GENETIC ARCHITECTURE OF LEUKAEMIA

Mel Greaves

Munich, 28 February 2011
GENETIC ARCHITECTURE OF CHILDHOOD LEUKAEMIA

En bloc
Inherited allelic variants

Sequentially acquired somatic cell mutations (‘drivers’)

RISK / Susceptibility

Initiation
Promotion
Progression
Tx
Relapse

Sub-clonal segregation
- stem cells

- stem cells
• Candidate gene studies
  - Immune response genes: HLA, IL-12
  - Folate metabolism: MTHFR
  - Activation / detox: GSTs, NQ01

- Lack of reproducibility
  (under-powered studies?)
VARIATION IN GENETIC SUSCEPTIBILITY TO CHILDHOOD ALL

- **Genome-Wide** Associated Studies (GWAS)
  - SNP-associated normal genomes (Ca. vs control)
    + agnostic, no prior assumptions
      - precedents with other cancers
    - incomplete genome coverage
      - biased towards common alleles
      - may not identify functional variant
      - large case series required
VARIATION IN GENETIC SUSCEPTIBILITY TO CHILDHOOD ALL

- **Genome-Wide Associated Studies (GWAS)**
  - SNP-associated normal genomes (Ca. vs control)

Richard Houlston et al  Childhood ALL
1° GWA 1
577 (1438 controls) 392 (960 controls)
Papaemmanuil E et al, Nature Genet, 2009; 41: 1006-1010

2° Validation 1 (for candidate genes)
1384 (1877 controls) Germany
Prasad R et al, Blood, 2010; 115: 1765-1767

3° Validation 2 (for candidate gene)
1428 (1516 controls) Germany
148 (137 controls) Spain
550 (450 controls) Hungary
260 (266 controls) Canada
Sherborne AL et al, Nature Genet, 2010; 42: 492-494
# GWAS FOR CHILDHOOD ALL

Common alleles of:

<table>
<thead>
<tr>
<th>Allele</th>
<th>OR</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IKZF1</strong></td>
<td>1.69</td>
<td>1.2 x 10^{-19} *</td>
</tr>
<tr>
<td><strong>ARID5B</strong> (hyperdiploid)</td>
<td>1.65</td>
<td>6.7 x 10^{-19} *</td>
</tr>
<tr>
<td><strong>CEBPE</strong></td>
<td>1.34</td>
<td>2.9 x 10^{-7}</td>
</tr>
<tr>
<td><strong>CDKN2A / p16</strong></td>
<td>1.42</td>
<td>3.0 x 10^{-11}</td>
</tr>
</tbody>
</table>

Per allele 1.53 3.5 x 10^{42} (trend)

= Additive risk ~10x

* Trevino LR *et al*, Nature Genet, 2009; 41: 1001-1005
GENETIC SUSCEPTIBILITY AND MOLECULAR PATHOGENESIS OF ALL

1. L/M stem cell
2. pre-pro-B
3. pro-B
4. pre-B
5. B

- **CEBP**
- **IKZF1**
- **ARID5B / RAGs**
- **CDKN2A / p16**

Myeloid lineage

ALL
ALLELIC ARCHITECTURE of GENETIC SUSCEPTIBILITY to ALL

- Same genes involved as somatic mutants

- Gene variants that impact on intrinsic vulnerability of ‘target’ cells, i.e. component of molecular pathogenesis

  - via expression levels of key regulators of differentiation and cell cycle
Normalized $\log_2$ expression of IKZF1

- CC ($n = 8$)
- CA ($n = 40$)
- AA ($n = 39$)

$P = 0.005$
FUTURE:

1. Full GWAS on German series (1,500 cases)
2. International consortium (4-5,000 cases). Deep mining
3. Identification of functional variants (re-sequencing)
4. Functional validation
5. Weighted contribution to risk of ALL?
CAUSAL MECHANISMS for LEUKAEMIA / CANCER

