコピー数はそれぞれのエクソン領域におけるシーケンスリード数を直接数える手法によって計測され、正常核型 648 例のうち 494 例 (76.2%) に関してはアレイ CGH によるコピー数解析も併せて行われた。

生存解析に関してはそれぞれの遺伝子異常ごとに COX 回帰法により調査され、予後予測モデル作成に関する因子の選択には LASSO 法 (the Least Absolute Shrinkage and Selection Operator) が用いられた(51, 52)。COX 回帰による線形モデルにより患者のリスク分類が決定された。予後予測モデルは探索コホートにおいて作成され、独立した検証コホートにおいて確認が行われ、その性能は赤池の基本統計量や J-テスト(53)により従来の予後予測モデルである IPSS-R と比較された。

表 2A. 標的遺伝子

Gene	Inclusion criteria	Gene	Inclusion criteria
<i>ABCG8</i>	Reported in MDS	<i>LAMB4</i>	Nonsense mutation in WES in MDS
<i>ALAS2</i>	Responsible gene for congenital sideroblastic anemia	<i>LUC7L2</i>	Reported in MDS; RNA splicing
<i>ARID1B</i>	Histone modification	<i>MED12</i>	Cohesin/CTCF pathway
<i>ASH1L</i>	Histone modification	<i>MLH3</i>	DNA repair
<i>ASXL1</i>	Reported in MDS; histone modification	<i>MPL</i>	Reported in MDS; cytokine receptor
<i>ATM</i>	Frameshift mutation in WES in MDS	<i>MRE11A</i>	Nonsense mutation in target sequencing in MDS
<i>ATRX</i>	Reported in MDS; chromatin regulation	<i>MSH6</i>	DNA repair
<i>BCOR</i>	Reported in myeloid malignancies	<i>NCOR1</i>	BCOR homologue
<i>BCORL1</i>	Reported in myeloid malignancies	<i>NCOR2</i>	BCOR homologue
<i>BRAF</i>	Reported in myeloid malignancies	<i>NF1</i>	Reported in myeloid malignancies
<i>BRCA1</i>	DNA repair	<i>NIPBL</i>	Cohesin/CTCF pathway
<i>BRCC3</i>	Nonsense mutation in WES in MDS	<i>NOTCH1</i>	Reported in myeloid malignancies
<i>CBL</i>	Reported in MDS; RAS signaling	<i>NOTCH2</i>	NOTCH1 homologue
<i>CBLB</i>	Reported in MDS; RAS signaling	<i>NPM1</i>	Reported in MDS
<i>CDCA5</i>	Nonsense mutation in WES in MDS	<i>NRAS</i>	Reported in MDS; RAS signaling
<i>CEBPA</i>	Reported in MDS; transcription	<i>PDS5B</i>	Cohesin/CTCF pathway
<i>CRKL</i>	Nonsense mutation in WES in MDS	<i>PHF6</i>	Reported in myeloid malignancies
<i>CTCF</i>	Cohesin/CTCF pathway	<i>PHF8</i>	Nonsense mutation in WES in MDS
<i>DCLRE1C</i>	DNA repair	<i>PIGA</i>	Reported in MDS
<i>DNMT3A</i>	Reported in MDS; DNA methylation	<i>PRPF40B</i>	Reported in MDS; RNA splicing
<i>DNMT3B</i>	Histone modification	<i>PRPF8</i>	Reported in MDS; RNA splicing
<i>DOT1L</i>	Nonsense mutation in WES in MDS	<i>PTPN11</i>	Reported in MDS; RAS signaling
<i>ELANE</i>	Homozygous mutations in WES of 19qUPD-positive case	<i>PTPRD</i>	PTPRT homologue
<i>ESCO2</i>	Cohesin/CTCF pathway	<i>PTPRT</i>	Reported in myeloid malignancies

<i>ETV6</i>	Reported in MDS; transcription	<i>PTTG1</i>	Nonsense mutation in WES in MDS
<i>EZH2</i>	Reported in MDS; histone modification	<i>RAD21</i>	Reported in myeloid malignancies
<i>FANCA</i>	Nonsense mutation in target sequencing in MDS	<i>RFC1</i>	DNA repair
<i>FANCB</i>	DNA repair	<i>RUNX1</i>	Reported in MDS; transcription
<i>FANCD2</i>	DNA repair	<i>SETD2</i>	Histone modification
<i>FANCE</i>	DNA repair	<i>SF1</i>	Reported in MDS; RNA splicing
<i>FANCF</i>	DNA repair	<i>SF3A1</i>	Reported in MDS; RNA splicing
<i>FANCG</i>	DNA repair	<i>SF3A3</i>	RNA splicing
<i>FANCI</i>	DNA repair	<i>SF3B1</i>	Reported in MDS; RNA splicing
<i>FANCL</i>	DNA repair	<i>SF3B14</i>	RNA splicing
<i>FANCM</i>	DNA repair	<i>SLC25A38</i>	Responsible gene for congenital sideroblastic anemia
<i>FBXW7</i>	Reported in hematologic malignancies; NOTCH signaling	<i>SMAD4</i>	Reported in myeloid malignancies
<i>FLT3</i>	Reported in MDS; cytokine receptor	<i>SMC1A</i>	Reported in myeloid malignancies
<i>GATA1</i>	Reported in myeloid malignancies	<i>SMC1B</i>	Cohesin/CTCF pathway
<i>GATA2</i>	Reported in myeloid malignancies	<i>SMC3</i>	Reported in myeloid malignancies
<i>GNAS</i>	Reported in MDS; signaling	<i>SRRM1</i>	RNA splicing
<i>GNB1</i>	Recurrent mutation in WES in MDS	<i>SRSF2</i>	Reported in MDS; RNA splicing
<i>GPRC5A</i>	Receptor/signaling	<i>STAG1</i>	Cohesin/CTCF pathway; Missense mutation in WES in MDS
<i>GRID1</i>	Nonsense mutation in WES in MDS	<i>STAG2</i>	Reported in myeloid malignancies
<i>HIF1A</i>	Homozygous mutations in WES of 14qUPD-positive case	<i>STAG3</i>	Cohesin/CTCF pathway
<i>IDH1</i>	Reported in MDS; DNA methylation, histone modification	<i>TET2</i>	Reported in MDS; DNA methylation
<i>IDH2</i>	Reported in MDS; DNA methylation, histone modification	<i>TP53</i>	Reported in MDS; transcription, DNA repair
<i>IRF1</i>	Reported in MDS; transcription	<i>TP53BP1</i>	Nonsense mutation in WES in MDS
<i>JAK2</i>	Reported in MDS; kinase signaling	<i>U2AF1</i>	Reported in MDS; RNA splicing
<i>KDM6A</i>	Reported in MDS; histone modification	<i>U2AF2</i>	Reported in MDS; RNA splicing

<i>KDR</i>	Homozygous mutations in WES of 4qUPD-positive case	<i>WAPAL</i>	Cohesin/CTCF pathway
<i>KIT</i>	Reported in MDS; cytokine receptor	<i>WT1</i>	Reported in myeloid malignancies
<i>KRAS</i>	Reported in MDS; RAS signaling	<i>ZRSR2</i>	Reported in MDS; RNA splicing

表 2B. 標的遺伝子のパスウェイと参考文献

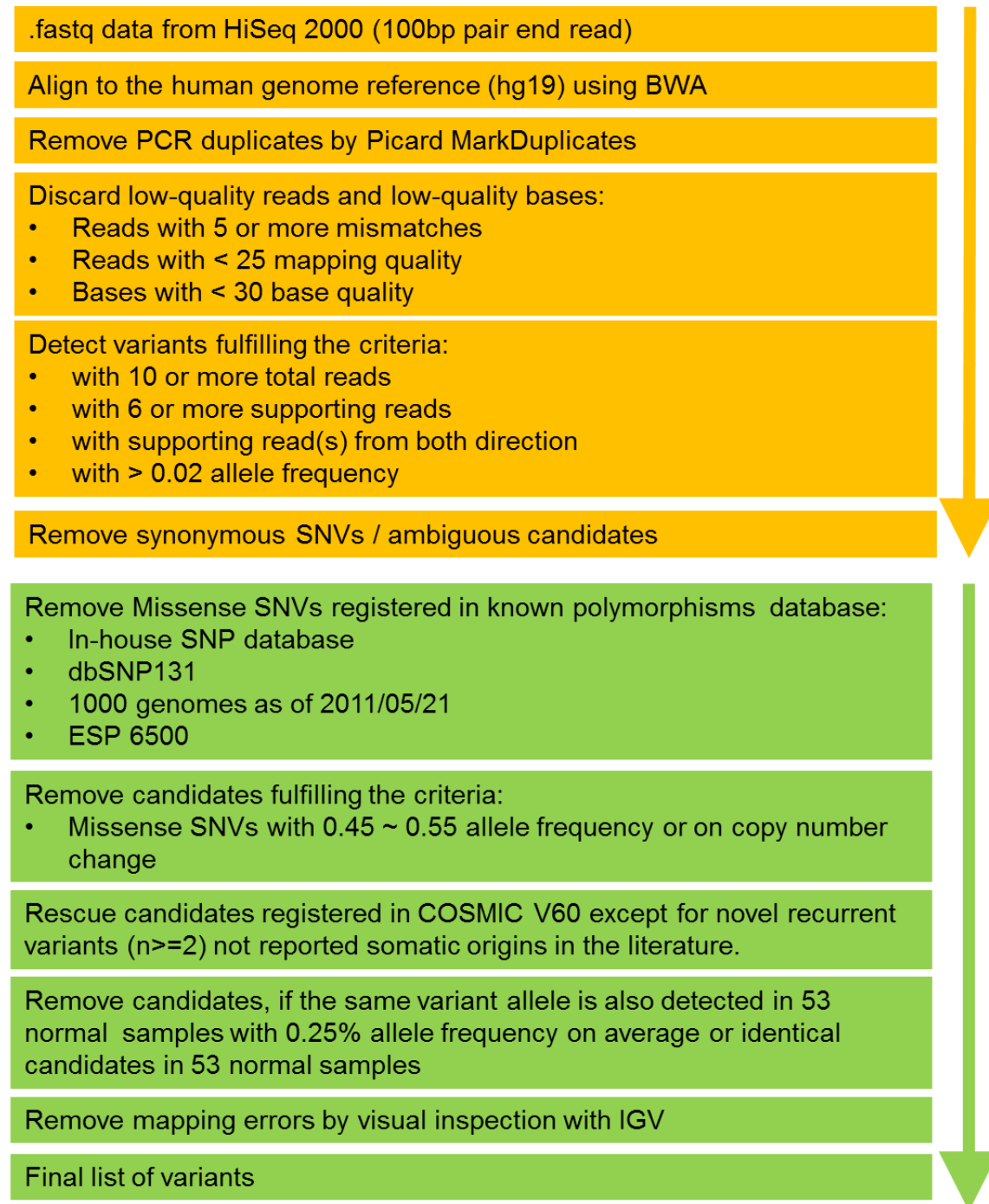
gene	Pathway	GO ID	Paper
<i>ASXL1</i>	Chromatin modification	-	Journal of Clinical Oncology 2011 Jun 20;29(18):2499-506.
<i>EZH2</i>	Chromatin modification	GO:0016568	
<i>KDM6A</i>	Chromatin modification	GO:0016568	
<i>ATRX</i>	Chromatin modification	-	European Journal of Haematology 2006 May;76(5):432-5, 453.
<i>ARID1B</i>	Chromatin modification	GO:0016568	
<i>SETD2</i>	Chromatin modification	GO:0016568	
<i>ASH1L</i>	Chromatin modification	GO:0016568	
<i>DOT1L</i>	Chromatin modification	GO:0016568	
<i>DNMT3B</i>	Chromatin modification	GO:0016568	
<i>PHF8</i>	Chromatin modification	-	Nucleic Acids Research. 2012 Oct 1;40(19):9429-40.
<i>STAG2</i>	Cohesin	GO:0007059	Cell. 2012 Jul 20;150(2):264-78.
<i>SMC3</i>	Cohesin	GO:0007059	Cell. 2012 Jul 20;150(2):264-78.
<i>CTCF</i>	Cohesin	GO:0007059	
<i>RAD21</i>	Cohesin	GO:0007059	Cell. 2012 Jul 20;150(2):264-78.
<i>SMC1A</i>	Cohesin	GO:0007059	Cell. 2012 Jul 20;150(2):264-78.
<i>PDS5B</i>	Cohesin	GO:0007059	
<i>NIPBL</i>	Cohesin	GO:0007059	
<i>STAG1</i>	Cohesin	GO:0007059	
<i>STAG3</i>	Cohesin	GO:0007059	
<i>SMC1B</i>	Cohesin	-	Annual Review of Genetics 2009;43:525-58.
<i>WAPAL</i>	Cohesin	-	Annual Review of Genetics 2009;43:525-58.
<i>CDCA5</i>	Cohesin	GO:0007059	
<i>ESCO2</i>	Cohesin	-	Annual Review of Genetics 2009;43:525-58.
<i>PTTG1</i>	Cohesin	GO:0007059	
<i>TET2</i>	DNA methylation	-	Nature Reviews Cancer. 2012 Sep;12(9):599-612
<i>DNMT3A</i>	DNA methylation	-	Nature Reviews Cancer. 2012 Sep;12(9):599-612
<i>IDH2</i>	DNA methylation	-	Nature Reviews Cancer. 2012 Sep;12(9):599-612
<i>IDH1</i>	DNA methylation	-	Nature Reviews Cancer. 2012 Sep;12(9):599-612
<i>ATM</i>	DNArepair	GO:0006281	
<i>BRCC3</i>	DNArepair	GO:0006281	

<i>FANCL</i>	DNArepair	GO:0006281	
<i>FANCA</i>	DNArepair	GO:0006281	
<i>MRE11A</i>	DNArepair	GO:0006281	
<i>FANCM</i>	DNArepair	GO:0006281	
<i>MSH6</i>	DNArepair	GO:0006281	
<i>FANCD2</i>	DNArepair	GO:0006281	
<i>FANCB</i>	DNArepair	GO:0006281	
<i>DCLRE1C</i>	DNArepair	GO:0006281	
<i>FANCI</i>	DNArepair	GO:0006281	
<i>BRCA1</i>	DNArepair	GO:0006281	
<i>FANCE</i>	DNArepair	GO:0006281	
<i>FANCG</i>	DNArepair	GO:0006281	
<i>MLH3</i>	DNArepair	GO:0006281	
<i>FANCF</i>	DNArepair	GO:0006281	
<i>RFC1</i>	DNArepair	GO:0006281	
<i>LAMB4</i>	Other	GO:0007155	
<i>PIGA</i>	Other	-	
<i>ALAS2</i>	Other	GO:0048534	
<i>GNB1</i>	Other	-	
<i>MED12</i>	Other	-	
<i>ABCG8</i>	Other	-	
<i>GRID1</i>	Other	-	
<i>SLC25A38</i>	Other	GO:0048534	
<i>CRKL</i>	Other	GO:0048534	
<i>ELANE</i>	Other	-	
<i>CBL</i>	RAS pathway	-	Nature. 2009 Aug 13;460(7257):904-8
<i>NRAS</i>	RAS pathway	GO:0048523	
<i>NF1</i>	RAS pathway	GO:0048523	
<i>KRAS</i>	RAS pathway	GO:0048523	
<i>PTPN11</i>	RAS pathway	GO:0048523	
<i>BRAF</i>	RAS pathway	GO:0048523	
<i>CBLB</i>	RAS pathway	GO:0048523	
<i>JAK2</i>	Receptors/Kinases	GO:0007166	
<i>MPL</i>	Receptors/Kinases	GO:0007166	

