Origins of cancer genome complexity revealed by haplotype-resolved genomic analysis of Barrett's esophagus to esophageal adenocarcinoma progression

Matthew Stachler

https://stachlerlab.ucsf.edu/

With Special guest Dr. Cheng-Zhong Zhang (DFCI)

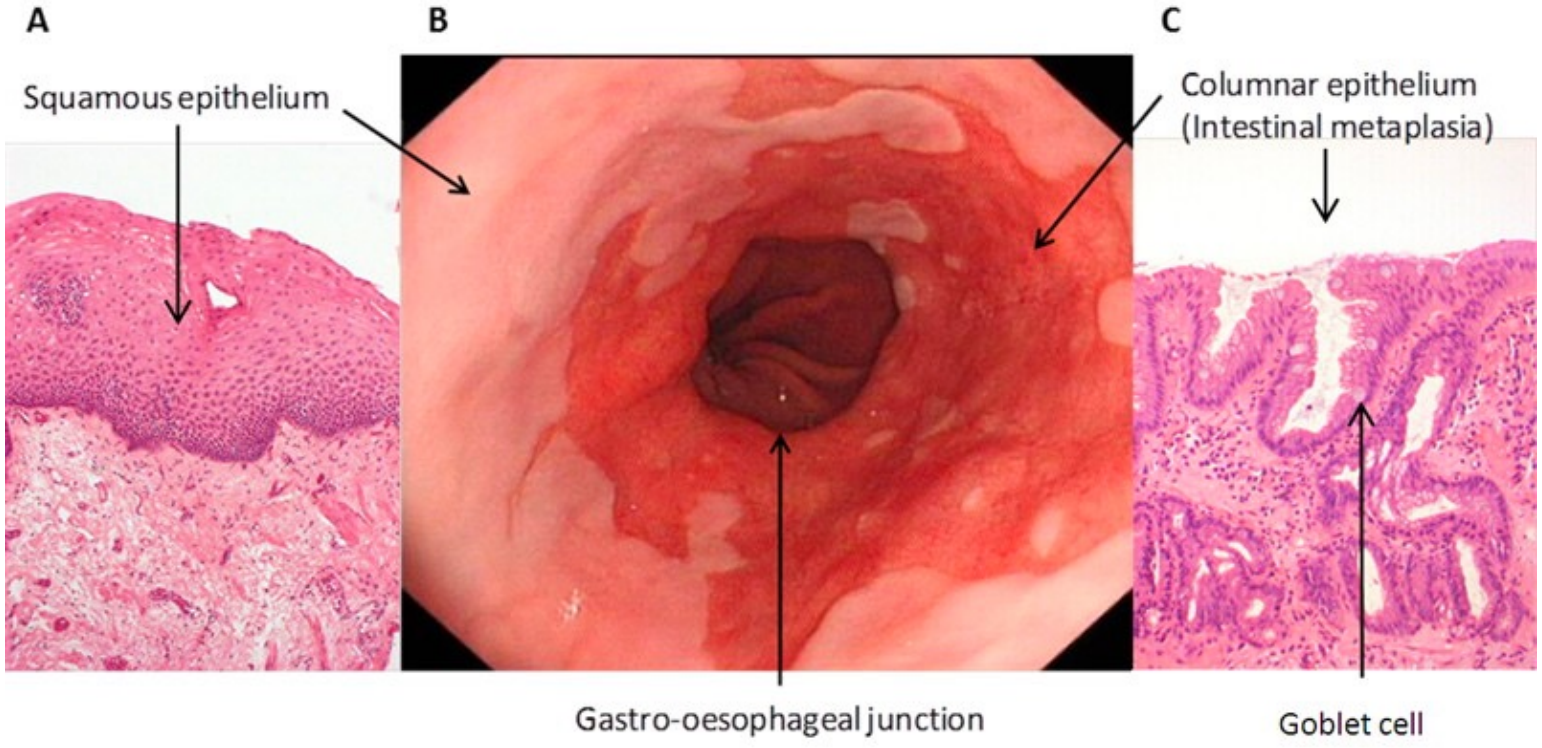
UCSF School of Medicine

UCSF Helen Diller Family Comprehensive Cancer Center

What is Barrett's esophagus?

Barrett's esophagus:

- Barrett's esophagus is the pre-cancerous lesion of esophageal adenocarcinoma
- Replacement of the normally squamous lined lower esophagus is with a columnar epithelium that develops intestinal differentiation
- Thought to be due to chronic reflux (heart burn) and inflammation



Barrett’s esophagus and esophageal adenocarcinoma

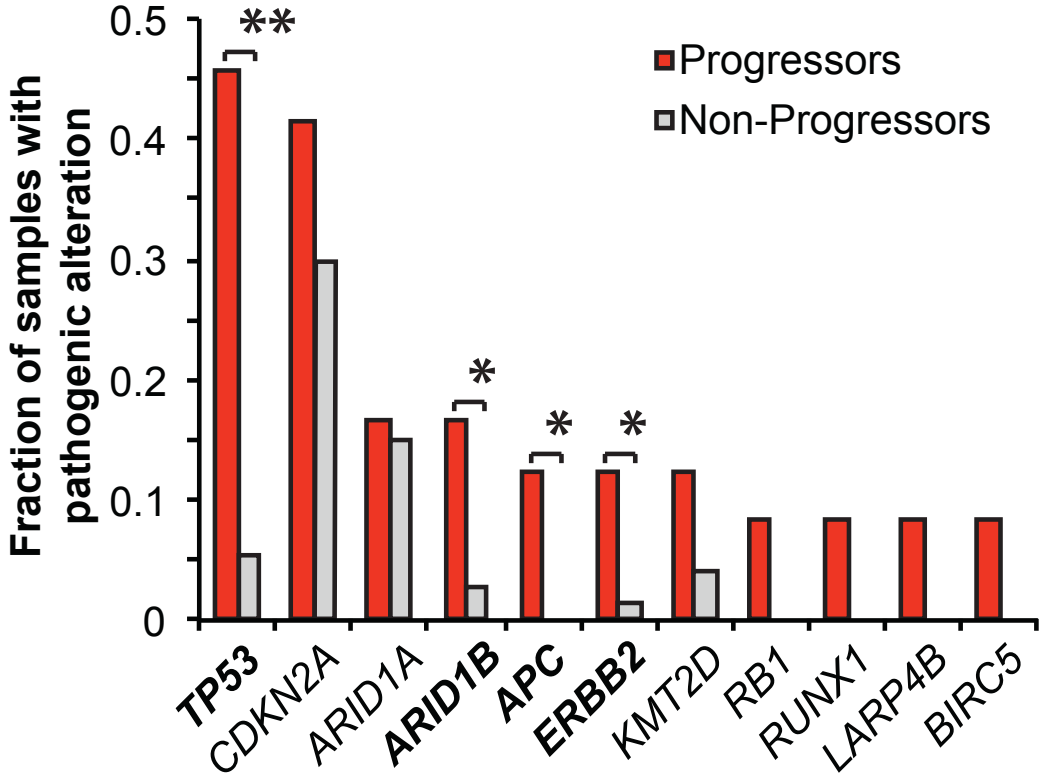
Barrett’s esophagus > Barrett's esophagus with dysplasia > esophageal adenocarcinoma

(very common) (less common) (rare, but rapidly growing incidence)