Mutagenic Exposures

Risk 'Modifiers'
- genetics
- diet

Mutation / clonal selection
CAUSAL MECHANISMS for LEUKAEMIA / CANCER

Mutagenic Exposures

Risk ‘Modifiers’
- genetics
- diet

Mutation / clonal selection

C = ALL : 1 in 2000
Speculations on the cause of childhood acute lymphoblastic leukemia. *Leukemia*, 2: 120-125
History!
BACKTRACKING THE PRE-NATAL ORIGINS OF CHILDHOOD LEUKAEMIA

• Monozygotic (monochorionic) twins with concordant ALL

• Archived neonatal blood spots (Guthrie cards) of patients with ALL / AML

• Frozen cord bloods
  - patients with ALL (rare)
  - unselected cohort
EARLY OR INITIATING EVENTS IN LEUKAEMOGENESIS

• Foetal haemopoiesis

• Chromosome translocation / gene fusions
  * MLL-AF4
  * ETV6-RUNX1 (*TEL-AML1*)
  * AML1-ETO

• Chromosomal hyperdiploidy

• Chromosomal instability (rare)

• Mutations - *GATA1* in TMD / AML in Down’s

SECONDARY, POST-NATAL MUTATIONS ARE ESSENTIAL FOR CLINICAL LEUKAEMIA

- Incidence of pre-leukaemic clones in cord blood (100x)
- Concordance rate in twins is 10-15%
- Other mutations are present and detectable at diagnosis
- Model systems:

  - \textit{ETV6-RUNX1} transgenics (Ford et al, \textit{J Clin Invest} 2009)
  - \textit{ETV6-RUNX1} murine stem cell transplants (Tsuzuki et al, \textit{PNAS} 2004)
  - \textit{ETV6-RUNX1} in human stem cells (Hong et al, \textit{Science} 2008)
### Deletions

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Gene(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>12p13.2</td>
<td>ETV6</td>
</tr>
<tr>
<td>9p13.2</td>
<td>PAX5</td>
</tr>
<tr>
<td>9p21.3</td>
<td>CDKN2A</td>
</tr>
<tr>
<td>3q13.2</td>
<td>CD200, BTLA</td>
</tr>
<tr>
<td>6q16.2-q16.3</td>
<td>16 genes, inc CCNC</td>
</tr>
<tr>
<td>6q21</td>
<td>ARMC2, SESN1</td>
</tr>
<tr>
<td>3q26.32</td>
<td>TBL1XR1</td>
</tr>
<tr>
<td>12q21.33</td>
<td>BTG1</td>
</tr>
<tr>
<td>1q31.3</td>
<td>TROVE2, GLRX2 etc</td>
</tr>
<tr>
<td>4q31.21</td>
<td>telomeric to NR3C2</td>
</tr>
<tr>
<td>5q31.3</td>
<td>NR3C1, LOC389335</td>
</tr>
<tr>
<td>5q33.3</td>
<td>EBF1</td>
</tr>
<tr>
<td>3p14.2</td>
<td>FHIT</td>
</tr>
<tr>
<td>8q12.1</td>
<td>5' of TOX</td>
</tr>
</tbody>
</table>

### Gains

- Xq
- 21q22.11-q22.12: 33 genes including RUNX1
- 1q23.3-q44

---

‘DRIVERS’ (recurrency / function) vs. ‘PASSENGERS’
SEQUENTIAL MODEL FOR ALL

1. ETV6-RUNX1\(^+\) (TEL-AML1) → covert pre-ALL

2. CNA (\(~4\))
   - ETV6\(^-\)
   - PAX5\(^-\)
   - p16\(^-\)
   - Chr21q\(^+\)
   - Secondary?

ALL 1-15 yrs (2-5)

SNP arrays
CONCORDANT ALL IN MONOZYGOTIC TWINS

(5x) ETV6-RUNX1 / Hyperdiploidy (3x)

f1

f2

single placenta

Tw1

Shared pre-ALL clone

+CNA (3/4) all distinct +CNA (3/4)

Tw2

ALL

ALL

birth

C Bateman

Blood, 2010, 115: 3553-8

ASH 2010
ETV6-RUNXI

ALL

DISCORDANT

Risk = ~1 in 10

Hong D et al
Science, 2008
‘ARRESTED’ ANCESTRAL CLONES IN MONOZYGOTIC TWINS *DISCORDANT FOR ALL*

Initiating mutation

ETV6-RUNX1  
BCR-ABL1  
Hyperdiploidy

\[ \text{f}1 \quad \text{f}2 \]