<i>PTPRD</i>	Receptors/Kinases	GO:0007166	
<i>GNAS</i>	Receptors/Kinases	-	Journal of Clinical Endocrinology Metabolism. 2012 Jul;97(7):2404-13.
<i>FBXW7</i>	Receptors/Kinases	-	Haematologica. 2011 Dec;96(12):1808-14
<i>KIT</i>	Receptors/Kinases	GO:0007166	
<i>KDR</i>	Receptors/Kinases	GO:0007166	
<i>FLT3</i>	Receptors/Kinases	GO:0007166	
<i>GPRC5A</i>	Receptors/Kinases	-	Cancer Biology & Therapy. 2009 May;8(10):951-62.
<i>PTPRT</i>	Receptors/Kinases	GO:0007166	
<i>NOTCH1</i>	Receptors/Kinases	GO:0007166	
<i>NOTCH2</i>	Receptors/Kinases	GO:0007166	
<i>SF3B1</i>	Splicing	GO:0008380	
<i>SRSF2</i>	Splicing	-	Nature. 2011 Sep 11;478(7367):64-9
<i>ZRSR2</i>	Splicing	GO:0008380	
<i>U2AF1</i>	Splicing	GO:0008380	
<i>PRPF8</i>	Splicing	GO:0008380	
<i>LUC7L2</i>	Splicing	-	Blood. 2012 Apr 5;119(14):3203-10.
<i>SRRM1</i>	Splicing	GO:0008380	
<i>U2AF2</i>	Splicing	GO:0008380	
<i>PRPF40B</i>	Splicing	GO:0008380	
<i>SF1</i>	Splicing	GO:0008380	
<i>SF3A3</i>	Splicing	GO:0008380	
<i>SF3A1</i>	Splicing	GO:0008380	
<i>SF3B14</i>	Splicing	GO:0008380	
<i>RUNX1</i>	Transcription	GO:0030528	
<i>TP53</i>	Transcription	GO:0030528	
<i>BCOR</i>	Transcription	GO:0030528	Haematologica. 2012 Jan;97(1):3-5
<i>PHF6</i>	Transcription	-	Nature Genetics. 2010 Apr;42(4):338-42.
<i>ETV6</i>	Transcription	GO:0030528	
<i>NCOR2</i>	Transcription	GO:0030528	
<i>NCOR1</i>	Transcription	GO:0030528	
<i>BCORL1</i>	Transcription	-	Haematologica. 2012 Jan;97(1):3-5
<i>TP53BP1</i>	Transcription	GO:0030528	
<i>NPM1</i>	Transcription	GO:0030528	

<i>CEBPA</i>	Transcription	GO:0030528	
<i>GATA2</i>	Transcription	GO:0030528	
<i>IRF1</i>	Transcription	GO:0030528	Science. 1993 Feb 12;259(5097):968-71.
<i>WT1</i>	Transcription	GO:0030528	
<i>HIF1A</i>	Transcription	GO:0030528	
<i>GATA1</i>	Transcription	GO:0030528	
<i>SMAD4</i>	Transcription	GO:0030528	

図 1. 体細胞変異を同定するパイプライン



解析パイプラインは二つのパートから成り立っている。初めのパートでは信頼のおけるアミノ酸置換を伴う一塩基置換や挿入・欠失を同定し、後半のパートでは複数の既知の SNPs データベースを用いて多型を除去し、既知の体細胞性変異データベース (COSMIC) (54) に登録のある変異を抽出することで最終的に信頼度の高い体細胞性変異を同定する。

結果

(1) 標的シーケンス

944 症例、104 遺伝子に関する標的シーケンスにおける平均の読み取り深度は 1066 回 (257-2306 回) であった。標的領域はエクソン領域と 6 塩基のスプライスサイトの総和である 357,861 塩基対であり、30 回以上の読み取り深度が得られている領域は 99%以上であった。シーケンスエラー、マッピングエラーを取り除き、既知もしくは可能性の高い SNPs を除外したところ、最終的に 2764 個の一塩基置換と短い挿入・欠失が同定された。(表 2) 96 個の遺伝子については少なくとも 1 個以上の変異を認め、それらは高率で体細胞性変異であることが示唆された。これらの検証のためランダムに 300 個の変異が選択され、サンガーシーケンス法やイルミナ社の Miseq によるディープシーケンス法など異なる手法による検証の結果、これらの内 299 個 (99.7%) の変異が確認された。53 例の正常人由来 DNA について同様の解析を行ったところ、わずか 10 個の変異 (正常人には変異は存在しないため偽陽性) しかコールされなかった。我々のパイプラインの正確性を評価する目的で 16 例の MDS 患者の腫瘍/正常ペア DNA 検体を解析したところ 75 個の変異がコールされ、ペア正常検体を解析することでその内 74 個 (98.7%) が真の体細胞性変異であることが確定した。最後に MLL 白血病研究所ですでに判明している遺伝子変異情報と照らし合わせを行ったところ、446 症例で全 1207 遺伝子領域が解析対象となり、結果の一致率は 99.4%

(1200/1207)であった。

(2) 遺伝子変異の頻度と分布

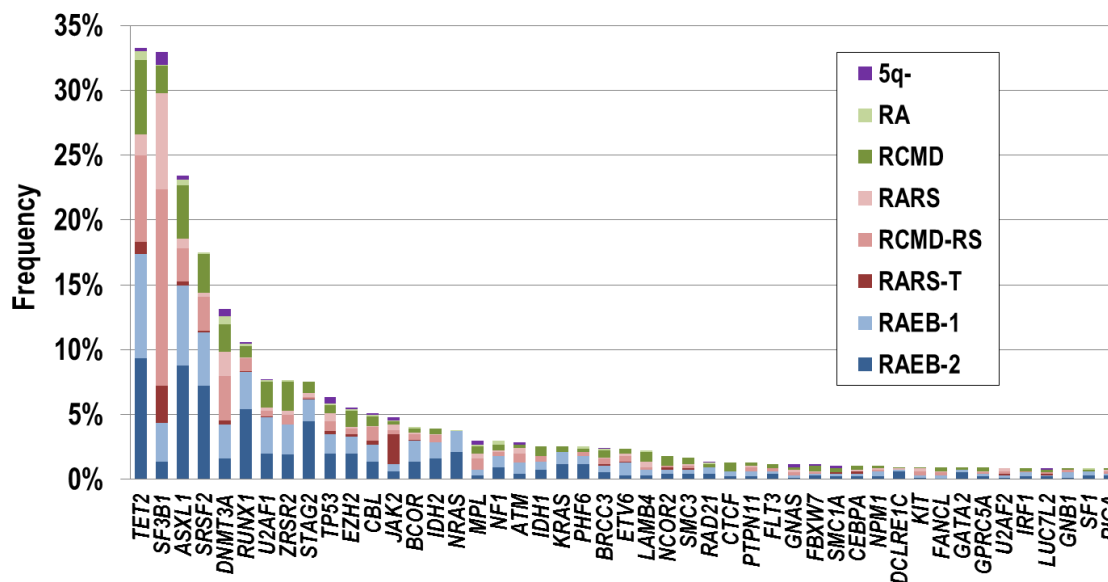
全 944 例の内、845 例 (89.5%) が少なくとも 1 個以上の遺伝子変異を有しており、一症例あたりの中央値は 3 個 (0–12 個) であった。この内 574 例 (67.9%) は正常核型であった。変異はなく、コピー数欠失のみが同定された症例は 19 例認められた。

68 個の遺伝子が 0.5%以上の頻度で変異しており、104 遺伝子のタイプ 1 エラーを計算し多重検定を Benjamini-Hochberg 法にて補正した(55)ところ、47 遺伝子が統計学的に有意であると判定された。(表 3) ($q < 0.01$) 10%以上の頻度で変異している遺伝子は 6 個であった。(TET2,SF3B1,ASXL1,SRSF2,DNMT3A,RUNX1) また、2-10%の頻度で変異しているのは U2AF1,ZRSR2,STAG2,TP53,EZH2,CBL,JAK2,BCOR,IDH2, NRAS,MPL,NF1,ATM,IDH1,KRAS,PHF6,BRCC3,ETV6,LAMB4 であった。有意に変異していた遺伝子の多くは MDS もしくは他の骨髄系腫瘍で変異が報告されているものであった(56-61)。しかし、BRCC3(32),FANCL,LUC7L2(62),STAG2 や RAD21,SMC3,SMC1A(63) などのコヒーシン複合体(37)、CTCF,GPRC5A,LAMB4 や IRF1(64, 65)など新規に同定された、もしくは高頻度に変異していることが再確認された遺伝子も含まれていた。

変異が認められた遺伝子はいくつかの機能的なパスウェイに分類され、それらは MDS の病態を特徴づけることが示唆された。これらの中で、最も高頻度に変異して

いるパスウェイは RNA スプライシングであり、約 64%の症例で変異が認められ以下 DNA メチル化、クロマチン修飾、転写、RAS/シグナル伝達、コヒーシン、DNA 修復の順であった。(図 3) 平均の変異個数や分布に関しては病型 (WHO 分類) により大きな特徴が認められた。*SF3B1* 変異は環状鉄芽球を有する病型の大多数 (82.2%、240/292) と 5q-症候群の多く (24.3%、9/37) に認められた(66)。*SF3B1*、*DNMT3A*、*JAK2*、*MPL* を除いて共通する変異遺伝子の多くは RA/RARS の病型と比較して、高リスクの病型 (RAEB-1,2) においてより高頻度であった。(図 2) よく似た結果として、平均の変異個数は高リスクの病型になるほど多い傾向が認められた。(図 4) 正常核型の RA や RCMD ですら 72.7% (133/183) に遺伝子変異を認めた。また、いくつかのがん抑制遺伝子については従来の染色体分析やアレイ CGH によりコピー数の欠失が報告されてきた。標的シーケンス技術を用いることで *EZH2*、*LUC7L2*、*ETV6*、*TET2*、*IRF1*、*TP53*、*PHF6* などがん抑制遺伝子のコピー数欠失について感度よく同定することが可能であった。(図 5) これら遺伝子変異と欠失を融合させることにより、MDS の遺伝学的異常の展望を描くことに成功した。(図 6)

図 2. MDS で有意に変異している遺伝子



様々な病型 944 例において有意に変異している 47 遺伝子についての変異頻度を示す。横軸は遺伝子、縦軸は変異頻度で、それぞれの病型ごとに色分けして記載をする。

表 3. 標的遺伝子の変異と欠失

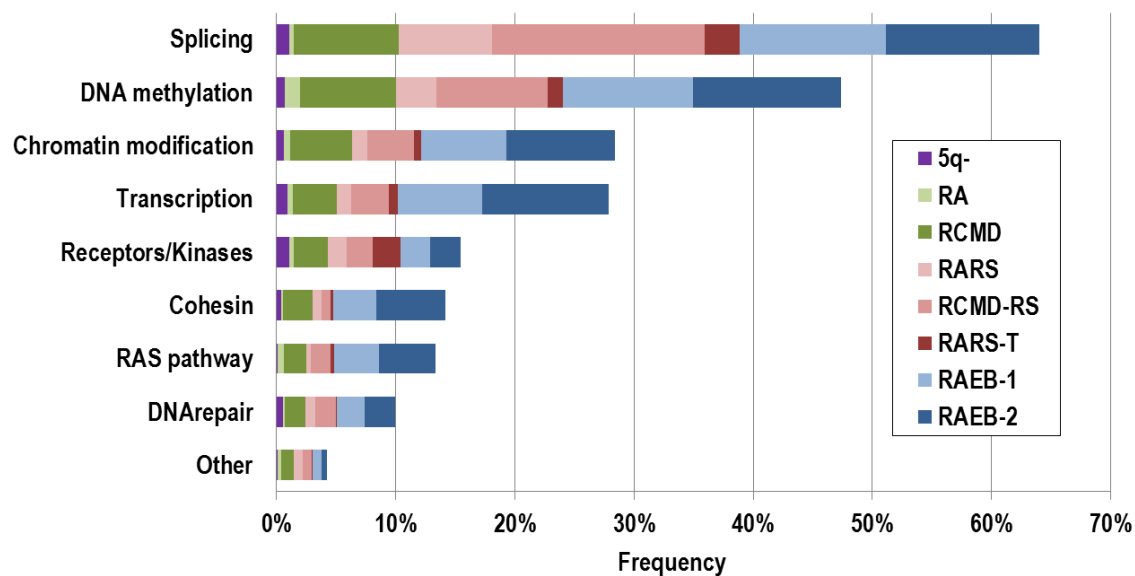
gene	# of mutations (missense mutations)	# of mutated cases	p-value	q-value	# of cases with copy number deletion	total
<i>TET2</i>	515 (108)	314	< 1.00E-300	< 1.00E-300	15	318
<i>SF3B1</i>	320 (317)	311	< 1.00E-300	< 1.00E-300	0	311
<i>SRSF2</i>	166 (138)	165	< 1.00E-300	< 1.00E-300	1	166
<i>ASXL1</i>	228 (0)	221	1.49E-295	3.87E-294	5	225
<i>RUNX1</i>	121 (44)	100	3.77E-184	7.84E-183	2	102
<i>DNMT3A</i>	129 (74)	124	1.04E-162	1.80E-161	5	127
<i>U2AF1</i>	75 (74)	73	7.95E-121	1.18E-119	0	73
<i>TP53</i>	76 (59)	60	3.24E-107	4.21E-106	17	67
<i>ZRSR2</i>	77 (19)	72	7.42E-103	8.57E-102	3	75
<i>STAG2</i>	96 (2)	71	3.98E-96	4.14E-95	7	77
<i>EZH2</i>	63 (36)	52	5.14E-67	4.86E-66	45	95
<i>NRAS</i>	42 (42)	36	8.92E-63	7.73E-62	0	36
<i>CBL</i>	61 (45)	48	3.67E-60	2.94E-59	0	48
<i>IDH2</i>	37 (37)	37	2.12E-40	1.57E-39	0	37
<i>PHF6</i>	32 (10)	24	5.16E-36	3.58E-35	4	27
<i>JAK2</i>	45 (45)	45	5.71E-35	3.71E-34	0	45
<i>KRAS</i>	25 (24)	24	1.68E-33	1.03E-32	0	24
<i>BRCC3</i>	24 (7)	23	5.47E-26	3.16E-25	4	26
<i>MPL</i>	29 (21)	28	2.39E-25	1.31E-24	0	28
<i>BCOR</i>	41 (6)	38	1.74E-24	9.05E-24	2	40
<i>IDH1</i>	24 (23)	24	1.06E-23	5.25E-23	0	24
<i>ETV6</i>	23 (7)	22	8.78E-22	4.15E-21	18	40
<i>ATM</i>	36 (23)	27	7.4E-12	3.35E-11	11	34
<i>PTPN11</i>	14 (14)	12	1.52E-09	6.59E-09	1	13
<i>NF1</i>	30 (13)	28	3.7E-09	1.54E-08	6	34
<i>RAD21</i>	13 (6)	13	2.13E-08	8.52E-08	0	13
<i>LAMB4</i>	21 (13)	21	6.27E-08	2.42E-07	43	64
<i>NPM1</i>	9 (0)	9	7.68E-08	2.85E-07	25	34