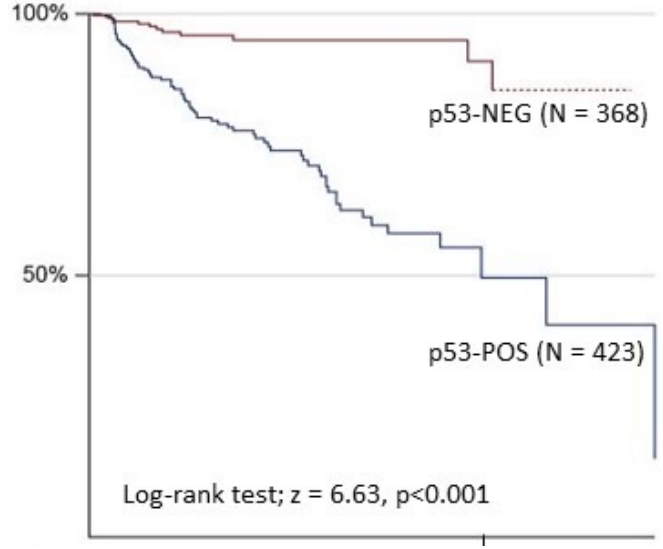
- EAC is a deadly cancer with a drastically increasing incidence
  - BE is the precursor (pre-cancerous lesion) and largest risk factor for developing EAC
  - Despite knowing this, currently we do a poor job at identifying the right patients to treat.
    - BE is extremely common
      - Low overall risk of progression
    - Risk currently determined by diagnosis of dysplasia
      - Challenging diagnosis with significant disagreement
    - Newer studies suggest the time between the development of dysplasia > cancer shorter than originally thought

Background

- *TP53* alterations occur early in progression process

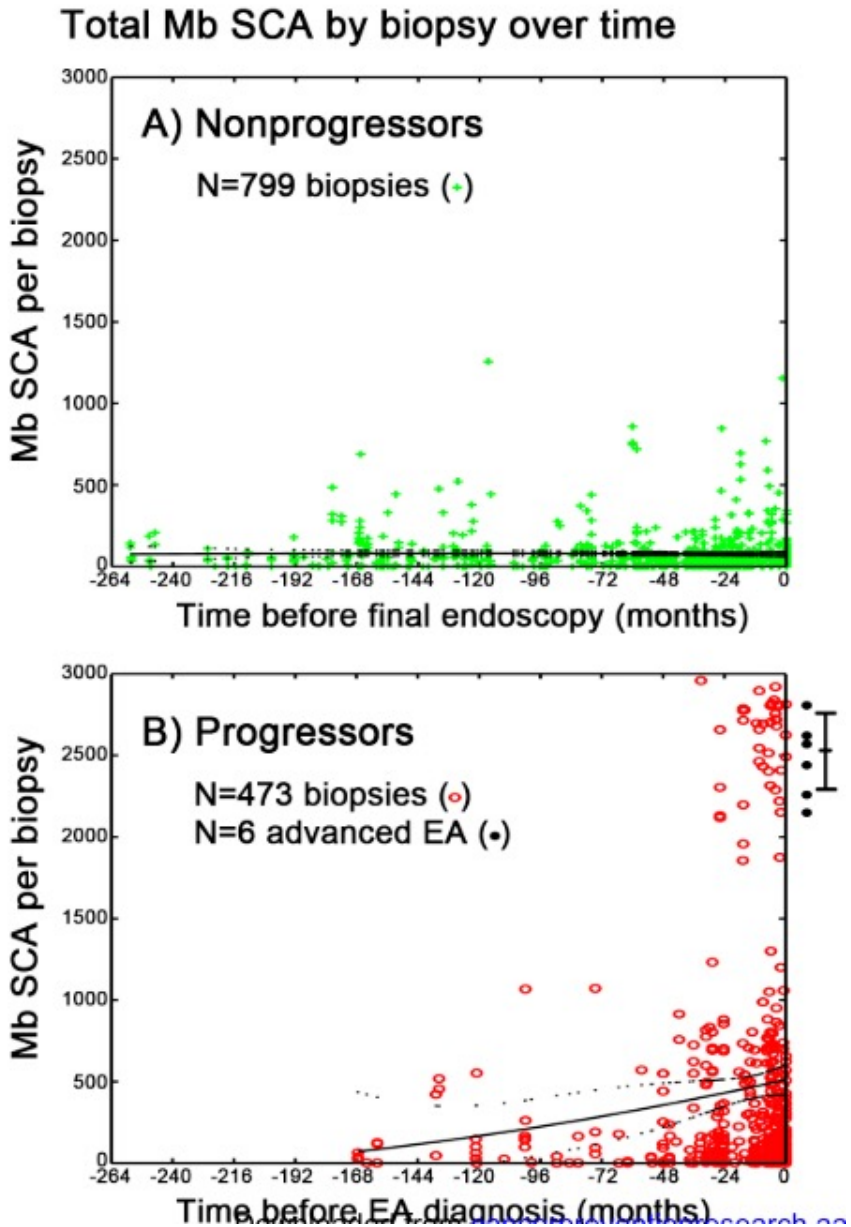


(Stachler, Gastro, 2018)



Background

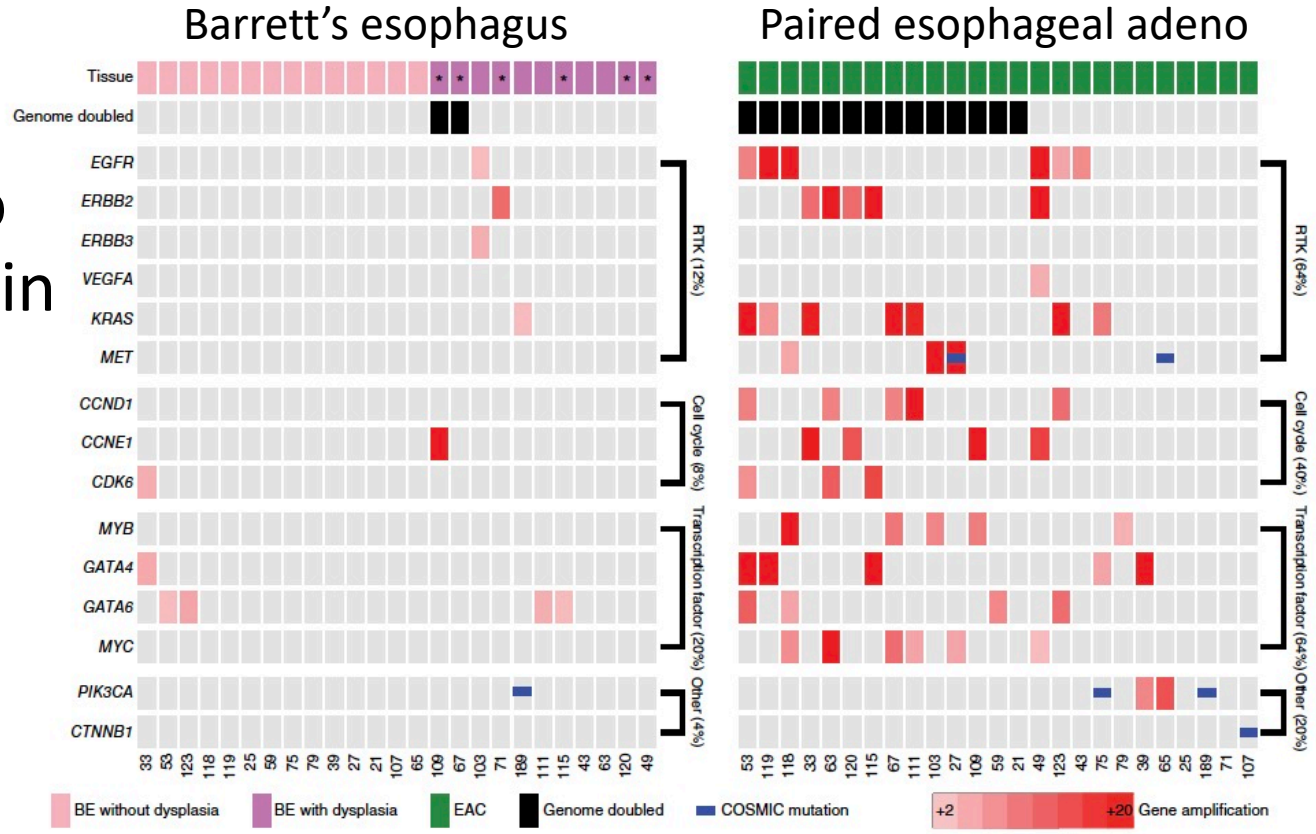
- *TP53* alterations occur early in progression process
- Copy number alterations also begin to form in NDBE but rapidly accumulate in dysplasia