Tw1  
Tw2  

ALL  
healthy

birth

single placenta

+ secondary mutations / CNA
‘ARRESTED’ ANCESTRAL CLONES IN MONOZYGOTIC TWINS _DISCORDANT FOR ALL_

**Initiating mutation**

ETV6-RUNX1  
BCR-ABL1  
Hyperdiploidy

**f1**  
**f2**  
single placenta

**Tw1**  
Shared pre-ALL clone

**Tw2**  

birth

ALL  
healthy

+ secondary mutations / CNA
‘ARRESTED’ ANCESTRAL CLONES IN MONOZYGOTIC TWINS DISCORDANT FOR ALL

Initiating mutation

- ETV6-RUNX1
- BCR-ABL1
- Hyperdiploidy

Single placenta

Tw1

Tw2

Shared pre-ALL clone

ALL

covert pre-ALL

absent

+ secondary mutations / CNA

~x10^-4
**BCR-ABL1 ALL IN MONOZYGOTIC TWINS**

**CONCORDANT**

- Tw1
- BCR-ABL1
- IKZF1
- ALL
- Hyperdiploidy
- alive 2 yrs +
- †

- Tw2
- BCR-ABL1
- ALL
- Hyperdiploidy
- alive 2 yrs +
- †

**DISCORDANT**

- Tw3
- BCR-ABL1
- ALL
- IKZF1
- Healthy
- †
- †
- ‘pre-ALL’
- 2 yrs +
SEQUENTIAL MODEL FOR ALL

1

ETV6-RUNX1⁺
(TEL-AML1)

birth

2

covet

pre-ALL

CNA
(~4)
‘driver’

ETV6⁻
PAX5⁻
p16⁻
Chr21q⁺

ALL

1-15 yrs
(2-5)
SEQUENTIAL MODEL FOR ALL: NUMBER OF MUTATIONS?

1. ETV6-RUNX1⁺ (TEL-AML1)
   - covert pre-ALL

2. CNA (~4) ‘driver’
   - ETV6⁻
   - PAX5⁻
   - p16⁻
   - Chr21q⁺

ALL

1-15 yrs (2-5)
CRYPTIC MUTATIONS IN ALL?

• 50 cases of \textit{ETV6-RUNX1}^{+} \textit{ALL} (+ remission / normal DNA)
  - Paired end sequencing (- cryptic rearrangements)
  - Exon pull-down, solexa sequencing (- sequence based mutations)

  (Collaboration with Sanger Centre; P Campbell \textit{et al})

• 7 cases of \textit{ALL} (+ remission / normal DNA)
  - Whole genome sequencing

  (Collaboration with R Houlston, ICR and Complete Genomics)
ETV6-RUNX1+ (TEL-AML1) → covert pre-ALL

CNA (~4) → ALL

ETV6-, PAX5-, p16-, Chr21q+

1-15 yrs (2-5)
A LINEAR CLONAL ARCHITECTURE?

Initiation

gene fusion

Promotion

Dominant sub-clone

CNA accrual

ALL
THE CANCER GENOME ‘LANDSCAPE’

Genetic variegation of clonal architecture and propagating cells in leukaemia

Kristina Anderson, Christoph Lutz, Frederik van Delft, Caroline Bateman, Yanping Guo, Susan Colman, Helena Kempski, Anthony Moorman, Ian Titley, John Swansbury, Lyndal Kearney, Tariq Enver, Mel Greaves

Anderson K et al
*Nature*, 2011, 469: 356-361

Kristina Anderson
Fusion gene

- *ETV6*
  + copies of fusion gene
  + copy of Chr 21q (*RUNX1*) / r+2
- copies of *PAX5* (-/+ or -/-)
- copies of *CDKN2a/p16* (-/+ or -/-)