<i>CEBPA</i>	9 (4)	9	3.41.E-07	1.22E-06	1	10
<i>FBXW7</i>	12 (4)	10	4.42.E-07	1.53E-06	0	10
<i>SMC3</i>	16 (11)	16	5.28.E-07	1.77E-06	1	17
<i>CTCF</i>	12 (9)	12	5.89.E-07	1.91E-06	7	19
<i>LUC7L2</i>	9 (2)	6	6.99.E-07	2.20E-06	45	49
<i>FLT3</i>	14 (11)	11	9.49.E-07	2.90.E-06	0	11
<i>GATA2</i>	9 (4)	7	2.90.E-06	8.62.E-06	3	10
<i>GPRC5A</i>	7 (0)	7	1.54.E-05	4.45.E-05	15	22
<i>PIGA</i>	6 (2)	4	1.60.E-05	4.50.E-05	2	6
<i>FANCL</i>	7 (0)	7	4.58.E-05	1.25.E-04	0	7
<i>GNAS</i>	11 (11)	11	9.30.E-05	2.48.E-04	0	11
<i>IRF1</i>	6 (1)	6	1.04.E-04	2.70.E-04	99	101
<i>U2AF2</i>	7 (3)	7	1.58.E-04	4.01.E-04	0	7
<i>DCLRE1C</i>	8 (2)	8	3.62.E-04	8.96.E-04	0	8
<i>GNB1</i>	5 (5)	5	9.75.E-04	2.36.E-03	2	7
<i>NCOR2</i>	17 (14)	17	1.05.E-03	2.49.E-03	1	18
<i>SMC1A</i>	10 (9)	10	1.85.E-03	4.26.E-03	4	14
<i>KIT</i>	8 (8)	8	3.84.E-03	8.69.E-03	0	8
<i>SF1</i>	6 (6)	5	0.004	0.009	1	6
<i>WT1</i>	5 (2)	5	0.006	0.012	0	5
<i>KDM6A</i>	9 (1)	9	0.013	0.027	4	13
<i>MRE11A</i>	5 (0)	5	0.027	0.055	7	12
<i>DNMT3B</i>	5 (4)	5	0.052	0.107	18	23
<i>PRPF8</i>	11 (10)	11	0.058	0.116	17	28
<i>SRRM1</i>	5 (2)	5	0.061	0.120	5	10
<i>KDR</i>	7 (7)	7	0.063	0.121	2	9
<i>BCORL1</i>	8 (3)	8	0.067	0.127	4	12
<i>PTPRD</i>	9 (6)	9	0.080	0.148	0	9
<i>PDS5B</i>	7 (0)	7	0.086	0.158	10	17
<i>BRAF</i>	4 (4)	4	0.093	0.167	0	4
<i>FANCA</i>	7 (3)	7	0.096	0.169	6	13
<i>ALAS2</i>	3 (2)	3	0.101	0.176	4	7
<i>NCOR1</i>	10 (9)	10	0.132	0.225	13	23
<i>TP53BP1</i>	8 (6)	8	0.144	0.241	1	9

<i>ABCG8</i>	3 (1)	3	0.153	0.253	0	3
<i>FANCM</i>	8 (3)	8	0.159	0.259	1	9
<i>CDCA5</i>	1 (0)	1	0.186	0.297	1	2
<i>STAG3</i>	5 (3)	5	0.193	0.304	0	5
<i>ARID1B</i>	8 (4)	8	0.220	0.342	1	9
<i>SF3A1</i>	3 (3)	3	0.228	0.348	0	3
<i>NOTCH1</i>	9 (8)	9	0.238	0.359	8	17
<i>SLC25A38</i>	1 (0)	1	0.243	0.361	5	6
<i>PRPF40B</i>	3 (3)	3	0.298	0.430	0	3
<i>PTPRT</i>	5 (4)	5	0.297	0.435	30	34
<i>ATRX</i>	8 (6)	8	0.332	0.460	4	11
<i>SF3B14</i>	0 (0)	0	0.331	0.465	0	0
<i>DOTIL</i>	5 (5)	5	0.327	0.466	0	5
<i>GRID1</i>	3 (3)	3	0.362	0.495	1	4
<i>FANCD2</i>	4 (4)	4	0.489	0.628	11	15
<i>SF3A3</i>	1 (1)	1	0.480	0.631	2	3
<i>FANCE</i>	1 (0)	1	0.486	0.632	1	2
<i>PTTG1</i>	0 (0)	0	0.474	0.632	0	0
<i>FANCB</i>	2 (0)	2	0.469	0.633	2	4
<i>ESCO2</i>	1 (0)	1	0.546	0.692	1	2
<i>ELANE</i>	0 (0)	0	0.558	0.699	1	1
<i>FANCG</i>	1 (1)	1	0.573	0.709	1	2
<i>CRKL</i>	0 (0)	0	0.591	0.724	0	0
<i>SETD2</i>	6 (4)	6	0.622	0.752	4	10
<i>PHF8</i>	2 (2)	2	0.631	0.755	4	6
<i>ASH1L</i>	7 (7)	7	0.642	0.758	1	8
<i>FANCF</i>	0 (0)	0	0.668	0.781	0	0
<i>HIF1A</i>	1 (0)	1	0.715	0.817	1	2
<i>GATA1</i>	0 (0)	0	0.711	0.821	4	4
<i>SMC1B</i>	2 (2)	2	0.728	0.823	0	2
<i>STAG1</i>	2 (2)	2	0.755	0.844	6	8
<i>CBLB</i>	1 (1)	1	0.801	0.877	0	1
<i>MLH3</i>	2 (1)	2	0.798	0.883	1	3
<i>SMAD4</i>	0 (0)	0	0.818	0.886	5	5

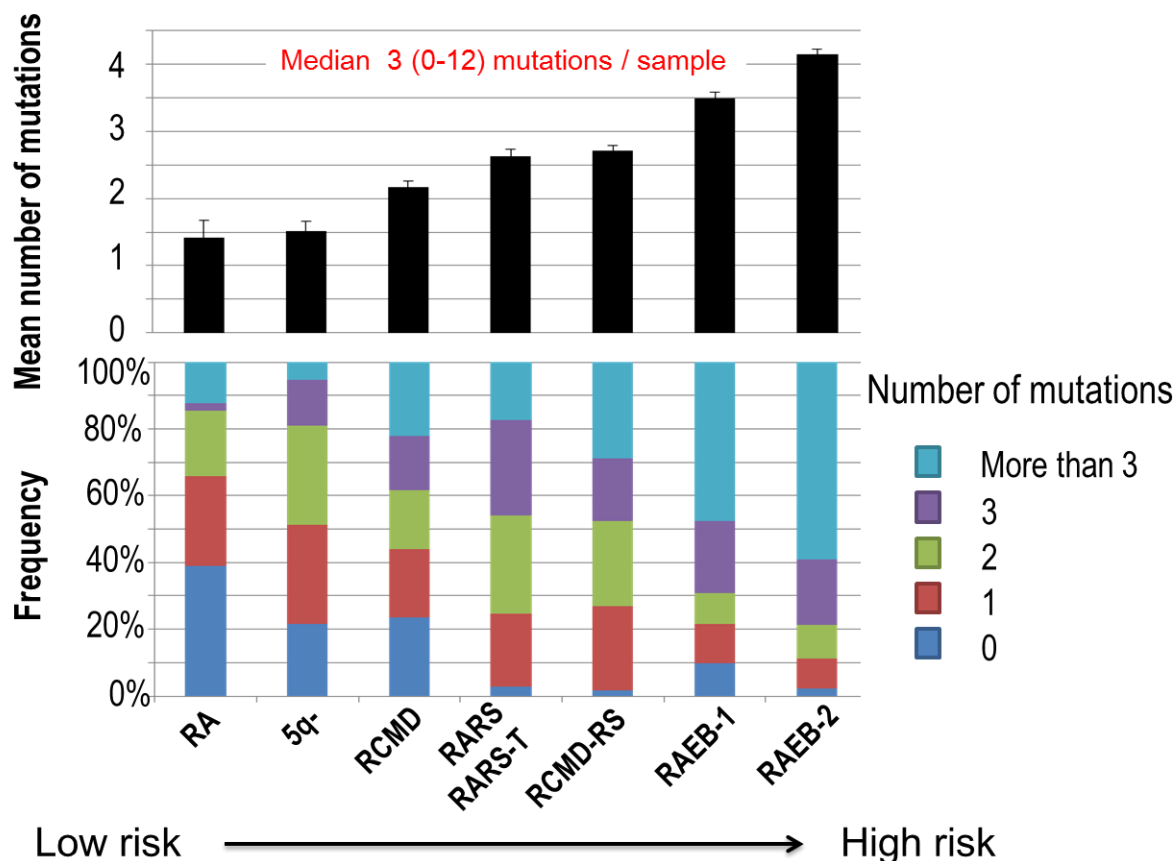
<i>NOTCH2</i>	4 (4)	4	0.859	0.921	1	5
<i>WAPAL</i>	1 (1)	1	0.873	0.927	0	1
<i>NIPBL</i>	4 (4)	4	0.925	0.943	1	5
<i>FANCI</i>	1 (0)	1	0.922	0.949	0	1
<i>MSH6</i>	1 (0)	1	0.906	0.952	0	1
<i>BRCA1</i>	2 (2)	2	0.918	0.955	1	3
<i>RFC1</i>	0 (0)	0	0.971	0.971	0	0
<i>MED12</i>	2 (1)	2	0.962	0.972	4	6
Total	2764 (1452)	2401			589	2949

図 3. パスウェイごとの変異頻度



変異遺伝子をそれぞれの機能的なパスウェイに分類し、変異頻度を比較した。縦軸にパスウェイ、横軸にパスウェイごとの変異頻度を示し、それぞれの病型ごとに色分けし記載をする。

図 4. WHO 分類ごとの平均変異個数とその分布



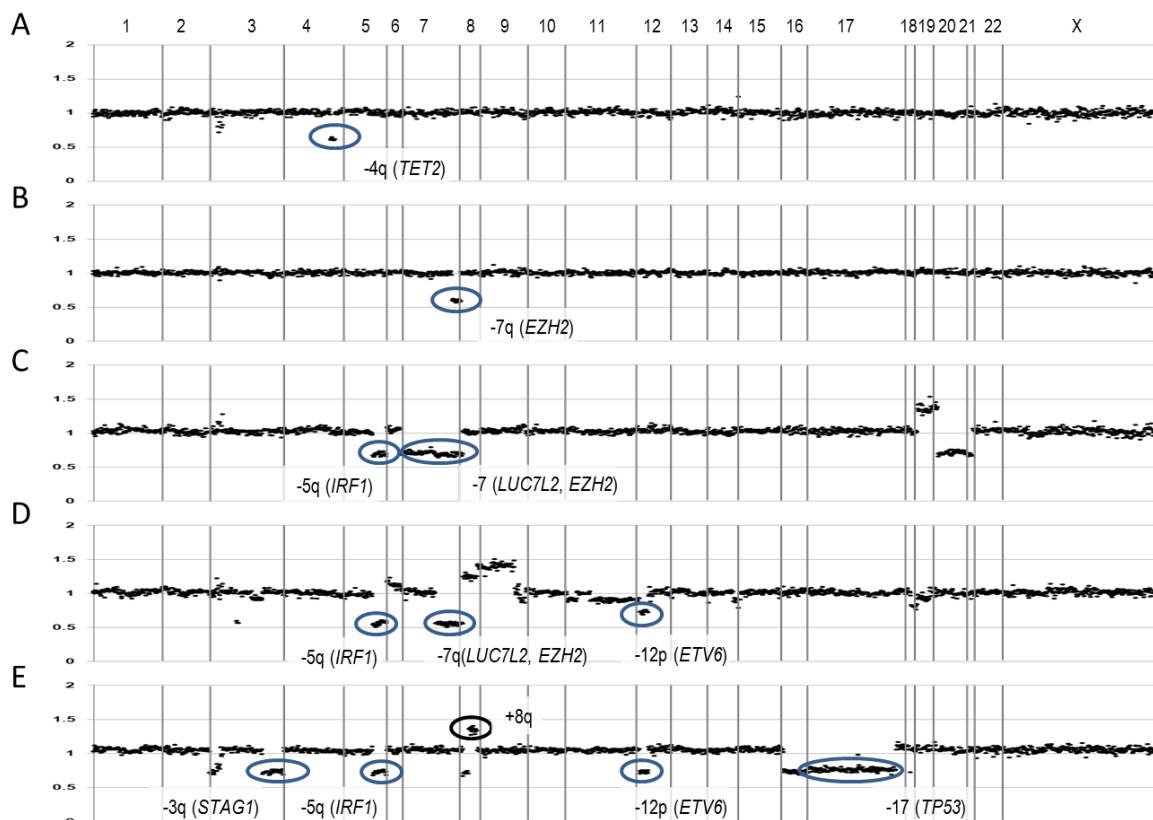
横軸に WHO 分類を低リスクから高リスクの順に記載する。

上段：縦軸に病型ごとの 1 症例あたりの平均変異個数を示す。エラーバーは標準誤差を表している。

下段：縦軸に病型ごとの平均変異数の分布割合を示す。

低リスクから高リスクになるにつれ、平均変異個数が増加し、変異個数を多く有する症例の割合も増えていくことが確認された。

図5. 標的シーケンスによるコピー数異常の解析



代表的な5症例について標的シーケンスによるコピー数解析の結果を表す。縦軸にコピー数、横軸に染色体を示す。

症例Aは4番染色体長腕(4q22)の欠失(*TET2*遺伝子)が認められた。

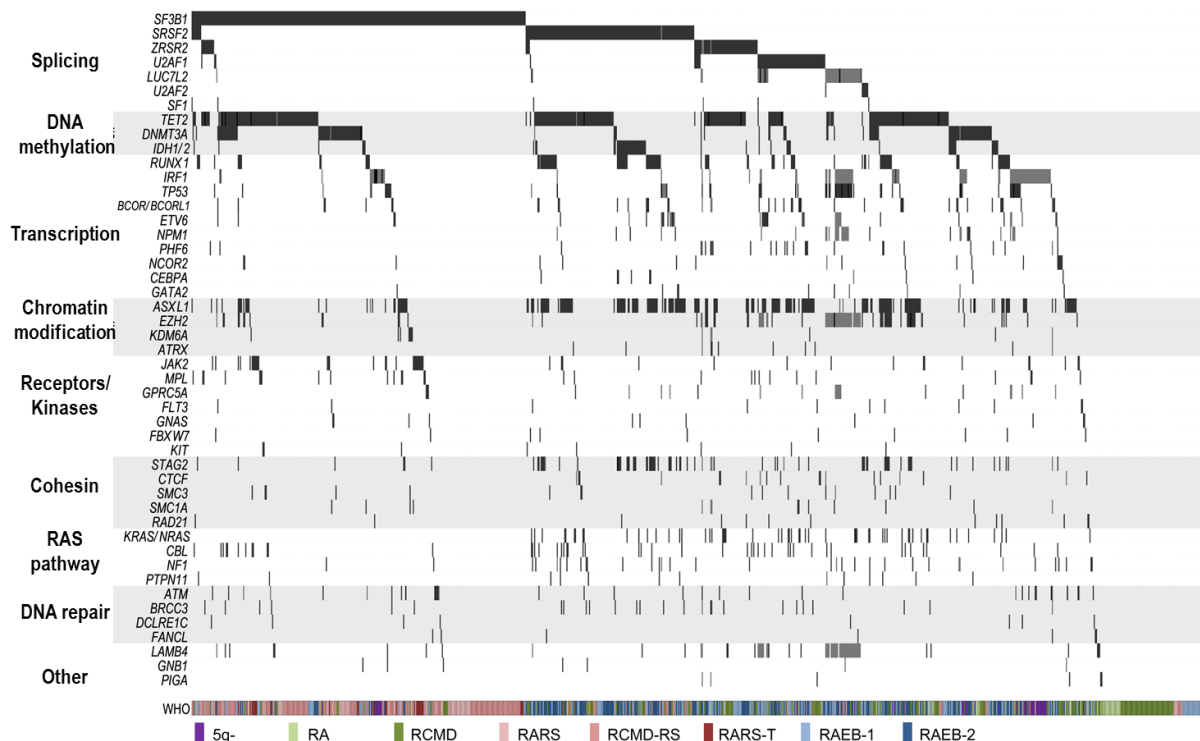
症例Bは7番染色体長腕(7q)の欠失(*EZH2*遺伝子)が認められた。

症例Cは5番染色体長腕(5q)(*IRF1*遺伝子)と7番染色体の欠失(*LUC7L2*遺伝子、*EZH2*遺伝子)が認められた。

症例Dは5番染色体長腕(5q)(*IRF1*遺伝子)、7番染色体(*LUC7L2*遺伝子、*EZH2*遺伝子)と12番染色体短腕(12p)の欠失(*ETV6*遺伝子)が認められた。