(Li, Cancer Prev Res, 2014)

Background

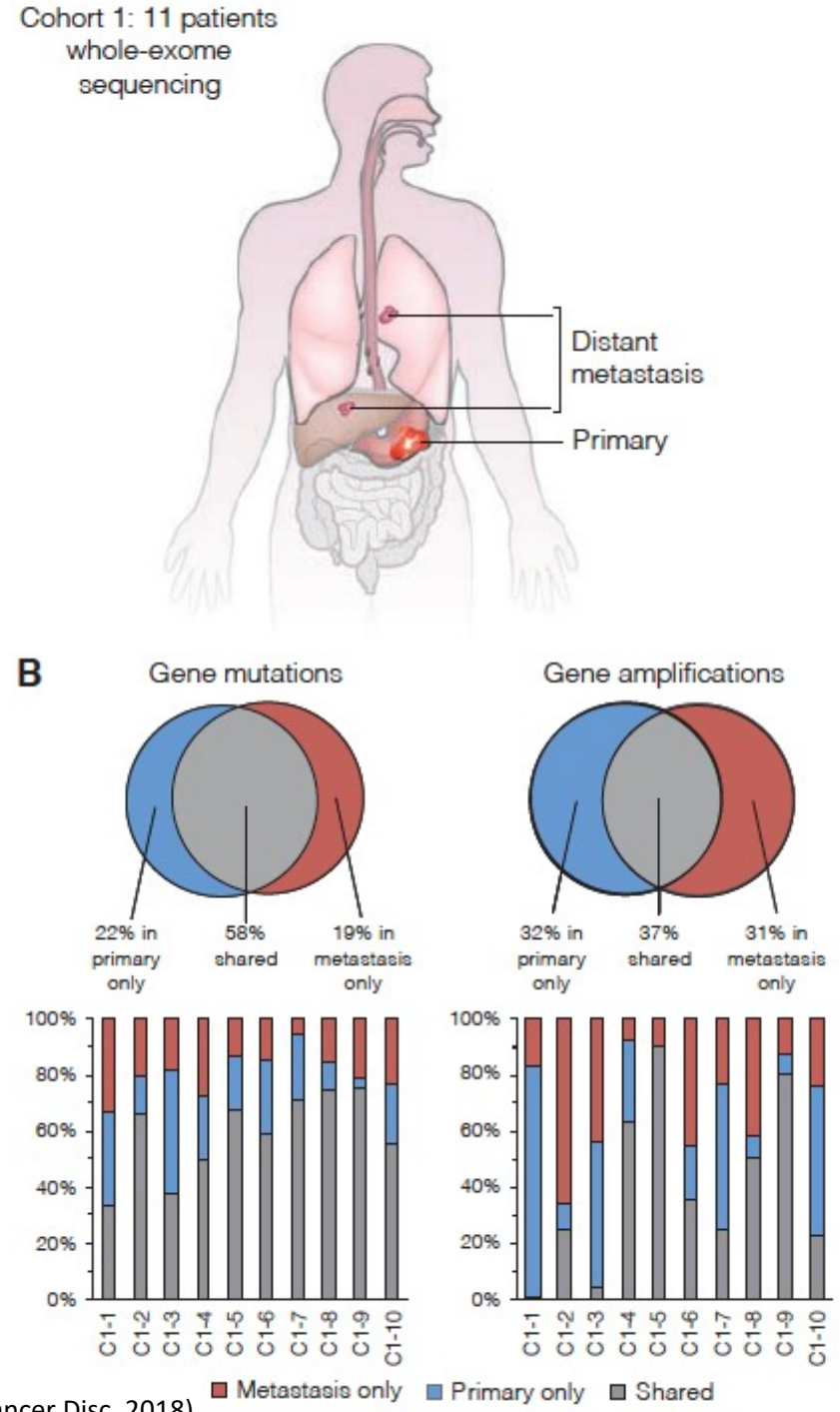
- TP53 alterations occur early in progression process
- Copy number alterations also begin to form in NDBE and rapidly accumulate in dysplasia
- High-level amplifications of driver oncogenes occur late



(Stachler, Nat Gen, 2015)

Background

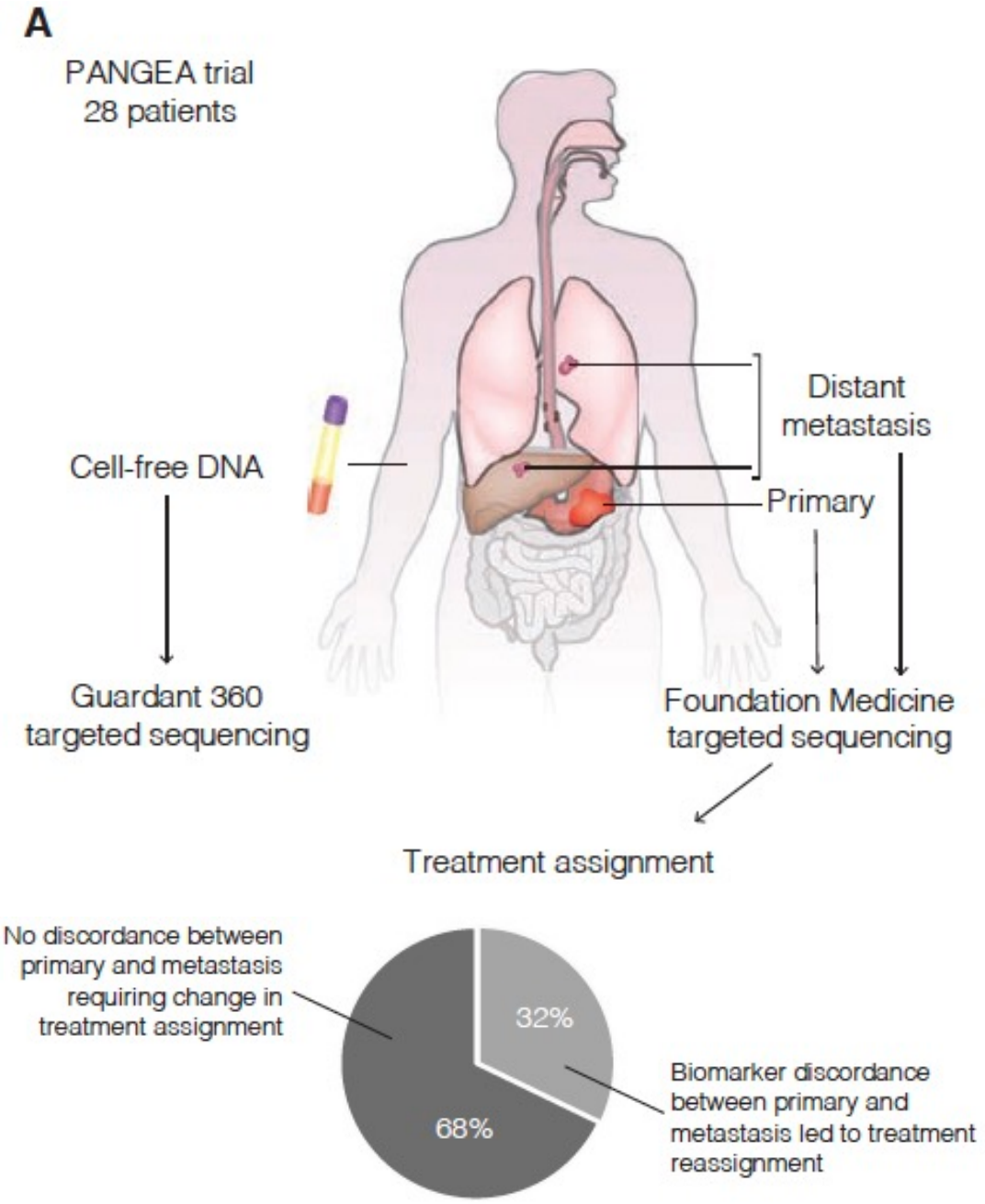
- TP53 alterations occur early in progression process
- Copy number alterations also begin to form in NDBE and rapidly accumulate in dysplasia
- High-level amplifications of driver oncogenes occur late
- Despite being ‘vital’ to EAC, driver oncogene amplifications can show significant intra-patient heterogeneity