→ Genotype distinct sub-clones (%)
→ Clonal architecture / ancestral tree

~30 cases of *ETV6-RUNX1*+ ALL
Pt #24: SNAPSHOT OF ANCESTRAL GENETIC TREES IN ALL

23%
1 F
2 RUNX1
1 ETV6
2 PAX5

64%
1 F
2 RUNX1
0 ETV6
2 PAX5

11%
1 F
2 RUNX1
0 ETV6
1 PAX5
Pt #36: SNAPSHOT OF ANCESTRAL GENETIC TREES IN ALL
Pt #33: SNAPSHOT OF ANCESTRAL GENETIC TREES IN ALL

1 F
3 RUNX1
1 ETV6
2 PAX5

1 F
2 RUNX1
1 ETV6
2 PAX5

1 F
3 RUNX1
1 ETV6
1 PAX5

1 F
2 RUNX1
0 ETV6
2 PAX5

1 F
2 RUNX1
1 ETV6
1 PAX5

2 F
2 RUNX1
1 ETV6
2 PAX5

1 F
3 RUNX1
0 ETV6
2 PAX5

1 F
3 RUNX1
0 ETV6
1 PAX5
Clonal Dynamics

D56 aplasia → 7 months → ALL

- 11% with 1 F, 2 RUNX1, 1 ETV6, 2 CDKN2A, 1 11q
- 16% with 1 F, 2 RUNX1, 1 ETV6, 1 CDKN2A, 1 11q
- 69% with 1 F, 2 RUNX1, 1 ETV6, 0 CDKN2A, 1 11q
- 4% with 1 F, 2 RUNX1, 1 ETV6, 0 CDKN2A, 0 11q
Clonal Dynamics

D56 aplasia → 7 months → ALL

1 F
2 RUNX1
1 ETV6
2 CDKN2A
2 11q
4%

1 F
2 RUNX1
1 ETV6
2 CDKN2A
1 11q
88%

1 F
2 RUNX1
1 ETV6
2 CDKN2A
0 11q
8%

1 F
2 RUNX1
1 ETV6
2 CDKN2A
1 11q
11%

1 F
2 RUNX1
1 ETV6
2 CDKN2A
1 11q
11%

1 F
2 RUNX1
1 ETV6
0 CDKN2A
1 11q
69%

1 F
2 RUNX1
1 ETV6
0 CDKN2A
1 11q
69%

1 F
2 RUNX1
1 ETV6
1 CDKN2A
1 11q
16%

1 F
2 RUNX1
1 ETV6
1 CDKN2A
1 11q
16%

1 F
2 RUNX1
1 ETV6
1 CDKN2A
1 11q
4%

1 F
2 RUNX1
1 ETV6
0 CDKN2A
0 11q
4%
Pt #6: ANCESTRAL TREE RECONSTRUCTION IN RELAPSE

Diagnosis

1 F
2 RUNX1
1 ETV6
0 CDKN2A
19%

2 F
2 RUNX1
1 ETV6
0 CDKN2A
2%

1 F
2 RUNX1
1 ETV6
1 CDKN2A
2%

4%

1 F
2 RUNX1
1 ETV6
0 CDKN2A
4%

73%

1 F
2 RUNX1
1 CDKN2A
0%

0%

1 F
2 RUNX1
1 ETV6
2 CDKN2A
19%

2%
Pt #6:  ANCESTRAL TREE RECONSTRUCTION IN RELAPSE

Diagnosis

- 1 F 2 RUNX1 1 ETV6 0 CDKN2A: 19%
- 1 F 2 RUNX1 1 ETV6 1 CDKN2A: 2% (→ 2 F 2 RUNX1 1 ETV6 0 CDKN2A: 73%)
- 1 F 2 RUNX1 1 ETV6 1 CDKN2A: 4%
- 1 F 2 RUNX1 0 ETV6 1 CDKN2A: 2% (→ 1 F 2 RUNX1 0 ETV6 0 CDKN2A: 89%)

Relapse

- 1 F 2 RUNX1 1 ETV6 2 CDKN2A: 0% (→ 1 F 2 RUNX1 1 ETV6 1 CDKN2A: 5%)
- 1 F 2 RUNX1 0 ETV6 1 CDKN2A: 5%
Pt #6: ANCESTRAL TREE RECONSTRUCTION IN RELAPSE

Diagnosis

Relapse

1 F 2 RUNX1 1 ETV6 1 CDKN2A

0%

5%

?  