症例Eは3番染色体長腕(3q)(*STAG1*遺伝子)、5番染色体長腕(5q)(*IRF1*遺伝子)12番染色体短腕(12p)(*ETV6*遺伝子)と17番染色体の欠失(*TP53*遺伝子)、8番染色体長腕(8q)の増幅が認められた。

図 6. 変異と欠失の分布



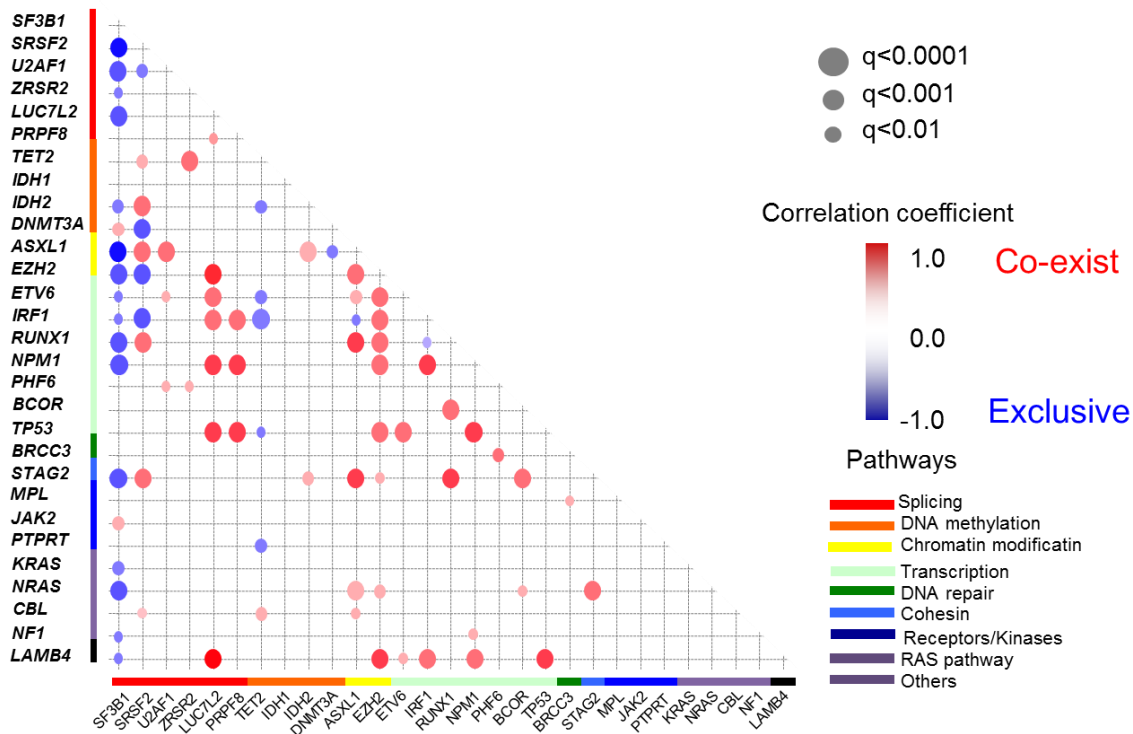
全 944 症例における変異・欠失の分布を示す。縦軸に変異遺伝子、横軸に症例を示し、黒の濃淡にて淡い順番にそれぞれ、欠失、変異もしくはその両方を表す。最下段に症例ごとの WHO 分類を色分けし記載をする。

これらはパスウェイごとの変異頻度順に並べられており、RNA スプライシングの遺伝子異常は既報のようにそれぞれ排他的な関係になっていることが確認された。

(3) 遺伝子変異の相関

有意な変異遺伝子 47 個について遺伝子変異の相関を調べたところ、82 個の組み合わせが統計学的に有意 ($q < 0.01$) であることが判明した。このことは MDS の病態においてこれらの遺伝子変異はランダムに起きているのではなく、機能的な相互作用を有する可能性が示唆された。*STAG2, IDH2, ASXL1, RUNX1, BCOR* の変異/欠失が共存関係にあることだけでなく、*SRSF2* 変異と *DNMT3A, EZH2, IRF1* の変異/欠失や *ASXL1* と *DNMT3A* 変異/欠失が排他関係にあることも示された。(図 7) *SF3B1* 変異症例は変異数が少ないことを反映しているのかもしれないが、*SF3B1* 変異は *JAK2* と *DNMT3A* を除き他の主要な標的遺伝子変異とは排他関係にあった。*ASXL1* 変異は *SF3B1, DNMT3A* と *IRF1* を除き他の変異と共存関係になることが多く、*TET2* 変異は *SRSF2* や *ZRSR2* 変異と共存関係になることが多いことも示された。

図 7. 主要な標的遺伝子変異の相関



縦軸、横軸とも有意に変異している遺伝子であり、それぞれの変異・欠失の相関について示す。円の大きさは相関関係の有意性、赤色は共存関係、青色は排他関係で色の濃淡でそれぞれの関連の程度を表している。またそれぞれの軸の内側に、遺伝子が分類されるパスウェイが異なる色で記載される。

スプライシング遺伝子変異は互いに排他関係であることが確認され、それ以外にも報告のない新たな関連性が見出された。

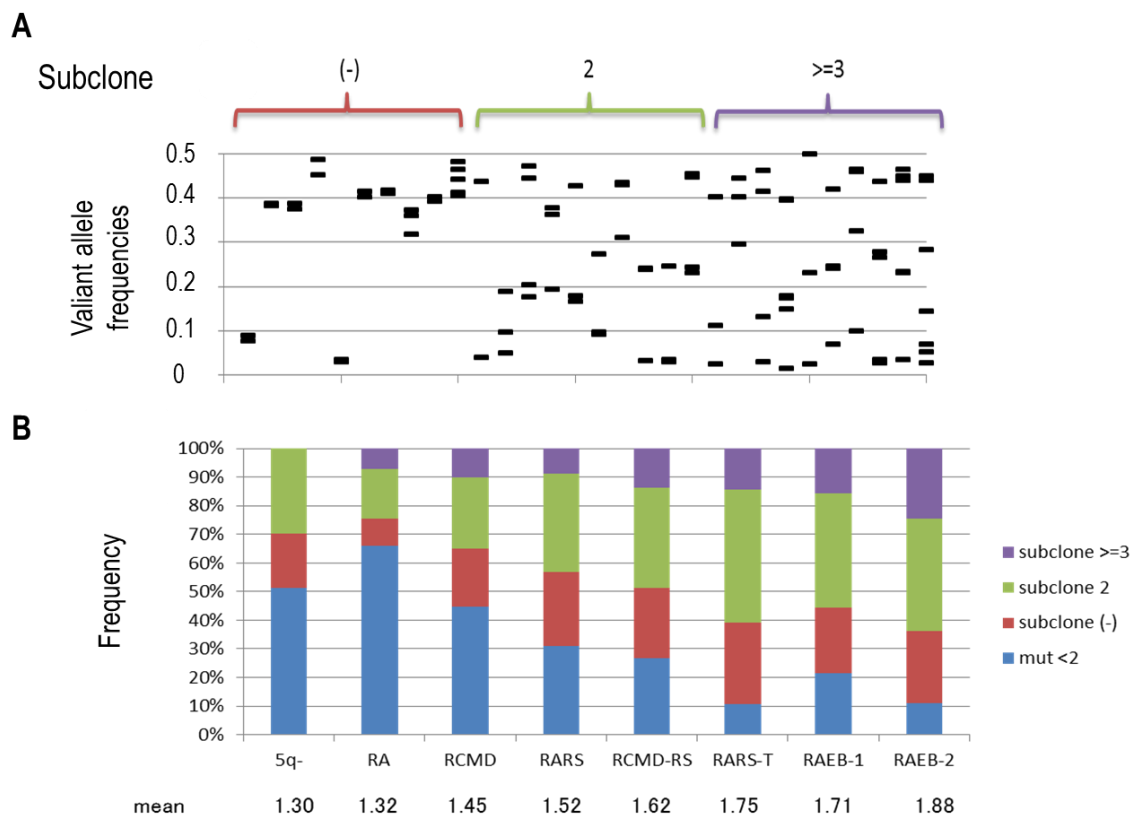
(4) 腫瘍内異質性の検定と変異アレル頻度の比較

腫瘍の発達には機能的な遺伝子変異がクローン拡大を起こす最初のヒットになり、そこに次々と他の遺伝子変異が蓄積することでより増殖が起こるモデルが示唆されている。(67-69) MDS から急性骨髄性白血病へ移行する際にクローンが拡大することが示されており(36)、それは化学療法後に急性骨髄性白血病が再発するときにも同様の報告がある(70)。変異アレル頻度よりその変異を有する腫瘍細胞比率を見積もることが可能となり、それぞれの遺伝子変異が、すべての細胞が有するメジャークローンに属するのか、あるいは一部の細胞のみが有するサブクローンに属するのか決定を行い、腫瘍内に異質性があるのか検討を行った(71)。変異が2個以上存在する症例を解析の対象とし、ディープシーケンスによる正確な変異アレル頻度と改変したカイ二乗検定法を用いて遺伝子変異の観点から腫瘍内に異質性が存在するかどうか検討を行った。解析対象遺伝子が少数のため、真の値より過小評価になっているものの、456症例(48.3%)については腫瘍内異質性ありと判定された。観察されたサブクローン数は同定された変異数に相関する傾向が認められ、進行したWHO分類や予後不良なグループでそれが高頻度に認められた。(図8)

平均の変異アレル頻度は主要な遺伝子においても様々なバリエーションが認められた。たとえば、RARS-T症例は*SF3B1*変異と*JAK2*変異を共有することが知られて

いるが、一例を除き *SF3B1* 変異は共存する *JAK2* 変異と比較すると有意にアレル頻度は高かった。(図 9A) このことは RARS や RCMD-RS などの環状鉄芽球を有する病型で *SF3B1* 変異が認められ、その後、*JAK2* 変異を新たに獲得して RARS-T が発症するモデルが強く示唆された。パスウェイ単位で考えると、RNA スプライシングや DNA メチル化は他のパスウェイと比較すると有意にアレル頻度が高く (図 9B)、これらの変異は腫瘍化の早期に起きている可能性が示唆された。クロマチン修飾や転写は RAS パスウェイよりも高いアレル頻度であった。有意な変異遺伝子 47 個もしくは標的パスウェイ 9 個に関してそれぞれ遺伝子ごと、パスウェイごとに共存している変異のアレル頻度を比較したところ、有意にアレル頻度が異なる組み合わせは 28 遺伝子ペア、18 パスウェイペアであった。(図 9C-D) これは腫瘍がクローナルな増殖を行う中で、主要な遺伝子変異に序列が存在する、すなわち絶対ではないもの遺伝子変異が起きる順番にはある程度の傾向が存在することが示唆された。(図 10)

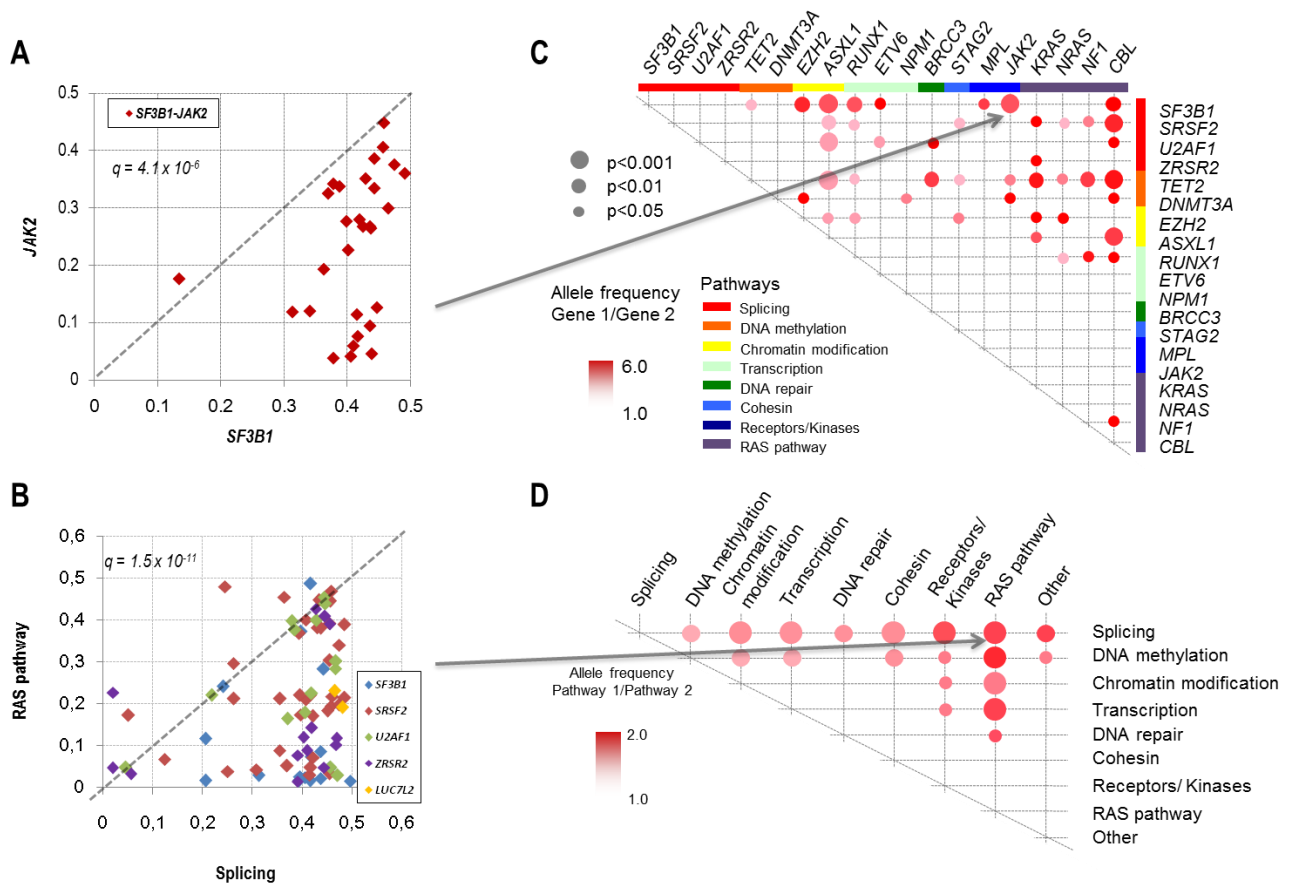
図 8. 腫瘍内異質性の検定



A は腫瘍内異質性の検定の結果、異質性なし、異質性ありでサブクローン数 2 個、異質性ありでサブクローン数 3 個以上、と判定された代表的な症例を 10 例ずつ示したもので、縦軸に変異を有する腫瘍細胞の割合、横軸に症例を示している。