(Pectasides, Cancer Disc, 2018)

Background

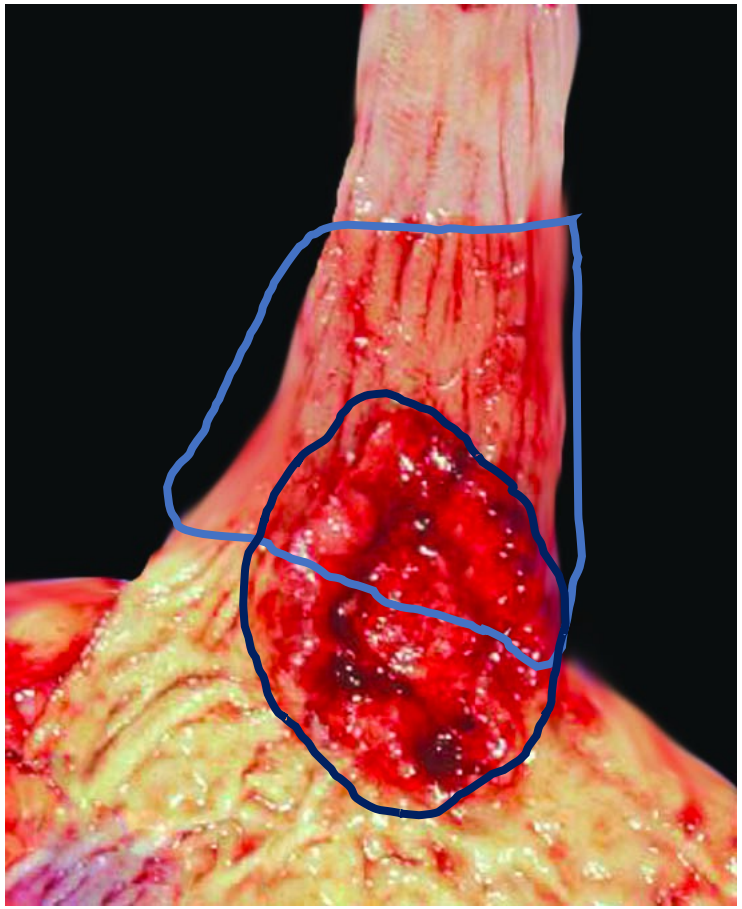
- TP53 alterations occur early in progression process
- Copy number alterations also begin to form in NDBE and rapidly accumulate in dysplasia
- High-level amplifications of driver oncogenes occur late
- Despite being ‘vital’ to EAC, driver oncogene amplifications can show significant intra-patient heterogeneity

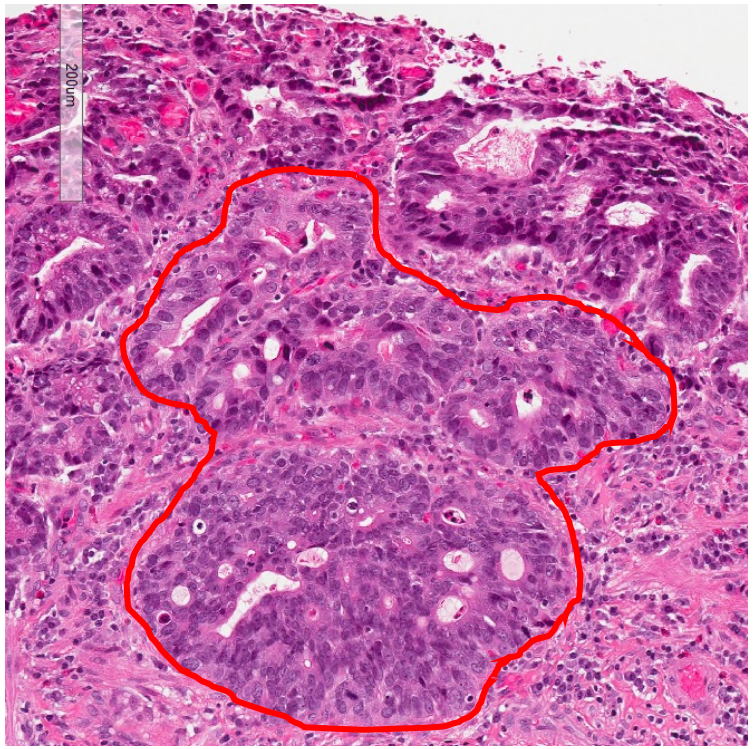


(Pectasides, Cancer Disc, 2018)

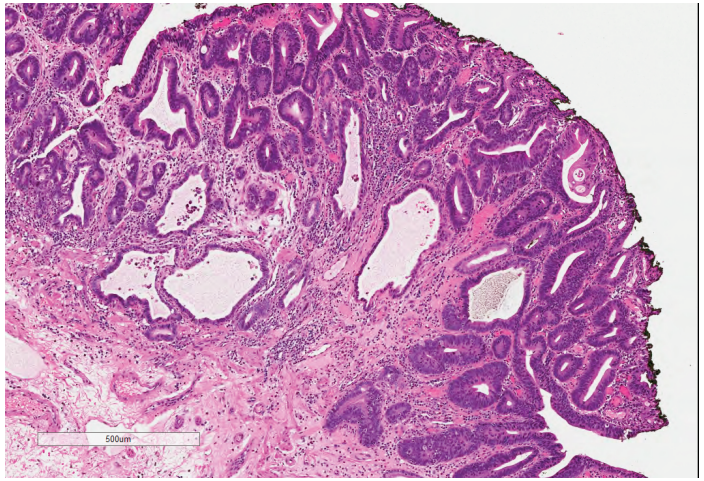
Understanding the transition into invasive cancer

- Previous studies have focused on esophagectomies with large tumors
- Area of dysplasia most closely related to cancer likely overgrown/destroyed
- Instead wanted to focus on very early (microscopic) cancers