19%

2%  

5%  

4%  

1 F 2 RUNX1 1 ETV6 0 CDKN2A

2 F 2 RUNX1 1 ETV6 0 CDKN2A

2%  

73%

1 F 2 RUNX1 1 ETV6 1 CDKN2A

89%  

2 F 2 RUNX1 1 ETV6 0 CDKN2A
INTRA-CLONAL GENETIC HETEROGENEITY IN ALL

- Multiple sub-clones (3 - 14) (under-estimate)
- Independent / multiple acquisition of recurrent CNA (mechanistic issues)
- No preferential order of CNA
- Relapse originating from major or minor clones at diagnosis
- Relapse clone diversifies and may reiterate evolution of sub-clones at diagnosis
- Non-linear dynamic / branching clonal architecture
Charles Darwin
(Transmutation notebook B)

1837

‘Ancestral tree’
INTRA-CLONAL GENETIC COMPLEXITY IN CANCER

AML: FLT3\textsuperscript{ITD} / RAS\textsuperscript{m} / sub-clonal

- Barrett’s oesophagus / oesophageal ca.
- Multi-focal bladder ca.
- Colon adenoma / carcinoma transition
- Genetic diversification of clonal metastases in pancreatic ca.
- Topographical genetic variation in prostate ca.
Prostate 2007

Map of Whole prostate slice: ERG rearrangement and PTEN loss

**EDel Tumour**

**2EDel Tumour** (poor prognosis)

**Esplits Tumour**

**PTEN loss** in 1EDel region of tumour

Clark *et al* Oncogene 2007
INTRA-CLONAL GENETIC COMPLEXITY IN CANCER

- Prognostic sample bias
- Substrate for drug resistance
- Identifying therapeutic targets
- Cancer stem cell heterogeneity?
CANCER STEM CELLS

CANCER CLONE DIVERSITY AND PROPAGATION

MODELS

• Stem cell
  (fixed; developmentally hierarchical)
  - or -

• Stochastic
  (random, variable)
  - or -

• Evolution
  (dominant sub-clone)

? Frequency / phenotype

Sustains cancer
Self-renews
Target for therapy
EVOLUTIONARY PROGRESSION

Bottleneck

- speciation
- antibiotic R.
- immune selection
- cancer
EVOLUTIONARY PROGRESSION

Bottleneck

- genetic diversity in units of selection (= leukaemia ‘stem’ / propagating cell)
ARE LEUKAEMIC STEM CELLS GENETICALLY DIVERSE?

Inter-tibial injection → NOD/SCID/γ mice → 8-12 weeks → Re-screen (M-FISH) for genetic complexity of leukaemic cells → re-transplant

ALL cells of defined genetic, sub-clonal complexity

Stem cell dependent regeneration of ALL
Pt #4

Diagnostic sample:

1 F 2 RUNX1 1 ETV6 2 CDKN2A

1 F 3 RUNX1 0 ETV6 2 CDKN2A

1 F 2 RUNX1 0 ETV6 1 CDKN2A

33%

21%

12%

33%
SNP array
FISH

+ chr. 22

[Diagram showing analysis of genetic markers with SNP array and FISH techniques for different chromosome segments, including primary and secondary transplant scenarios.]
THE ‘BACK TO DARWIN’ MODEL: STEM CELL HIERARCHIES IN ALL

pre-LSC

LSCs'

LSCs''

Initiation/
pre-leukaemia

Diagnosis

Tx

Relapse
(metastasis)

= genetically distinct stem cells
Glialblastoma Multiforme

Clonally related but genetically karyotypically distinct

All 4 segments transplant (CNS) in NOD/SCID

PANCREATIC CANCER

1°

A
B
C

metastases

Clonally related / genetically distinct

THE ‘BACK TO DARWIN’ MODEL: IMPLICATIONS

• Cancer stem cells (CSC) are genetically diverse – phenotypic diversity and frequency variation

• CSC are the units of selection in evolutionary progression of cancer and therapeutic resistance

• CSC are a diverse, moving and elusive therapeutic target
  – founder mutation is only universal target