B は横軸に WHO 分類を低リスクから高リスクの順になるように記載し、それぞれの分類内で異質性検定結果の分布を示している。また、平均のサブクローン数を WHO 分類の下に表す。高リスクの病型では、異質性を有する頻度が高く、サブクローン数も多かった。

図 9. 腫瘍アレル頻度の比較

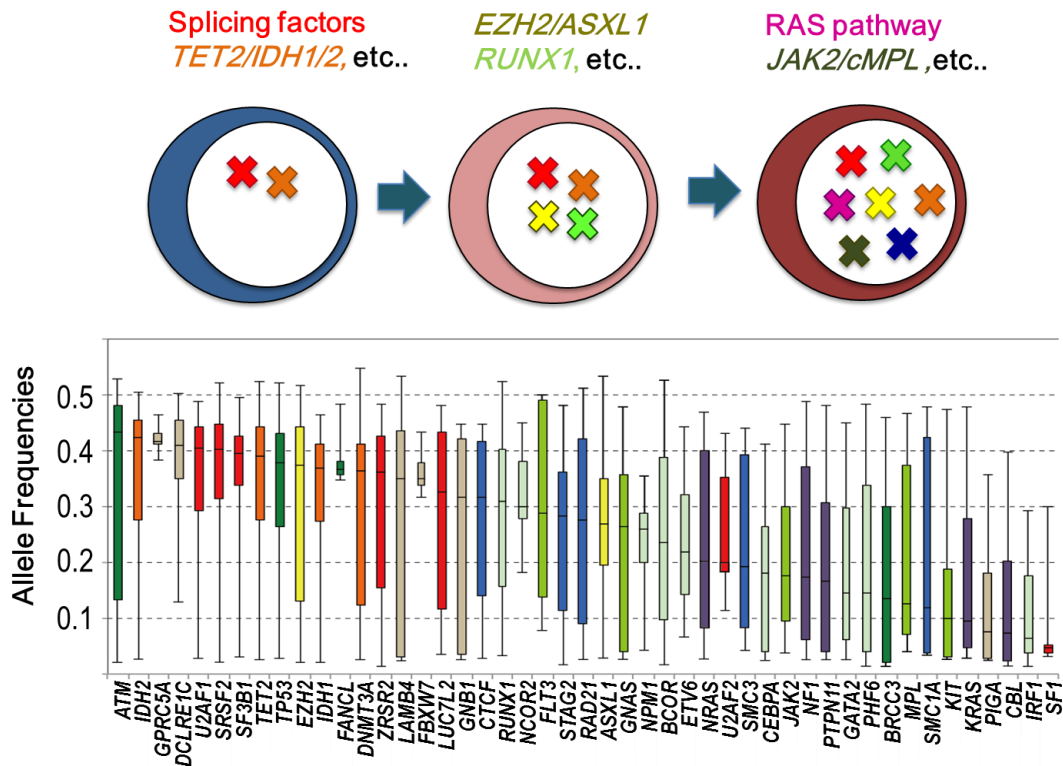


A は *SF3B1* 変異と *JAK2* 変異を共存している症例においてそれぞれ腫瘍アレル頻度を対角プロットにて示したものである。*SF3B1* 変異は *JAK2* 変異と比較して有意にアレル頻度が高いことが確認された。

B はパスウェイごとに考慮し、スプライシングと RAS パスウェイとが共存している症例においてそれぞれのアレル頻度を対角プロットにて示しており、スプライシング遺伝子をそれぞれ色分けして記載する。スプライシングは RAS パスウェイと比較して有意にアレル頻度が高いことが確認された。

C、D は主要な遺伝子ごと、パスウェイごとにそれぞれの平均のアレル頻度が比較された結果を表す。いずれも円の大きさはアレル頻度差の有意性、色はアレル頻度の相対比率を濃淡で示している。

図 10. 想定される遺伝子変異の序列



上段は想定される遺伝子変異の模範的な序列を図に表したもので、低リスク MDS の段階で初期に起きる遺伝子異常はスプライシングに関わる遺伝子や *TET2*, *IDH1/2* 変異のことが多く、その後、*EZH2/ASXL1* などのクロマチン修飾、*RUNX1* などの転写に関わる遺伝子変異が蓄積することでより高リスク MDS に移行し、最後に RAS パスウェイや *JAK2* 変異が起きることでクローナルな増殖が認められ、急性骨髄性白血病が引き起こされるモデルが示唆された。

下段はそれぞれの遺伝子変異の平均のアレル頻度を高い順に並べたものであり、縦軸にアレル頻度、横軸に遺伝子、パスウェイごとに色分けして表示している。

(5) 遺伝子異常を含んだ新たな予後予測モデルの作成

これらの遺伝学的な異常が臨床情報に及ぼす影響を検討するため、予後データが判明している 875 症例を対象に解析が行われた。有意な 47 の変異遺伝子に *PRPF8* を加えた 48 遺伝子について単変量解析が行われ、25 個の遺伝子が有意 ($p < 0.05$) に生存に影響を及ぼしていた。(図 11A) 唯一、*SF3B1* 変異のみが有意に予後良好群に関わっていた。全コホートの中で *SF3B1* 変異を有する症例では有意に予後の改善が認められ、変異を 1 個のみ有する 160 症例に絞った場合に置いても再現性が確認された。(図 11B) 次にこれらの遺伝子異常と IPSS-R など従来の予後予測に用いられる染色体や末梢血などの臨床学的因子を組み合わせる効果を検証するため、従来の COX 回帰により予後予測モデルが作成された。この目的のために、生存情報と染色体、末梢血情報が判明している 786 例について、611 例の探索コホートと残りの 175 例の検証コホートに分けてそれぞれ解析が行われた。想定される説明因子数を減らすために、探索コホート内で 15 患者以上に異常が認められた 32 遺伝子についてのみ回帰分析が行われ、因子選択には LASSO 法が用いられた。(方法を参照)

全体で 14 個の遺伝子と年齢、性別、IPSS-R で用いられるような末梢血パラメータ(白血球、ヘモグロビン、血小板)や悪性細胞数、染色体情報が選択され、COX の比例ハザードモデルを用いた回帰分析が行われた。(図 12A) 回帰モデルが線形予測になることに基づいて、我々は遺伝学的情報と臨床学的因子を組み合わせた新たな予

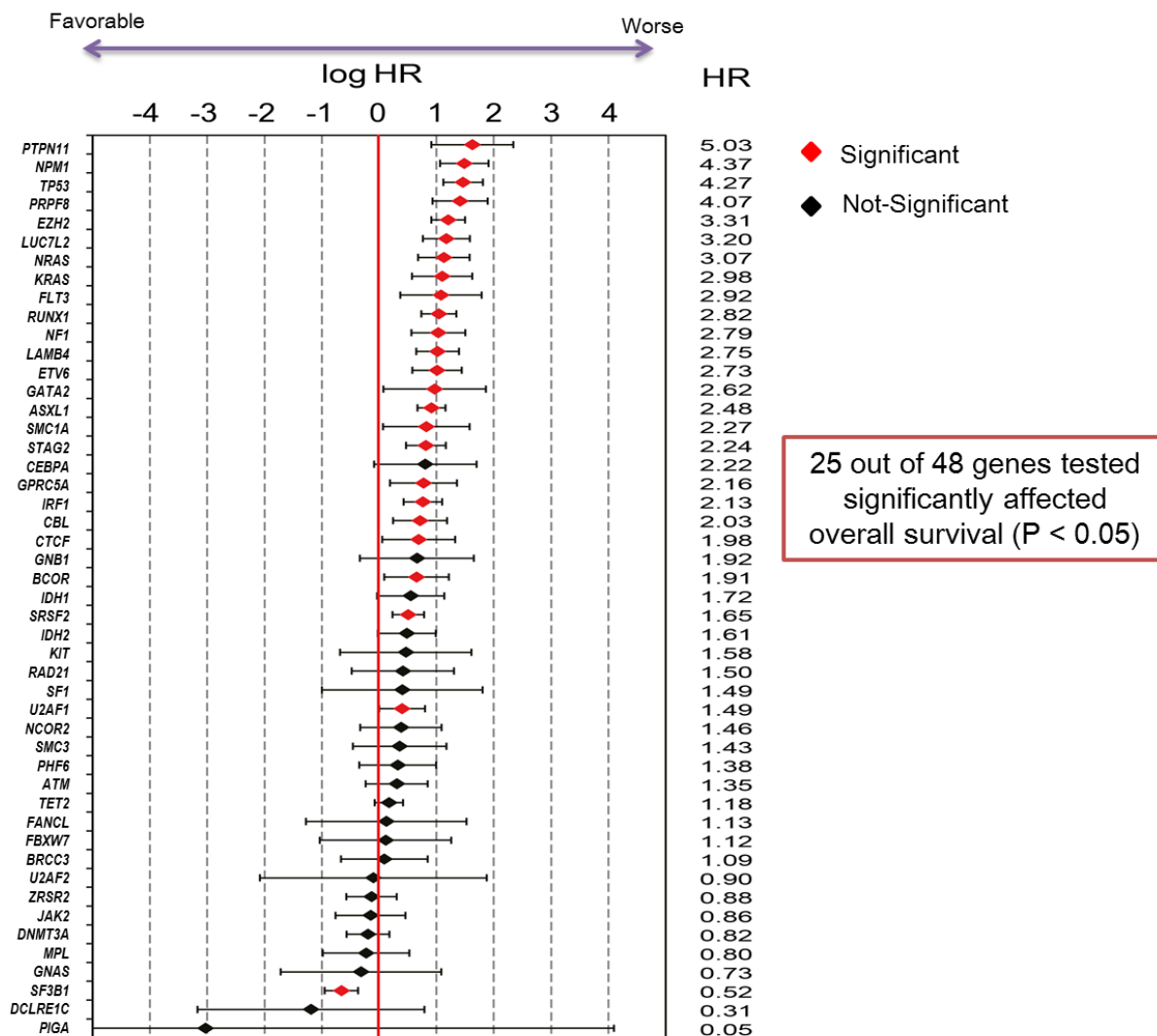
後予測モデル（モデル 1）を作成することに成功し、これは探索コホート内の患者を有意に異なる生存を示す 4 群に分離することが可能であった。（図 13A）このモデル 1 は 175 症例の検証コホートでも大部分において再現性が確認され（図 13D）、症状緩和治療のみが行われた 463 症例においてもまた同様の結果であった。また、遺伝子異常のみに着目し LASSO 法による因子選択を行ったところ、こちらも 14 個の遺伝子が選択され新たな後予測モデル（モデル 2）を作成することが可能であった。（図 12B）モデル 2 も探索コホート内の患者を有意に生存の異なる 4 群に分離可能であり（図 13B）、検証コホートにおいても統計学的に再現性が確認された（図 13E）。興味深いことに、モデル 1 で選択された 14 個の遺伝子の内、13 遺伝子については再びモデル 2 作成時においても選択された。また比較のために、探索コホート、検証コホートそれぞれにおいて従来の後予測モデルである IPSS-R による後分類の結果を示された。（図 13C,F）

これらの結果は特定遺伝子の変異/欠失は臨床学的因子から独立した後予測因子として使用可能であることを証明している。検証コホートにおいてこれらのモデルが比較されたとき、モデル 1 は IPSS-R やモデル 2 よりも優れた性能を示す可能性が示唆された。（表 4）赤池の情報量基準（AIC）はモデル 1 が最も低く、モデルの当てはまりが最もよいことを表している。また、モデル 1 の優位性は J-test によっても示

されている。すなわち、IPSS-R にモデル 1 を上乗せすることで有意 ($P<0.001$) に予後が改善するのに対し、モデル 1 に IPSS-R を上乗せしても有意な改善は認められなかった。 ($P=0.070$) 期待された結果ではあったが、遺伝子異常のみを用いたモデル 2 と IPSS-R が比較されたとき、お互いに上乗せをすることで予後が改善することが示された。これらの結果は検証コホート内で症状緩和治療のみを受けた症例に限定しても大きな相違は認められなかった。

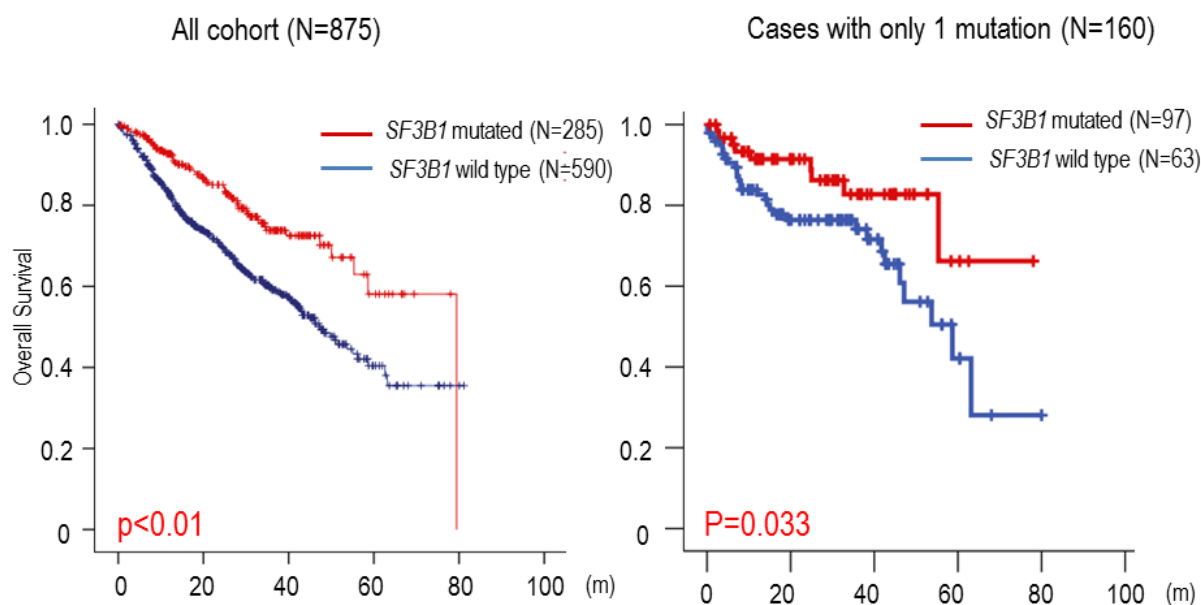
遺伝子変異と WHO 分類や年齢、性別、悪性細胞数など臨床情報についての今まで報告のない関連についても探索コホート内で調査された。 *ASXL1* 変異や *TET2* 変異は有意に男性に多く、 *TET2* 変異を有する症例は変異がない症例と比較して、高齢であった。主要な遺伝子変異の多くは高リスク群の病型もしくは悪性細胞が増加している MDS 患者に高頻度に認められた。