http://www.pathologyoutlines.com/topic/esophagusadenocarcinoma.html

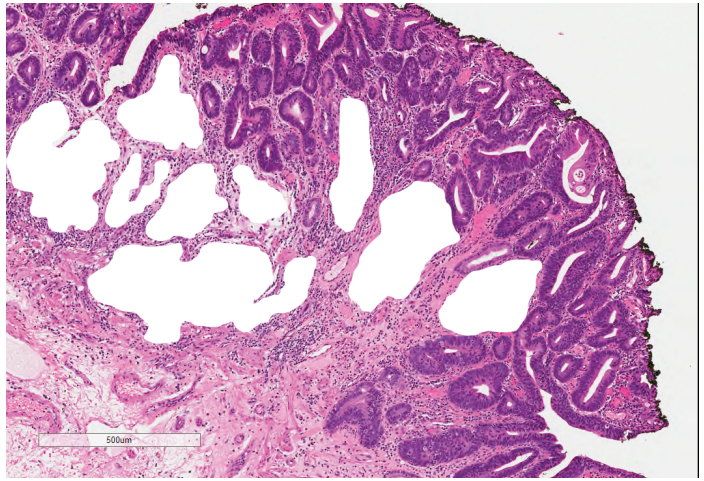


Histologic review

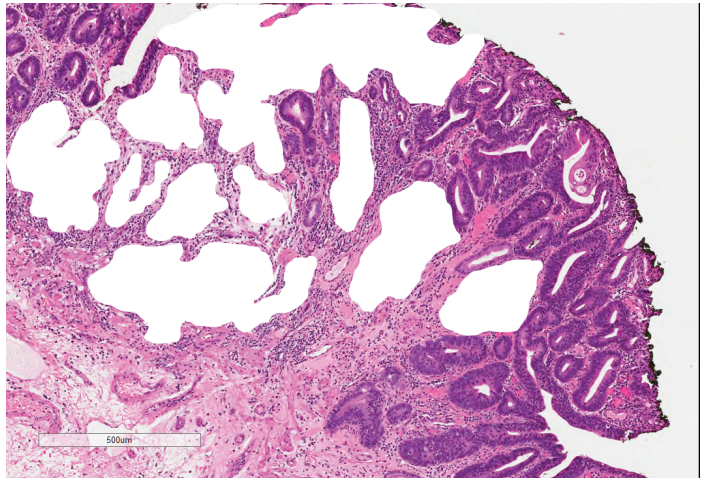
Pathology archieve search to include only small, early cancers
- 15 patients with Intramucosal adenocarcinoma or T1 adenocarcinoma
- No prior therapy

Sequential Laser Capture Microdissection (EAC > HGD > LGD > NDBE), 76 total samples