図 11A. 遺伝子異常の単変量解析



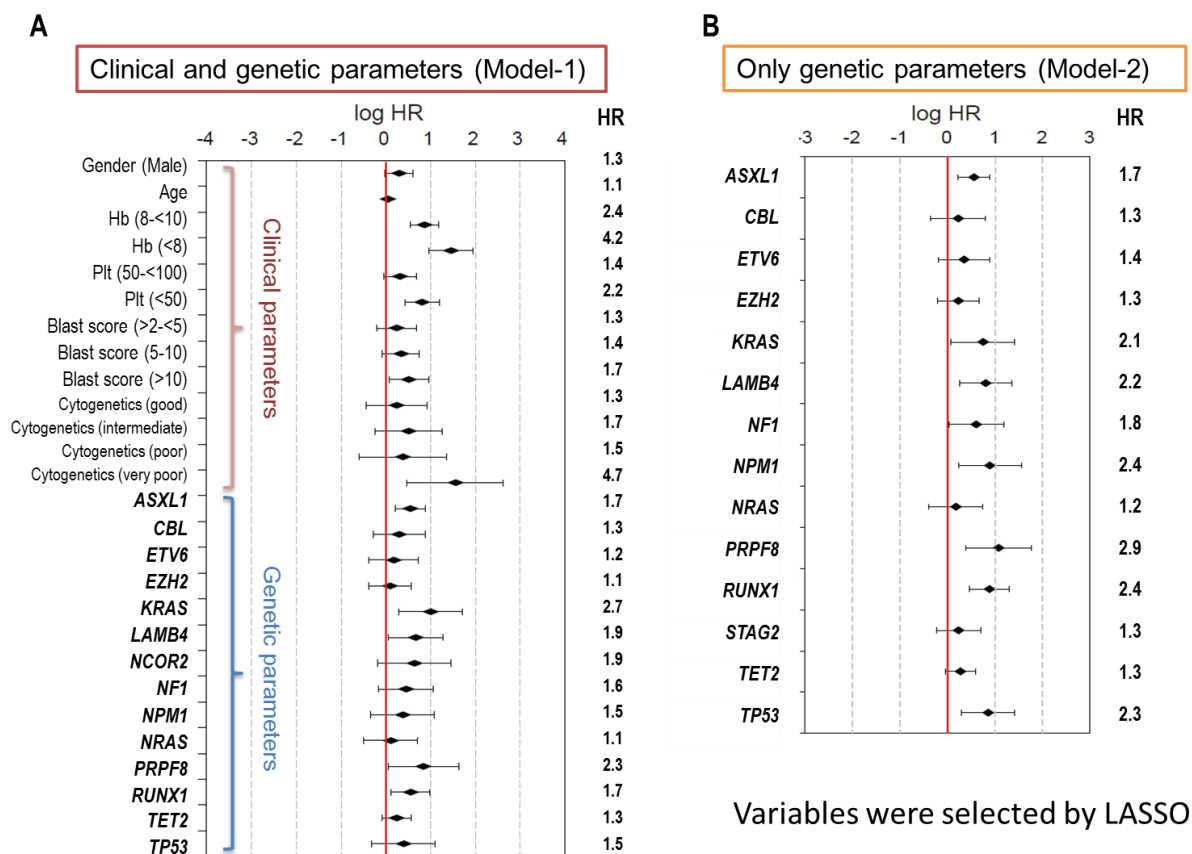
COX 回帰による単変量解析の結果を示したもので、縦軸に有意な遺伝子 47 個に *PRPF8* を加えた 48 遺伝子、横軸に自然対数ハザード比 (log HR) とその 95%信頼区間をエラーバーで表す。それぞれの遺伝子に関するハザード比 (HR) についても示す。ハザード比が高い、予後不良な因子から順番に並べ、有意な因子に関しては赤ひし形でその値を示す。最終的に 25 個の遺伝子が有意となり、予後良好を示したものは *SF3B1* 変異のみであった。

図 11B. SF3B1 変異と予後との関連



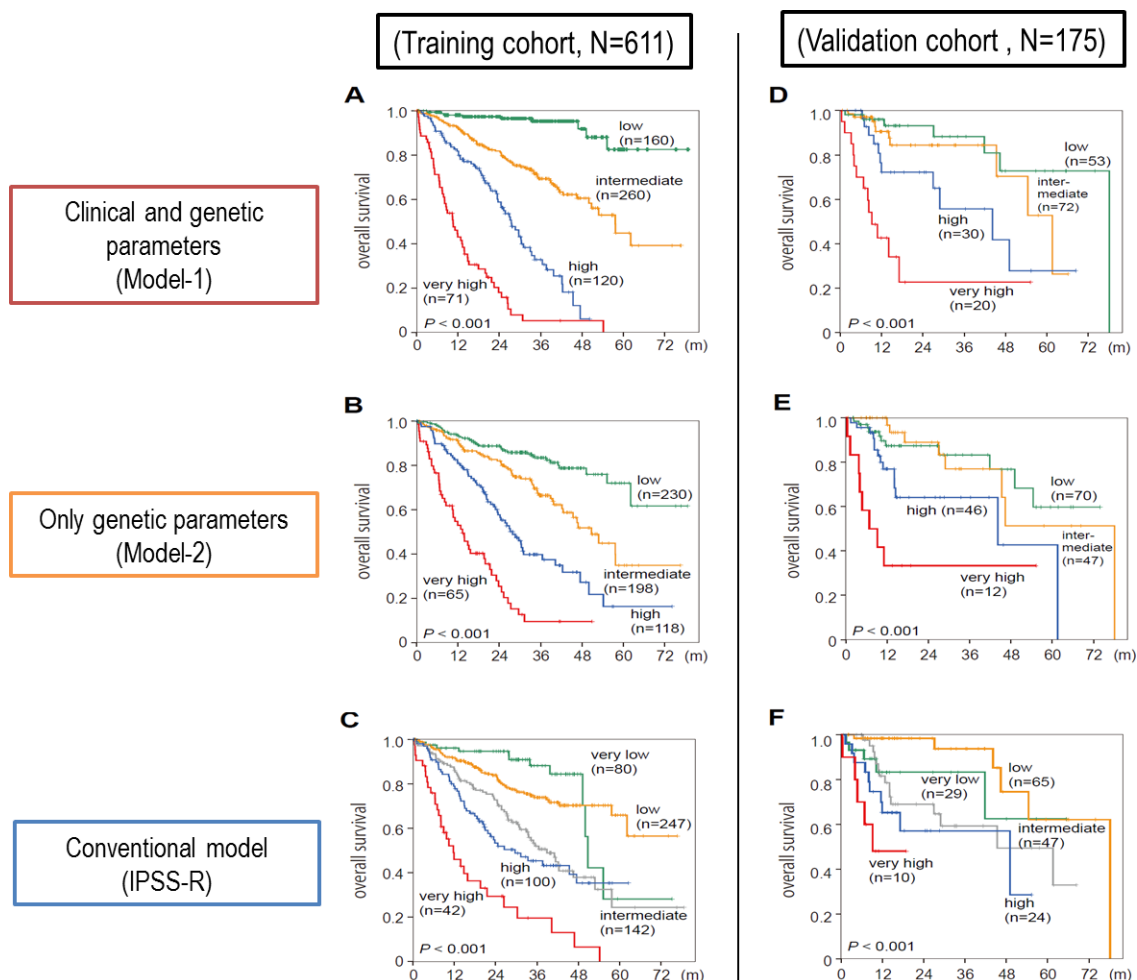
944 例の全コホートの中で予後が判明している 875 例について解析を行ったところ、*SF3B1* 変異を有する症例では有意に予後の改善が認められた。本解析結果は変異を 1 個のみ有する 160 症例に絞った場合においても再現性が確認され、*SF3B1* 変異を有する群はない群と比較し有意に予後の改善が認められた。

図 12. モデル 1 と 2 の多変量解析



A、Bそれぞれモデル1、2を作成する際にLASSO法により選択された因子におけるCOX回帰による多変量解析の結果を示したもので、それぞれの因子における自然対数ハザード比(log HR)とその95%信頼区間をエラーバーで表す。またハザード比(HR)についても示す。遺伝学的な異常と従来の臨床的予後予測因子を組み合わせたモデル1では性別、年齢、末梢血、悪性細胞数、染色体などIPSS-Rで用いられてきた因子と14個の遺伝子が予後予測に用いられた。一方、遺伝学的な異常のみに着目したモデル2においては、14個の遺伝子が選択された。興味深いことにモデル2の作成に用いられた遺伝子はモデル1作成に用いられた遺伝子とほとんど同じであった。

図 13. モデル 1, 2 と IPSS-R の探索コホート、検証コホートにおける生存曲線



モデル1,2とIPSS-Rについて探索コホートと検証コホートそれぞれにおける Kaplan-Meier法による生存曲線の比較を示す。モデル1は探索コホートにおいて生存の異なる4群に患者を分類可能であり(A)、検証コホートにおいてもやや予後良好群で分離が悪いものの有意差を認めた(D)。モデル2においても1と同様生存の異なる4群に患者を分類可能で(B)、検証コホートでも再現可能であった(E)。比較のため従来の予後予測因子であるIPSS-Rによる生存曲線も示された。(C,F)

これらの結果より、遺伝学的な異常は予後予測に有用である可能性が示唆された。検証コホート内で高リスク群を分類できた症例数はモデル1, 2, IPSS-Rでそれぞれ、20人、12人、10人であり、モデル1が最も多くの予後不良患者を抽出可能であった。

表 4. 予後予測モデルの検証コホートにおける比較

Model	Validation cohort (n=175)			Validation cohort – supportive treatment only (n=129)		
	AIC	J test		AIC	J test	
		H ₀ : Model x H ₁ : IPSS-R	H ₀ : IPSS-R H ₁ : Model x		H ₀ : Model x H ₁ : IPSS-R	H ₀ : IPSS-R H ₁ : Model x
Model-1	327.2	<i>P</i> =0.070	<i>P</i> <0.001	209.5	<i>P</i> =0.201	<i>P</i> <0.001
Model-2	337.5	<i>P</i> =0.003	<i>P</i> =0.015	212.6	<i>P</i> =0.002	<i>P</i> =0.039
IPSS-R	334.8	-	-	219.0	-	-

また、我々はモデル 1、2、IPSS-R のどのモデルがより生存を分類するのに優れているのか比較検証を行った。検証コホート内で赤池の情報量基準 (AIC) はモデル 1 が最も低く、モデルの当てはまりが最もよいことを表していた。また、モデル 1 の優位性は J-test によっても示されている。すなわち、IPSS-R にモデル 1 を上乗せすることで有意 ($P < 0.001$) に予後が改善するのに対し、モデル 1 に IPSS-R を上乗せしても改善は認められなかった ($P = 0.070$)。この結果は、症状緩和治療のみが行われた患者群でも同様の結果であった。

考察

944 症例、104 遺伝子について次世代シーケンス技術を用いてディープシーケンスを行い、ゲノムワイドなコピー数解析を組み合わせることで主要な遺伝子変異と欠失が判明し、多数症例における MDS の遺伝学的異常の展望を解明することが可能となった。コピー数異常を組み合わせると遺伝子異常は 864 症例 (91.5%) に認められた。最近異なる施設より MDS 多数症例における解析結果が報告されているが(41)、本結果はそれを上回るものであった。検証のため一部の症例ではあるものの、次世代シーケンサーの結果と他の分子診断によってもたらされた結果との間に極めて高い一致率が確認されたことは非常に重要と考えられる。標的配列全体に渡って均一に深い読み取り深度で解析できたことを考慮すると、長い欠失・挿入を除いて非常に高い感度で遺伝子変異を同定していることが期待された。このことは今まで変異報告のない遺伝子で 0.45 を超えるような高いアレル頻度を有する変異や非常に小さなサブクローンのみが存在する遺伝子変異についても真実である可能性が高いことを示唆している。現行のシーケンス機械や試薬には技術的な検出限界があり、今回の解析では正常ペア検体を使用できないため、既知の SNPs やシーケンスエラーをどんなに除外しても最終的な結果にいくつかの偽陽性が含まれる可能性は排除できない。

RA や他の低リスク MDS では遺伝子変異が少ないことが特徴的で予後良好であり、一方高リスク MDS では変異数が多いことが判明した。MDS における腫瘍内不均一性

の存在は極めて一般的であることが報告され(37)、進行した WHO 分類や予後不良群では変異数が多く腫瘍内サブクローン数も多いためその存在はより顕著となる。これらをまとめると、遺伝子変異が蓄積し、腫瘍内異質性が増加することは MDS における腫瘍クローン拡大の特徴であり、予後不良と関連している。変異序列の存在はもう一つの重要な発見であり、MDS の病態だけでなく、病勢の進行をモニターするような分子と臨床学的な関連についても理解を深める可能性が示唆された。我々の研究において同定された遺伝子変異は有意な共存性や排他性などの相関関係を持ち、たとえば、*TET2* 変異と *ZRSR2* 変異の共存性や *ASXL1* 変異と *SF3B1* 変異の排他性などこれらの多くは最近の論文によっても同様の報告がされている。(41) また、我々の研究室ではメジャーなスプライシング変異遺伝子の一つである *U2AF1* に変異を導入した変異体を細胞にトランスフェクションさせ、実際に変異の有無でスプライシング機構の異常が出現するのか検討を行った。RNA シークエンスにより、エクソンスキッピングや余計なイントロンが挿入されるなどスプライシング機構の異常が観察されたが、特定の遺伝子にこれらの異常が蓄積される結果は認められなかった。スプライシング遺伝子異常がスプライシング機構の異常を惹起することは同定されたが、その異常を通して直接腫瘍化に関わっているかどうかは現時点では不明である。解析コホートの 75%の症例で症状緩和治療のみが選択されており、化学療法やメチル化阻害剤など重

点的な治療が行われているのは約 15%である。*SF3B1* 変異は症状緩和治療が施行されている群で高頻度であり、*ASXL1,STAG2,SRSF2* や *RUNX1* 変異は重点的な治療が行われている群で高頻度であった。免疫抑制療法単独が施行されている症例は本研究には含まれていなかった。

しかしながら、本研究の最も重要な発見は 100 を超える多数の遺伝子解析が実現可能となり(41)、そのことでより鋭敏で正確な MDS 患者の予後層別化が図られ臨床判断の大きな手助けとなることである。実際に複数の遺伝子変異と予後との関連は非常に強く、今までにない新たな予後予測モデルの構築が可能となった。14 個の予後予測遺伝子と年齢、性別、IPSS-R に用いられる染色体情報など従来用いられている臨床的予後予測因子を組み合わせることにより新たな予後予測モデル(モデル 1)を作成し、これは MDS 患者を生存の異なる 4 群に分類することが可能であった。このモデルは遺伝子のみで作成したモデルや IPSS-R と比較してより鋭敏で優れた予後予測能を示した。具体的にモデル 1 は、MDS 患者を良好、中間、不良、超不良の 4 群に分類可能であり、探索コホートにおける 3 年生存率はそれぞれ 95.2%、69.3%、32.8%、5.3% と有意にはっきりと予後を予測することが可能であった。(P<0.001) 我々の研究に登録された患者は探索コホートと検証コホートのいずれにおいても一部の症例においてはメチル化阻害剤など新規薬物により治療されていたが、77%以上の症例が治療と

して症状緩和療法のみが施行されていた。従来の予後予測モデルである IPSS-R は症状緩和療法のみを受けている症例において予後予測が可能であることが報告されているが、検証コホート内で症状緩和療法のみを受けていた症例においてもモデル 1 は優れた予後予測能を有することが確認できた。

結論としてディープシーケンスを用いて総合的に分子遺伝学的なプロファイルを明らかにすることは実現可能であり、MDS 患者の正確な診断、病因、生物学的な下位分類、予後予測とリスク分類を詳細に解明する非常に前途有望なアプローチである。この技術に必要なコストは劇的に低下しており、それゆえ近い将来、臨床の現場で多数の患者がこのような複数の遺伝子プロファイリングを決定することが見込まれ、MDS 患者が個別化治療を選択できる状況になることが期待される。

謝辞

本研究の全般について直接ご指導いただいた、京都大学腫瘍生物学教授（研究当時
東京大学がんゲノミクス特任准教授）小川誠司先生、患者検体を多数御提供頂きま
したミュンヘン白血病研究所の Torsten Haferlach 先生を始め、詳細な染色体分析を解
析頂きました Claudia Haferlach 先生など多数のドイツの共同研究者に心より感謝申
し上げます。次世代シーケンサーデータ解析でスーパーコンピューターを用いた
解析を御指導頂きました東京大学医科学研究所教授 宮野悟先生、予後予測モデル
の作成について御指導頂きましたミュンスター大学の Hans-Ulrich Klein 先生, Martin
Dugas 先生にも深く感謝申し上げます。また、本研究室実験助手、森有加氏、市村
静江氏、溝田紀子氏、中村真紀氏は優れた実験技術をもってサポートいただき、心
より御礼申し上げます。

参考文献

1. Brunning R, Orazi A, Germing U, Le Beau MM, Porwit A, B. I. (2008), pp. pp 87–93.
2. A. Tefferi, J. W. Vardiman, Myelodysplastic syndromes. *N Engl J Med* **361**, 1872-1885 (2009).
3. M. Cazzola, M. G. Della Porta, E. Travaglino, L. Malcovati, Classification and prognostic evaluation of myelodysplastic syndromes. *Semin Oncol* **38**, 627-634 (2011).
4. A. List, G. Dewald, J. Bennett, A. Giagounidis, A. Raza, E. Feldman, B. Powell, P. Greenberg, D. Thomas, R. Stone, C. Reeder, K. Wride, J. Patin, M. Schmidt, J. Zeldis, R. Knight, M. S.-S. Investigators, Lenalidomide in the myelodysplastic syndrome with chromosome 5q deletion. *New England Journal of Medicine* **355**, 1456-1465 (2006).
5. A. F. List, J. M. Bennett, M. A. Sekeres, B. Skikne, T. Fu, J. M. Shammo, S. D. Nimer, R. D. Knight, A. Giagounidis, Extended survival and reduced risk of AML progression in erythroid-responsive lenalidomide-treated patients with lower-risk del(5q) MDS. *Leukemia*, (2013).
6. P. Font, J. Loscertales, C. Benavente, A. Bermejo, M. Callejas, L. Garcia-Alonso, A.

- Garcia-Marcilla, S. Gil, M. Lopez-Rubio, E. Martin, C. Muñoz, P. Ricard, C. Soto, P. Balsalobre, A. Villegas, Inter-observer variance with the diagnosis of myelodysplastic syndromes (MDS) following the 2008 WHO classification. *Annals of Hematology* **92**, 19-24 (2013).
7. P. Greenberg, C. Cox, M. M. LeBeau, P. Fenaux, P. Morel, G. Sanz, M. Sanz, T. Vallespi, T. Hamblin, D. Oscier, K. Ohyashiki, K. Toyama, C. Aul, G. Mufti, J. Bennett, International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood* **89**, 2079-2088 (1997).
8. L. Malcovati, U. Germing, A. Kuendgen, M. G. Della Porta, C. Pascutto, R. Invernizzi, A. Giagounidis, B. Hildebrandt, P. Bernasconi, S. Knipp, C. Strupp, M. Lazzarino, C. Aul, M. Cazzola, Time-dependent prognostic scoring system for predicting survival and leukemic evolution in myelodysplastic syndromes. *Journal of Clinical Oncology* **25**, 3503-3510 (2007).
9. P. L. Greenberg, H. Tuechler, J. Schanz, G. Sanz, G. Garcia-Manero, F. Solé, J. M. Bennett, D. Bowen, P. Fenaux, F. Dreyfus, Revised international prognostic scoring system for myelodysplastic syndromes. *Blood* **120**, 2454-2465 (2012).

10. B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz, K. W. Kinzler, Cancer genome landscapes. *Science* **339**, 1546-1558 (2013).
11. L. A. Garraway, Genomics-driven oncology: framework for an emerging paradigm. *Journal of Clinical Oncology* **31**, 1806-1814 (2013).
12. R. Bejar, R. Levine, B. L. Ebert, Unraveling the molecular pathophysiology of myelodysplastic syndromes. *Journal of Clinical Oncology* **29**, 504-515 (2011).
13. M. Sanada, T. Suzuki, L. Y. Shih, M. Otsu, M. Kato, S. Yamazaki, A. Tamura, H. Honda, M. Sakata-Yanagimoto, K. Kumano, H. Oda, T. Yamagata, J. Takita, N. Gotoh, K. Nakazaki, N. Kawamata, M. Onodera, M. Nobuyoshi, Y. Hayashi, H. Harada, M. Kurokawa, S. Chiba, H. Mori, K. Ozawa, M. Omine, H. Hirai, H. Nakauchi, H. P. Koeffler, S. Ogawa, Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms. *Nature* **460**, 904-908 (2009).
14. F. H. Grand, C. E. Hidalgo-Curtis, T. Ernst, K. Zoi, C. Zoi, C. McGuire, S. Kreil, A. Jones, J. Score, G. Metzgeroth, D. Oscier, A. Hall, C. Brandts, H. Serve, A. Reiter, A. J. Chase, N. C. Cross, Frequent CBL mutations associated with 11q acquired uniparental disomy in myeloproliferative neoplasms. *Blood* **113**, 6182-6192 (2009).

15. R. L. Levine, M. Loriaux, B. J. Huntly, M. L. Loh, M. Beran, E. Stoffregen, R. Berger, J. J. Clark, S. G. Willis, K. T. Nguyen, N. J. Flores, E. Estey, N. Gattermann, S. Armstrong, A. T. Look, J. D. Griffin, O. A. Bernard, M. C. Heinrich, D. G. Gilliland, B. Druker, M. W. Deininger, The JAK2V617F activating mutation occurs in chronic myelomonocytic leukemia and acute myeloid leukemia, but not in acute lymphoblastic leukemia or chronic lymphocytic leukemia. *Blood* **106**, 3377-3379 (2005).
16. S. M. Langemeijer, R. P. Kuiper, M. Berends, R. Knops, M. G. Aslanyan, M. Massop, E. Stevens-Linders, P. van Hoogen, A. G. van Kessel, R. A. Raymakers, E. J. Kamping, G. E. Verhoef, E. Verburgh, A. Hagemeijer, P. Vandenberghe, T. de Witte, B. A. van der Reijden, J. H. Jansen, Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nature Genetics* **41**, 838-842 (2009).
17. F. Delhommeau, S. Dupont, V. D. Valle, C. James, S. Trannoy, A. Massé, O. Kosmider, J.-P. Le Couedic, F. Robert, A. Alberdi, Mutation in TET2 in myeloid cancers. *New England Journal of Medicine* **360**, 2289-2301 (2009).
18. S. Weissmann, T. Alpermann, V. Grossmann, A. Kowarsch, N. Nadarajah, C. Eder, F. Dicker, A. Fasan, C. Haferlach, T. Haferlach, W. Kern, S. Schnittger, A. Kohlmann,

- Landscape of TET2 mutations in acute myeloid leukemia. *Leukemia* **26**, 934-942 (2012).
19. M. Ko, Y. Huang, A. M. Jankowska, U. J. Pape, M. Tahiliani, H. S. Bandukwala, J. An, E. D. Lamperti, K. P. Koh, R. Ganetzky, X. S. Liu, L. Aravind, S. Agarwal, J. P. Maciejewski, A. Rao, Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature* **468**, 839-843 (2010).
20. M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, A. Rao, Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**, 930-935 (2009).
21. E. R. Mardis, New strategies and emerging technologies for massively parallel sequencing: applications in medical research. *Genome Medicine* **1**, 40 (2009).
22. K. Yoshida, M. Sanada, M. Kato, R. Kawahata, A. Matsubara, J. Takita, L. Shih, H. Mori, H. Koefler, S. Ogawa, A nonsense mutation of IDH1 in myelodysplastic syndromes and related disorders. *Leukemia* **25**, 184-186 (2010).
23. W. Xu, H. Yang, Y. Liu, Y. Yang, P. Wang, S. H. Kim, S. Ito, C. Yang, M. T. Xiao, L. X.

- Liu, W. Q. Jiang, J. Liu, J. Y. Zhang, B. Wang, S. Frye, Y. Zhang, Y. H. Xu, Q. Y. Lei, K. L. Guan, S. M. Zhao, Y. Xiong, Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of α -ketoglutarate-dependent dioxygenases. *Cancer Cell* **19**, 17-30 (2011).
24. M. Sasaki, C. B. Knobbe, J. C. Munger, E. F. Lind, D. Brenner, A. Brüstle, I. S. Harris, R. Holmes, A. Wakeham, J. Haight, A. You-Ten, W. Y. Li, S. Schalm, S. M. Su, C. Virtanen, G. Reifenberger, P. S. Ohashi, D. L. Barber, M. E. Figueroa, A. Melnick, J. C. Zúñiga-Pflücker, T. W. Mak, IDH1(R132H) mutation increases murine haematopoietic progenitors and alters epigenetics. *Nature* **488**, 656-659 (2012).
25. T. J. Ley, L. Ding, M. J. Walter, M. D. McLellan, T. Lamprecht, D. E. Larson, C. Kandoth, J. E. Payton, J. Baty, J. Welch, DNMT3A mutations in acute myeloid leukemia. *New England Journal of Medicine* **363**, 2424-2433 (2010).
26. M. J. Walter, L. Ding, D. Shen, J. Shao, M. Grillot, M. McLellan, R. Fulton, H. Schmidt, J. Kalicki-Veizer, M. O'Laughlin, C. Kandoth, J. Baty, P. Westervelt, J. F. DiPersio, E. R. Mardis, R. K. Wilson, T. J. Ley, T. A. Graubert, Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. *Leukemia* **25**, 1153-1158

- (2011).
27. V. Gelsi - Boyer, V. Trouplin, J. Roquain, J. Adelaïde, N. Carbuccia, B. Esterni, P. Finetti, A. Murati, C. Arnoulet, H. Zerazhi, ASXL1 mutation is associated with poor prognosis and acute transformation in chronic myelomonocytic leukaemia. *British journal of haematology* **151**, 365-375 (2010).
28. U. H. Park, S. K. Yoon, T. Park, E. J. Kim, S. J. Um, Additional sex comb-like (ASXL) proteins 1 and 2 play opposite roles in adipogenesis via reciprocal regulation of peroxisome proliferator-activated receptor γ . *The Journal of Biological Chemistry* **286**, 1354-1363 (2011).
29. C. L. Fisher, N. Pineault, C. Brookes, C. D. Helgason, H. Ohta, C. Bodner, J. L. Hess, R. K. Humphries, H. W. Brock, Loss-of-function Additional sex combs like 1 mutations disrupt hematopoiesis but do not cause severe myelodysplasia or leukemia. *Blood* **115**, 38-46 (2010).
30. G. Nikoloski, S. M. Langemeijer, R. P. Kuiper, R. Knops, M. Massop, E. R. Tönnissen, A. van der Heijden, T. N. Scheele, P. Vandenberghe, T. de Witte, Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. *Nature*

genetics **42**, 665-667 (2010).

31. T. Ernst, A. J. Chase, J. Score, C. E. Hidalgo-Curtis, C. Bryant, A. V. Jones, K. Waghorn, K. Zoi, F. M. Ross, A. Reiter, Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nature genetics* **42**, 722-726 (2010).
32. K. Yoshida, M. Sanada, Y. Shiraishi, D. Nowak, Y. Nagata, R. Yamamoto, Y. Sato, A. Sato-Otsubo, A. Kon, M. Nagasaki, G. Chalkidis, Y. Suzuki, M. Shiosaka, R. Kawahata, T. Yamaguchi, M. Otsu, N. Obara, M. Sakata-Yanagimoto, K. Ishiyama, H. Mori, F. Nolte, W. K. Hofmann, S. Miyawaki, S. Sugano, C. Haferlach, H. P. Koefler, L. Y. Shih, T. Haferlach, S. Chiba, H. Nakauchi, S. Miyano, S. Ogawa, Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64-69 (2011).
33. E. Papaemmanuil, M. Cazzola, J. Boulton, L. Malcovati, P. Vyas, D. Bowen, A. Pellagatti, J. S. Wainscoat, E. Hellstrom-Lindberg, C. Gambacorti-Passerini, A. L. Godfrey, I. Rapado, A. Cvejic, R. Rance, C. McGee, P. Ellis, L. J. Mudie, P. J. Stephens, S. McLaren, C. E. Massie, P. S. Tarpey, I. Varela, S. Nik-Zainal, H. R.