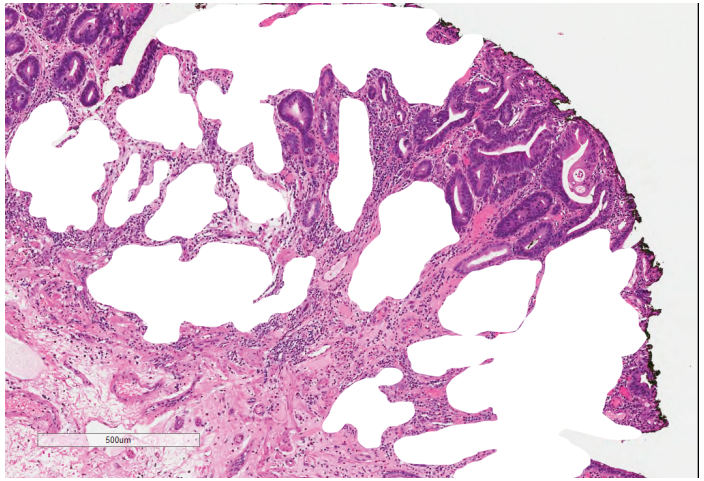
Cut IMEAC



Cut first HGD

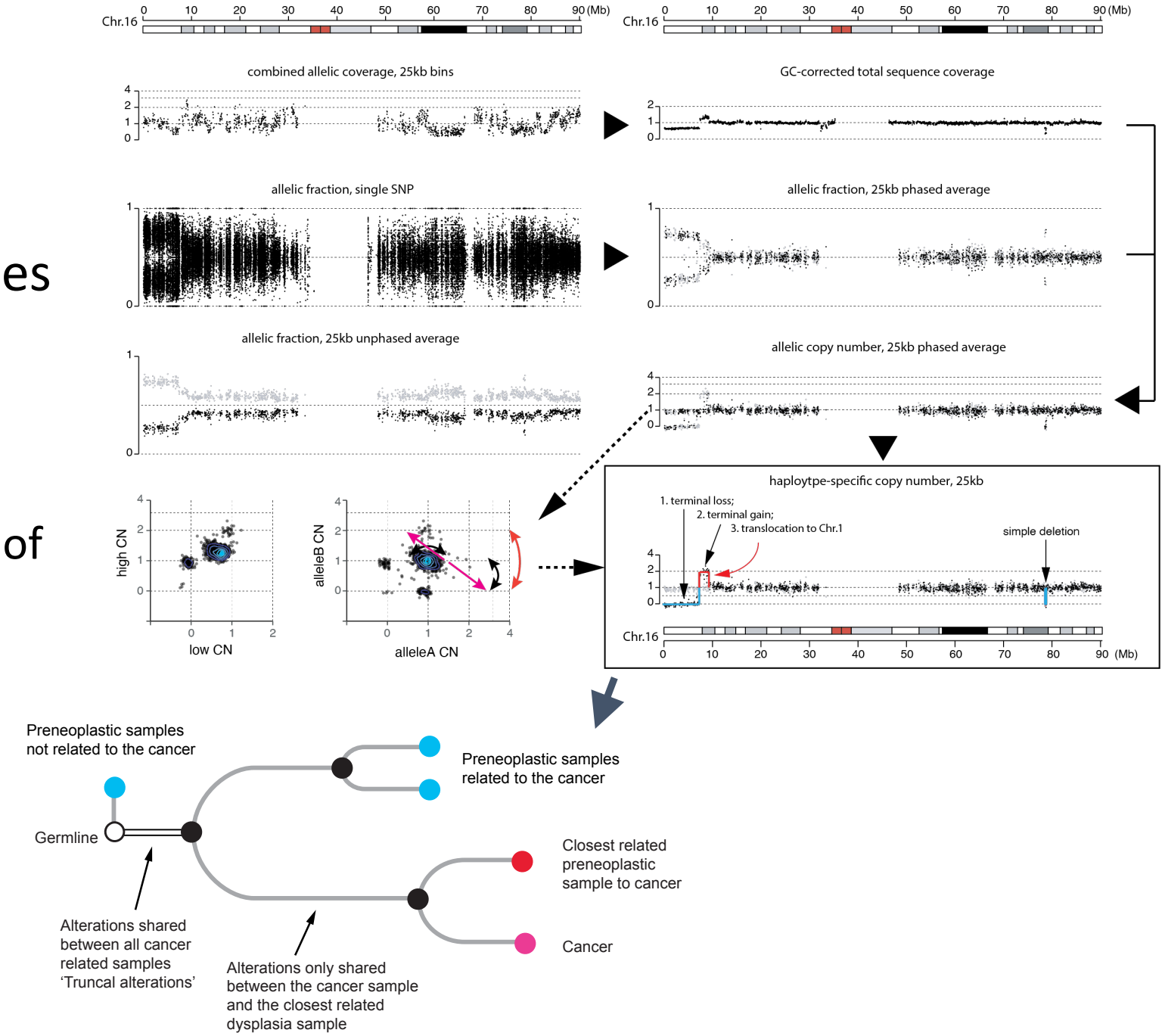


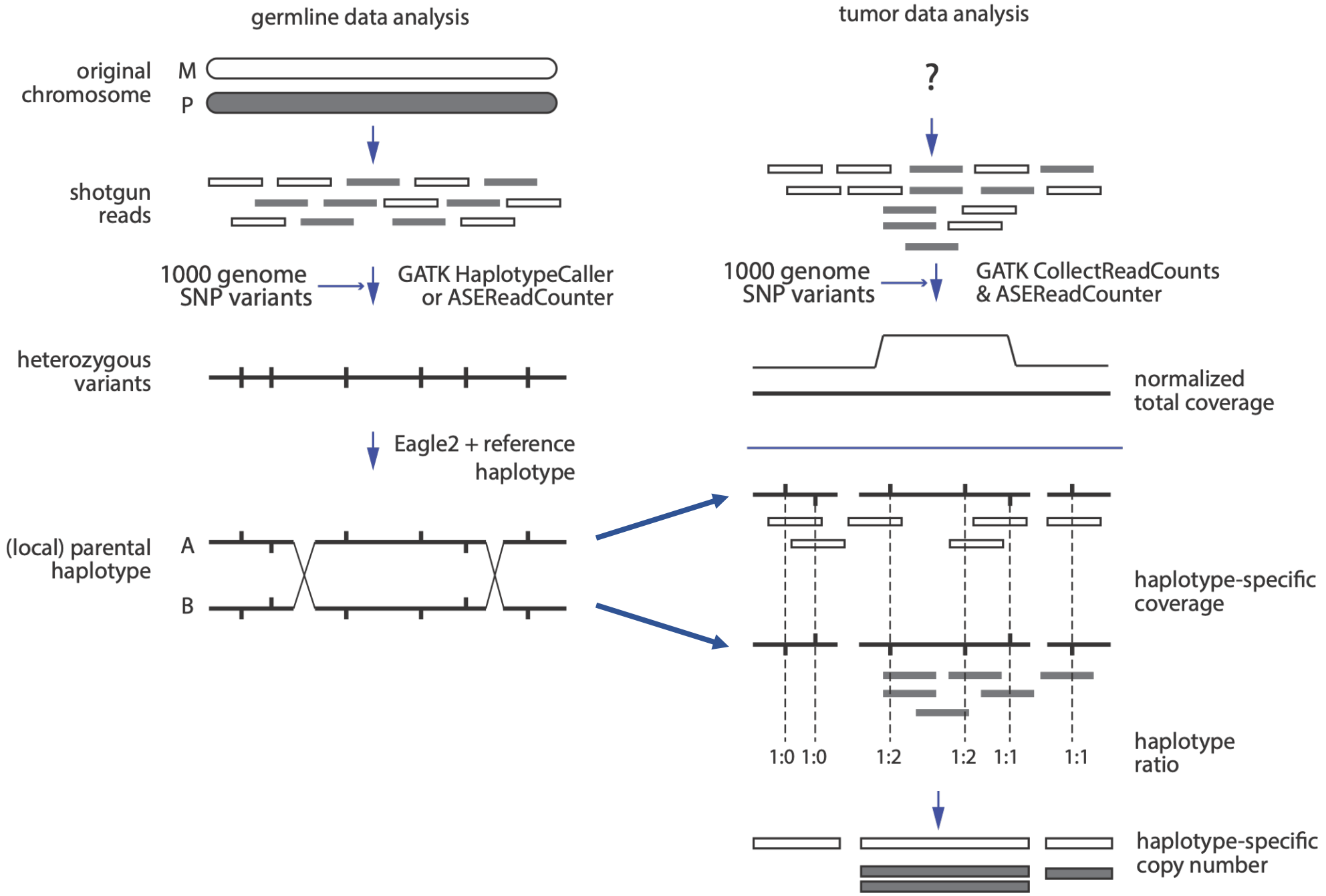
Cut second HGD

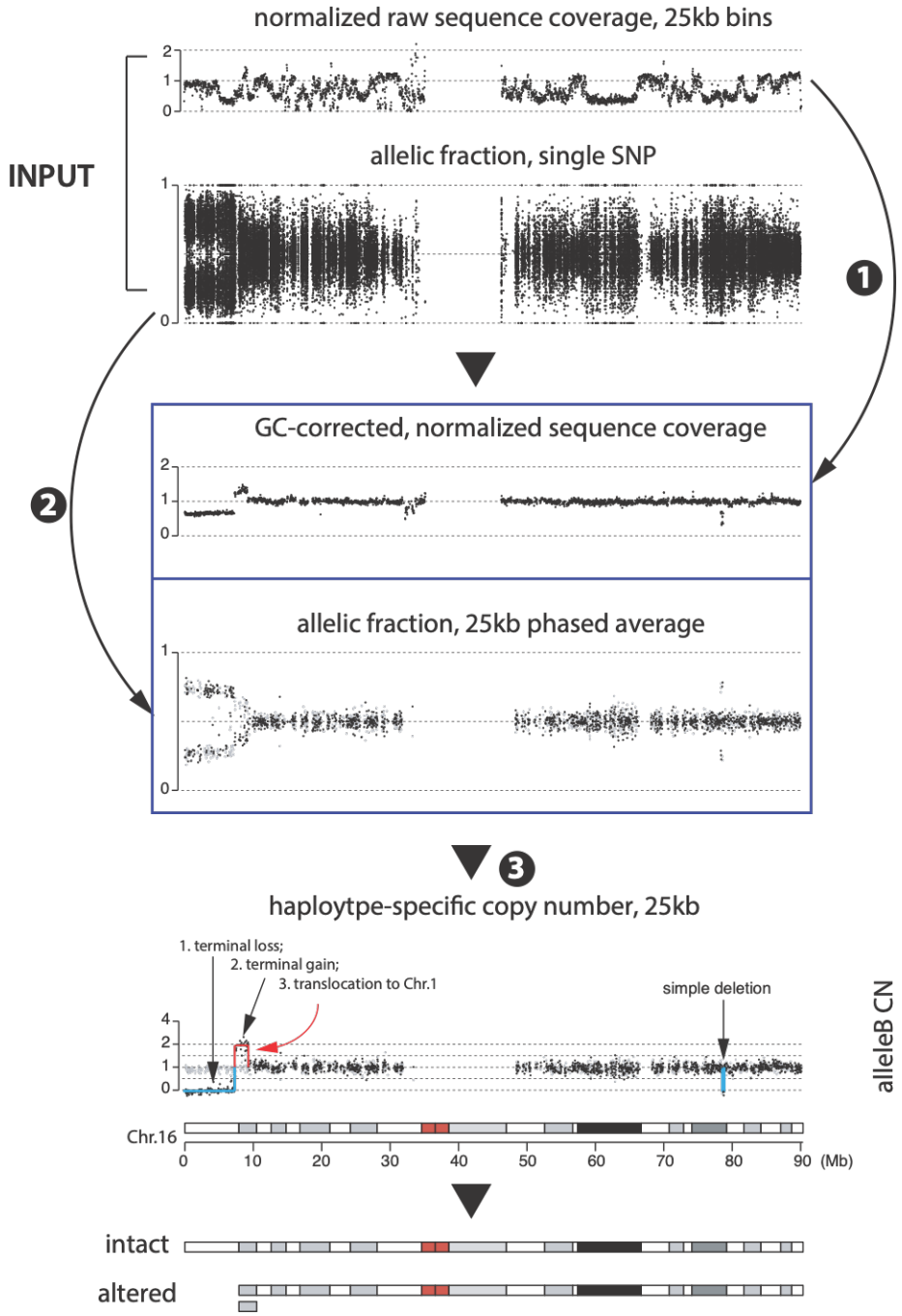


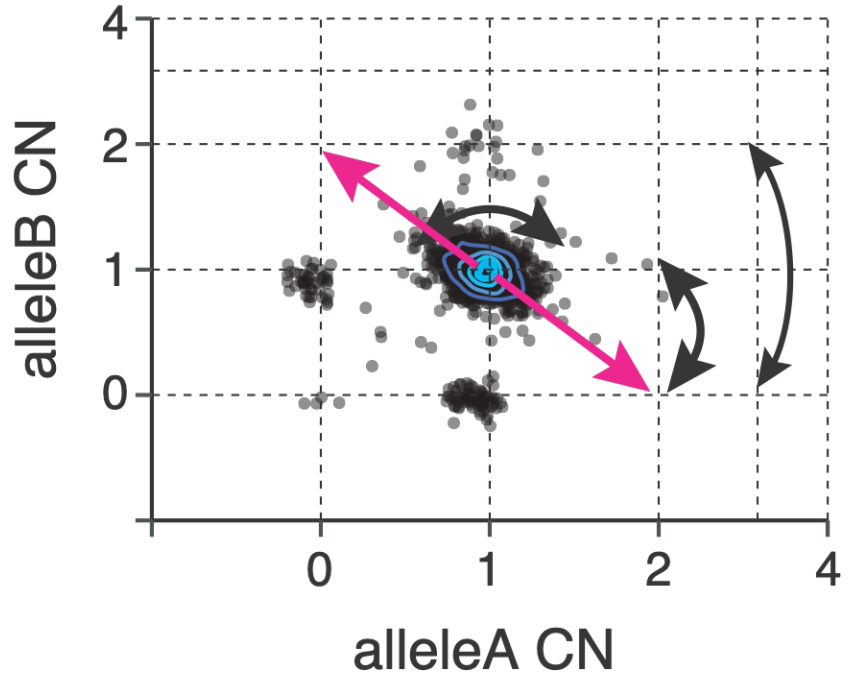
Workflow

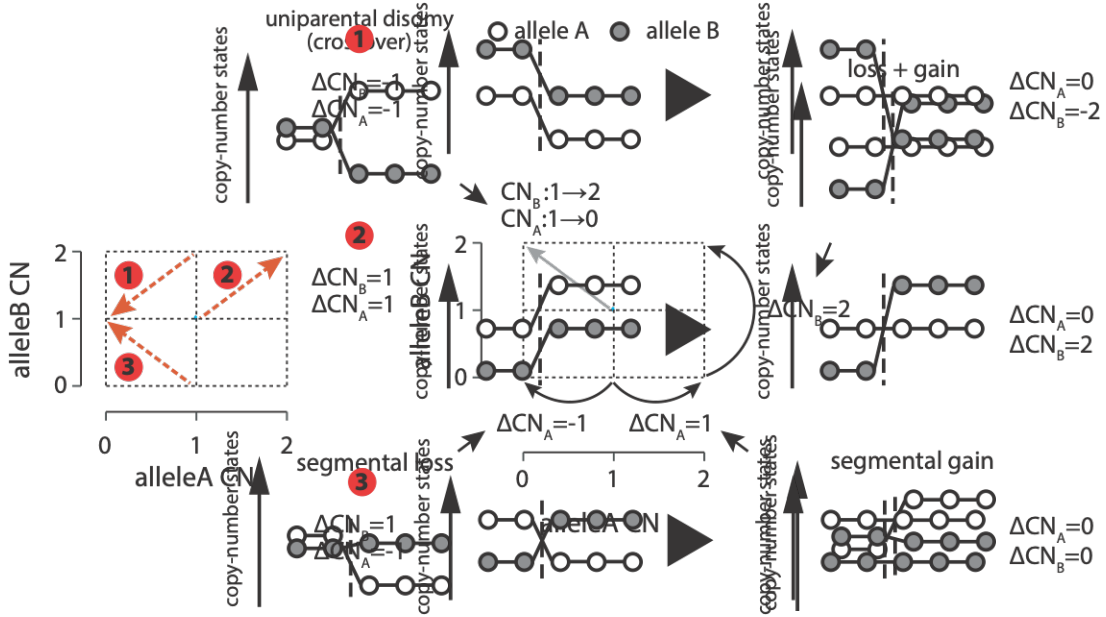
- Performed WGS on 76 samples + paired normal
- Haplotype-specific copy number calling
  - allowed a refined assessment of both sCNV and structural variants
- In combination with mutational calling, a detailed phylogenetic ‘tree’ for each patient was constructed



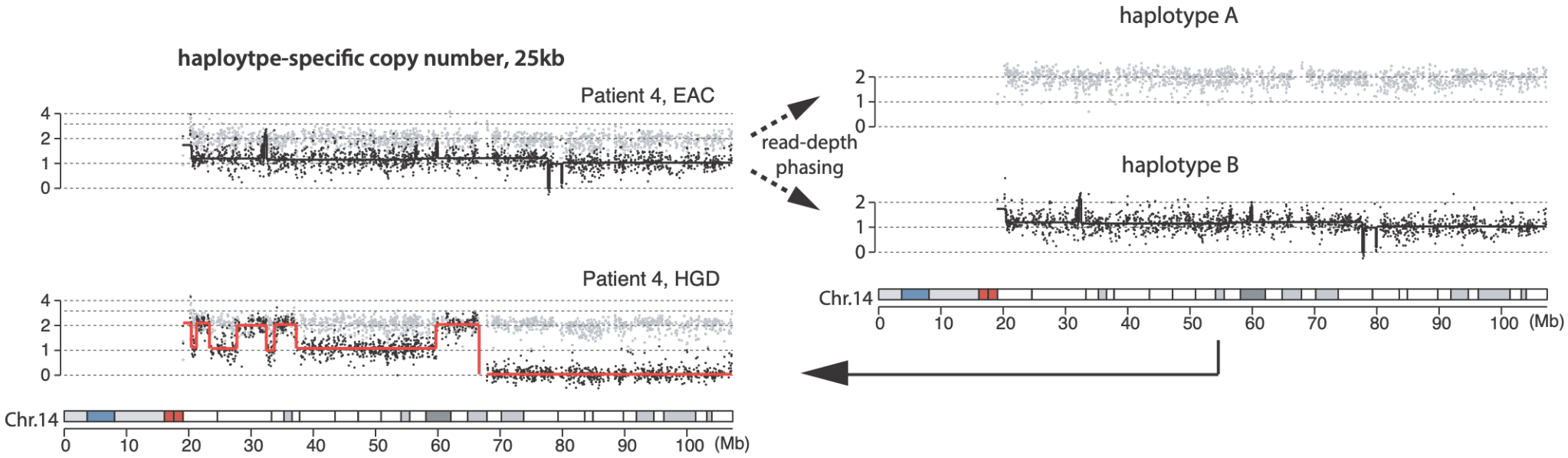

germline data analysis
tumor data analysis
original chromosome
M
P
shotgun reads
?
1000 genome SNP variants
GATK HaplotypeCaller or ASEReadCounter
1000 genome SNP variants
GATK CollectReadCounts & ASEReadCounter
heterozygous variants
normalized total coverage
Eagle2 + reference haplotype
(local) parental haplotype
A
B
haplotype-specific coverage
1:0 1:0
1:2
1:2
1:1
1:1
haplotype ratio
haplotype-specific copy number