- Davies, A. Shlien, D. Jones, K. Raine, J. Hinton, A. P. Butler, J. W. Teague, E. J.
- Baxter, J. Score, A. Galli, M. G. Della Porta, E. Travaglino, M. Groves, S. Tauro, N. C.
- Munshi, K. C. Anderson, A. El-Naggar, A. Fischer, V. Mustonen, A. J. Warren, N. C.
- Cross, A. R. Green, P. A. Futreal, M. R. Stratton, P. J. Campbell, C. M. D. W. G. o. t. I.
C. G. Consortium, Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts.
New England Journal of Medicine **365**, 1384-1395 (2011).
34. T. A. Graubert, D. Shen, L. Ding, T. Okeyo-Owuor, C. L. Lunn, J. Shao, K. Krysiak, C. C. Harris, D. C. Koboldt, D. E. Larson, Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nature genetics* **44**, 53-57 (2011).
35. R. Karni, E. de Stanchina, S. W. Lowe, R. Sinha, D. Mu, A. R. Krainer, The gene encoding the splicing factor SF2/ASF is a proto-oncogene. *Nature Structural & Molecular Biology* **14**, 185-193 (2007).
36. M. J. Walter, D. Shen, L. Ding, J. Shao, D. C. Koboldt, K. Chen, D. E. Larson, M. D. McLellan, D. Dooling, R. Abbott, Clonal architecture of secondary acute myeloid leukemia. *New England Journal of Medicine* **366**, 1090-1098 (2012).
37. M. J. Walter, D. Shen, J. Shao, L. Ding, B. S. White, C. Kandoth, C. A. Miller, B. Niu,

- M. D. McLellan, N. D. Dees, R. Fulton, K. Elliot, S. Heath, M. Grillot, P. Westervelt, D. C. Link, J. F. DiPersio, E. Mardis, T. J. Ley, R. K. Wilson, T. A. Graubert, Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. *Leukemia* **27**, 1275-1282 (2013).
38. H. Makishima, K. Yoshida, N. Nguyen, B. Przychodzen, M. Sanada, Y. Okuno, K. P. Ng, K. O. Gudmundsson, B. A. Vishwakarma, A. Jerez, I. Gomez-Segui, M. Takahashi, Y. Shiraishi, Y. Nagata, K. Guinta, H. Mori, M. A. Sekeres, K. Chiba, H. Tanaka, H. Muramatsu, H. Sakaguchi, R. L. Paquette, M. A. McDevitt, S. Kojima, Y. Sauntharajah, S. Miyano, L. Y. Shih, Y. Du, S. Ogawa, J. P. Maciejewski, Somatic SETBP1 mutations in myeloid malignancies. *Nature Genetics* **45**, 942-946 (2013).
39. A. Kon, L. Y. Shih, M. Minamino, M. Sanada, Y. Shiraishi, Y. Nagata, K. Yoshida, Y. Okuno, M. Bando, R. Nakato, S. Ishikawa, A. Sato-Otsubo, G. Nagae, A. Nishimoto, C. Haferlach, D. Nowak, Y. Sato, T. Alpermann, M. Nagasaki, T. Shimamura, H. Tanaka, K. Chiba, R. Yamamoto, T. Yamaguchi, M. Otsu, N. Obara, M. Sakata-Yanagimoto, T. Nakamaki, K. Ishiyama, F. Nolte, W. K. Hofmann, S. Miyawaki, S. Chiba, H. Mori, H. Nakauchi, H. P. Koefler, H. Aburatani, T. Haferlach, K. Shirahige, S. Miyano, S. Ogawa, Recurrent mutations in multiple components of

- the cohesin complex in myeloid neoplasms. *Nature Genetics* **45**, 1232-1237 (2013).
40. R. Piazza, S. Valletta, N. Winkelmann, S. Redaelli, R. Spinelli, A. Pirola, L. Antolini, L. Mologni, C. Donadoni, E. Papaemmanuil, S. Schnittger, D. W. Kim, J. Boultonwood, F. Rossi, G. Gaipa, G. P. De Martini, P. F. di Celle, H. G. Jang, V. Fantin, G. R. Bignell, V. Magistrini, T. Haferlach, E. M. Pogliani, P. J. Campbell, A. J. Chase, W. J. Tapper, N. C. Cross, C. Gambacorti-Passerini, Recurrent SETBP1 mutations in atypical chronic myeloid leukemia. *Nature Genetics* **45**, 18-24 (2013).
41. E. Papaemmanuil, M. Gerstung, L. Malcovati, S. Tauro, G. Gundem, P. Van Loo, C. J. Yoon, P. Ellis, D. C. Wedge, A. Pellagatti, Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* **122**, 3616-3627 (2013).
42. R. Bejar, K. Stevenson, O. Abdel-Wahab, N. Galili, B. Nilsson, G. Garcia-Manero, H. Kantarjian, A. Raza, R. L. Levine, D. Neuberg, Clinical effect of point mutations in myelodysplastic syndromes. *New England Journal of Medicine* **364**, 2496-2506 (2011).
43. R. Bejar, K. E. Stevenson, B. A. Caughey, O. Abdel-Wahab, D. P. Steensma, N. Galili, A. Raza, H. Kantarjian, R. L. Levine, D. Neuberg, G. Garcia-Manero, B. L. Ebert,

- Validation of a prognostic model and the impact of mutations in patients with lower-risk myelodysplastic syndromes. *Journal of Clinical Oncology* **30**, 3376-3382 (2012).
44. L. Malcovati, E. Hellstrom-Lindberg, D. Bowen, L. Ades, J. Cermak, C. Del Canizo, M. G. Della Porta, P. Fenaux, N. Gattermann, U. Germing, J. H. Jansen, M. Mittelman, G. Mufti, U. Platzbecker, G. F. Sanz, D. Selleslag, M. Skov-Holm, R. Stauder, A. Symeonidis, A. A. van de Loosdrecht, T. de Witte, M. Cazzola, Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukemiaNet. *Blood* **122**, 2943-2964 (2013).
45. M. L. Metzker, Sequencing technologies—the next generation. *Nature Reviews Genetics* **11**, 31-46 (2009).
46. M. Meyerson, S. Gabriel, G. Getz, Advances in understanding cancer genomes through second-generation sequencing. *Nature Reviews Genetics* **11**, 685-696 (2010).
47. H. Löffler, J. Rastetter, T. Haferlach, H. Begemann, *Atlas of clinical hematology*. (Springer, 2005).
48. K. Wang, M. Li, H. Hakonarson, ANNOVAR: functional annotation of genetic

- variants from high-throughput sequencing data. *Nucleic Acids Research* **38**, e164 (2010).
49. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, G. P. D. P. Subgroup, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
50. S. A. Forbes, N. Bindal, S. Bamford, C. Cole, C. Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, J. W. Teague, P. J. Campbell, M. R. Stratton, P. A. Futreal, COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Research* **39**, D945-950 (2011).
51. R. Tibshirani, Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288 (1996).
52. J. J. Goeman, L1 penalized estimation in the cox proportional hazards model. *Biometrical Journal* **52**, 70-84 (2010).
53. R. Davidson, J. G. MacKinnon, Several tests for model specification in the presence of alternative hypotheses. *Econometrica: Journal of the Econometric Society*, 781-793 (1981).

54. S. A. Forbes, N. Bindal, S. Bamford, C. Cole, C. Y. Kok, D. Beare, M. Jia, R. Shepherd, K. Leung, A. Menzies, J. W. Teague, P. J. Campbell, M. R. Stratton, P. A. Futreal, COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Res* **39**, D945-950 (2011).
55. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300 (1995).
56. O. Abdel-Wahab, A. Pardanani, J. Patel, M. Wadleigh, T. Lasho, A. Heguy, M. Beran, D. G. Gilliland, R. L. Levine, A. Tefferi, Concomitant analysis of EZH2 and ASXL1 mutations in myelofibrosis, chronic myelomonocytic leukemia and blast-phase myeloproliferative neoplasms. *Leukemia* **25**, 1200-1202 (2011).
57. M. E. Figueroa, O. Abdel-Wahab, C. Lu, P. S. Ward, J. Patel, A. Shih, Y. Li, N. Bhagwat, A. Vasanthakumar, H. F. Fernandez, M. S. Tallman, Z. Sun, K. Wolniak, J. K. Peeters, W. Liu, S. E. Choe, V. R. Fantin, E. Paietta, B. Löwenberg, J. D. Licht, L. A. Godley, R. Delwel, P. J. Valk, C. B. Thompson, R. L. Levine, A. Melnick, Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2

- function, and impair hematopoietic differentiation. *Cancer Cell* **18**, 553-567 (2010).
58. F. Thol, S. Kade, C. Schlarmann, P. Löffeld, M. Morgan, J. Krauter, M. W. Wlodarski, B. Kölking, M. Wichmann, K. Görlich, G. Göhring, G. Bug, O. Ottmann, C. M. Niemeyer, W. K. Hofmann, B. Schlegelberger, A. Ganser, M. Heuser, Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. *Blood* **119**, 3578-3584 (2012).
59. F. Damm, O. Kosmider, V. Gelsi-Boyer, A. Renneville, N. Carbuccia, C. Hidalgo-Curtis, V. Della Valle, L. Couronné, L. Scourzic, V. Chesnais, A. Guerci-Bresler, B. Slama, O. Beyne-Rauzy, A. Schmidt-Tanguy, A. Stamatoullas-Bastard, F. Dreyfus, T. Prébet, S. de Botton, N. Vey, M. A. Morgan, N. C. Cross, C. Preudhomme, D. Birnbaum, O. A. Bernard, M. Fontenay, G. F. d. Myelodysplasies, Mutations affecting mRNA splicing define distinct clinical phenotypes and correlate with patient outcome in myelodysplastic syndromes. *Blood* **119**, 3211-3218 (2012).
60. F. Damm, V. Chesnais, Y. Nagata, K. Yoshida, L. Scourzic, Y. Okuno, R. Itzykson, M. Sanada, Y. Shiraishi, V. Gelsi-Boyer, A. Renneville, S. Miyano, H. Mori, L. Y. Shih, S.

- Park, F. Dreyfus, A. Guerci-Bresler, E. Solary, C. Rose, S. Cheze, T. Prébet, N. Vey, M. Legentil, Y. Duffourd, S. de Botton, C. Preudhomme, D. Birnbaum, O. A. Bernard, S. Ogawa, M. Fontenay, O. Kosmider, BCOR and BCORL1 mutations in myelodysplastic syndromes and related disorders. *Blood* **122**, 3169-3177 (2013).
61. V. Grossmann, E. Tiacci, A. B. Holmes, A. Kohlmann, M. P. Martelli, W. Kern, A. Spanhol-Rosseto, H. U. Klein, M. Dugas, S. Schindela, V. Trifonov, S. Schnittger, C. Haferlach, R. Bassan, V. A. Wells, O. Spinelli, J. Chan, R. Rossi, S. Baldoni, L. De Carolis, K. Goetze, H. Serve, R. Peceny, K. A. Kreuzer, D. Oruzio, G. Specchia, F. Di Raimondo, F. Fabbiano, M. Sborgia, A. Liso, L. Farinelli, A. Rambaldi, L. Pasqualucci, R. Rabadan, T. Haferlach, B. Falini, Whole-exome sequencing identifies somatic mutations of BCOR in acute myeloid leukemia with normal karyotype. *Blood* **118**, 6153-6163 (2011).
62. H. Makishima, V. Visconte, H. Sakaguchi, A. M. Jankowska, S. A. Kar, A. Jerez, B. Przychodzen, M. Bupathi, K. Guinta, M. G. Afable, Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. *Blood* **119**, 3203-3210 (2012).

63. J. S. Welch, T. J. Ley, D. C. Link, C. A. Miller, D. E. Larson, D. C. Koboldt, L. D. Wartman, T. L. Lamprecht, F. Liu, J. Xia, The origin and evolution of mutations in acute myeloid leukemia. *Cell* **150**, 264-278 (2012).
64. C. L. Willman, C. E. Sever, M. G. Pallavicini, H. Harada, N. Tanaka, M. L. Slovak, H. Yamamoto, K. Harada, T. C. Meeker, A. F. List, Deletion of IRF-1, mapping to chromosome 5q31. 1, in human leukemia and preleukemic myelodysplasia. *Science* **259**, 968-971 (1993).
65. J. Boultonwood, C. Fidler, S. Lewis, A. MacCarthy, H. Sheridan, S. Kelly, D. Oscier, V. Buckle, J. Wainscoat, Allelic loss of IRF1 in myelodysplasia and acute myeloid leukemia: retention of IRF1 on the 5q-chromosome in some patients with the 5q-syndrome. *Blood* **82**, 2611-2616 (1993).
66. L. Malcovati, E. Papaemmanuil, D. T. Bowen, J. Boultonwood, M. G. Della Porta, C. Pascutto, E. Travaglino, M. J. Groves, A. L. Godfrey, I. Ambaglio, A. Gallì, M. C. Da Vià, S. Conte, S. Tauro, N. Keenan, A. Hyslop, J. Hinton, L. J. Mudie, J. S. Wainscoat, P. A. Futreal, M. R. Stratton, P. J. Campbell, E. Hellström-Lindberg, M. Cazzola, C. M. D. W. G. o. t. I. C. G. C. a. o. t. A. I. p. l. R. s. C. G. I. M. Mieloproliferative, Clinical

- significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. *Blood* **118**, 6239-6246 (2011).
67. M. Gerlinger, A. J. Rowan, S. Horswell, J. Larkin, D. Endesfelder, E. Gronroos, P. Martinez, N. Matthews, A. Stewart, P. Tarpey, I. Varela, B. Phillimore, S. Begum, N. Q. McDonald, A. Butler, D. Jones, K. Raine, C. Latimer, C. R. Santos, M. Nohadani, A. C. Eklund, B. Spencer-Dene, G. Clark, L. Pickering, G. Stamp, M. Gore, Z. Szallasi, J. Downward, P. A. Futreal, C. Swanton, Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *New England Journal of Medicine* **366**, 883-892 (2012).
68. P. J. Campbell, E. D. Pleasance, P. J. Stephens, E. Dicks, R. Rance, I. Goodhead, G. A. Follows, A. R. Green, P. A. Futreal, M. R. Stratton, Subclonal phylogenetic structures in cancer revealed by ultra-deep sequencing. *Proceedings of the National Academy of Sciences* **105**, 13081-13086 (2008).
69. L. Ding, M. J. Ellis, S. Li, D. E. Larson, K. Chen, J. W. Wallis, C. C. Harris, M. D. McLellan, R. S. Fulton, L. L. Fulton, R. M. Abbott, J. Hoog, D. J. Dooling, D. C. Koboldt, H. Schmidt, J. Kalicki, Q. Zhang, L. Chen, L. Lin, M. C. Wendl, J. F.

- McMichael, V. J. Magrini, L. Cook, S. D. McGrath, T. L. Vickery, E. Appelbaum, K. Deschryver, S. Davies, T. Guintoli, R. Crowder, Y. Tao, J. E. Snider, S. M. Smith, A. F. Dukes, G. E. Sanderson, C. S. Pohl, K. D. Delehaunty, C. C. Fronick, K. A. Pape, J. S. Reed, J. S. Robinson, J. S. Hodges, W. Schierding, N. D. Dees, D. Shen, D. P. Locke, M. E. Wiechert, J. M. Eldred, J. B. Peck, B. J. Oberkfell, J. T. Lolofie, F. Du, A. E. Hawkins, M. D. O'Laughlin, K. E. Bernard, M. Cunningham, G. Elliott, M. D. Mason, D. M. Thompson, J. L. Ivanovich, P. J. Goodfellow, C. M. Perou, G. M. Weinstock, R. Aft, M. Watson, T. J. Ley, R. K. Wilson, E. R. Mardis, Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* **464**, 999-1005 (2010).
70. L. Ding, T. J. Ley, D. E. Larson, C. A. Miller, D. C. Koboldt, J. S. Welch, J. K. Ritchey, M. A. Young, T. Lamprecht, M. D. McLellan, J. F. McMichael, J. W. Wallis, C. Lu, D. Shen, C. C. Harris, D. J. Dooling, R. S. Fulton, L. L. Fulton, K. Chen, H. Schmidt, J. Kalicki-Veizer, V. J. Magrini, L. Cook, S. D. McGrath, T. L. Vickery, M. C. Wendl, S. Heath, M. A. Watson, D. C. Link, M. H. Tomasson, W. D. Shannon, J. E. Payton, S. Kulkarni, P. Westervelt, M. J. Walter, T. A. Graubert, E. R. Mardis, R. K. Wilson, J. F. DiPersio, Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* **481**, 506-510 (2012).

71. S. Nik-Zainal, P. Van Loo, D. C. Wedge, L. B. Alexandrov, C. D. Greenman, K. W. Lau, K. Raine, D. Jones, J. Marshall, M. Ramakrishna, A. Shlien, S. L. Cooke, J. Hinton, A. Menzies, L. A. Stebbings, C. Leroy, M. Jia, R. Rance, L. J. Mudie, S. J. Gamble, P. J. Stephens, S. McLaren, P. S. Tarpey, E. Papaemmanuil, H. R. Davies, I. Varela, D. J. McBride, G. R. Bignell, K. Leung, A. P. Butler, J. W. Teague, S. Martin, G. Jönsson, O. Mariani, S. Boyault, P. Miron, A. Fatima, A. Langerød, S. A. Aparicio, A. Tutt, A. M. Sieuwerts, Å. Borg, G. Thomas, A. V. Salomon, A. L. Richardson, A. L. Børresen-Dale, P. A. Futreal, M. R. Stratton, P. J. Campbell, B. C. W. G. o. t. I. C. G. Consortium, The life history of 21 breast cancers. *Cell* **149**, 994-1007 (2012).