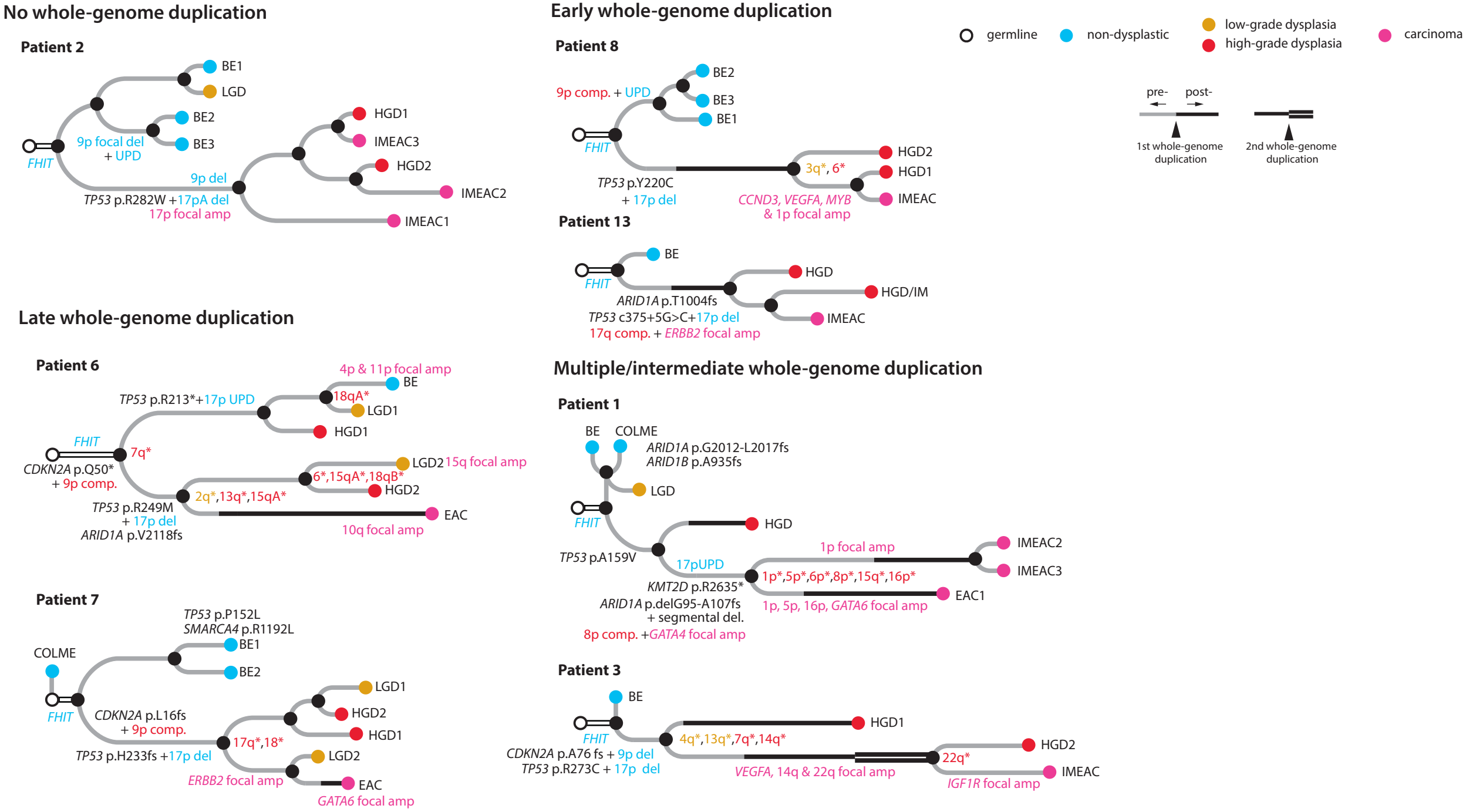

normalized raw sequence coverage, 25kb bins
INPUT
allelic fraction, single SNP
1
GC-corrected, normalized sequence coverage
2
allelic fraction, 25kb phased average
3
haploytpe-specific copy number, 25kb
1. terminal loss;
2. terminal gain;
3. translocation to Chr.1
simple deletion
Chr.16
0 10 20 30 40 50 60 70 80 90 (Mb)
intact
altered


alleleB CN
alleleA CN
0 1 2 4

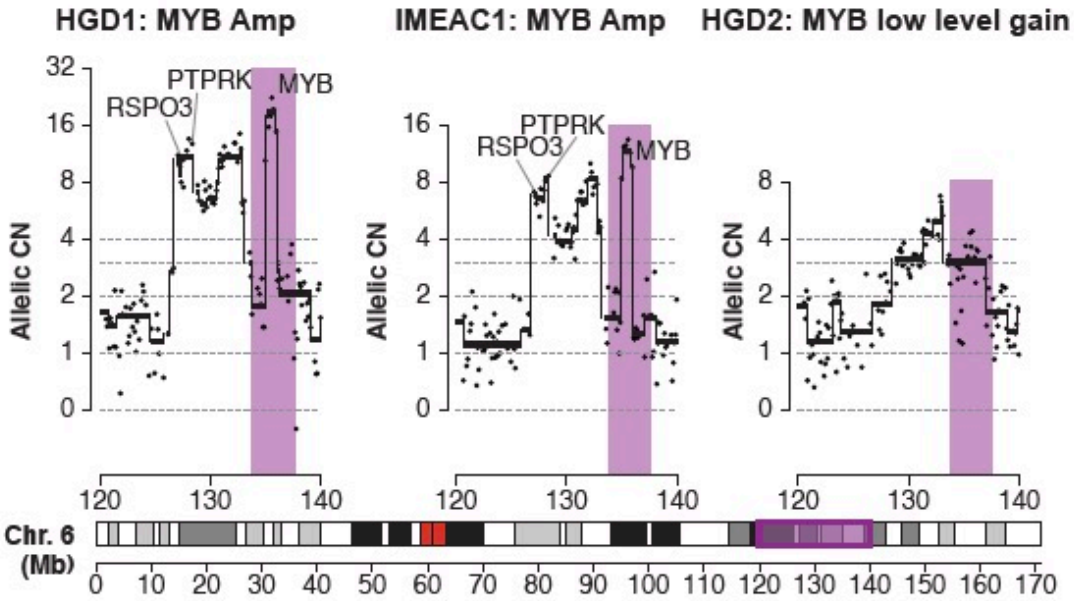

uniparental disomy (cross-over)
allele A
allele B
copy-number states
ΔCN_B=-1
ΔCN_A=-1
CN_B:1→2
CN_A:1→0
loss + gain
ΔCN_A=0
ΔCN_B=-2
alleleB CN
alleleA CN
ΔCN_B=1
ΔCN_A=1
ΔCN_B=2
ΔCN_A=0
ΔCN_B=2
ΔCN_A=-1
ΔCN_A=1
segmental loss
segmental gain
ΔCN_B=1
ΔCN_A=-1
ΔCN_A=0
ΔCN_B=0

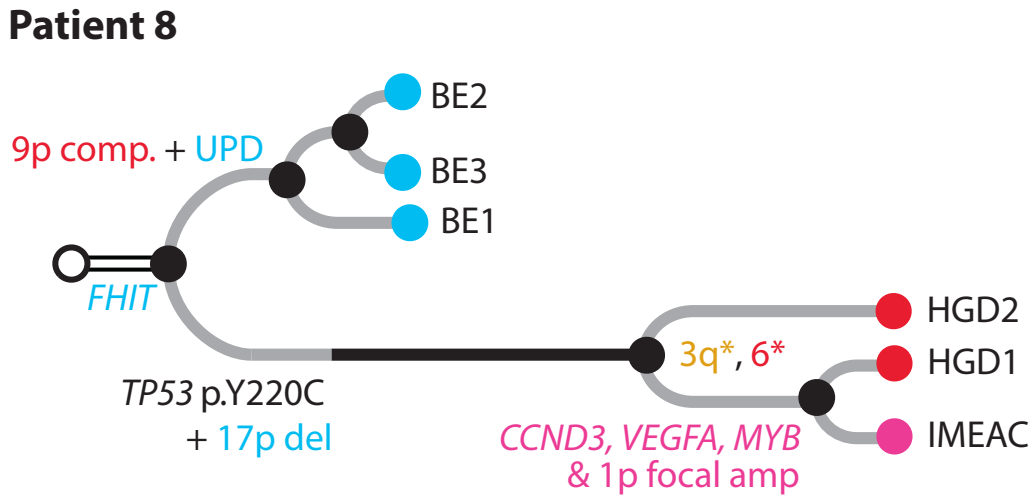
Multi-sample joint analysis




No whole-genome duplication
Patient 2
BE1
LGD
BE2
BE3
9p focal del
+ UPD
FHIT
HGD1
IMEAC3
HGD2
9p del
IMEAC2
TP53 p.R282W +17pA del
17p focal amp
IMEAC1
Early whole-genome duplication
Patient 8
BE2
9p comp. + UPD
BE3
BE1
FHIT
HGD2
3q*, 6*
HGD1
TP53 p.Y220C
+ 17p del
CCND3, VEGFA, MYB
& 1p focal amp
IMEAC
Patient 13
BE
FHIT
HGD
HGD/IM
ARID1A p.T1004fs
TP53 c375+5G>C+17p del
IMEAC
17q comp. + ERBB2 focal amp
germline
non-dysplastic
low-grade dysplasia
high-grade dysplasia
carcinoma
pre-
post-
1st whole-genome duplication
2nd whole-genome duplication
Late whole-genome duplication
Patient 6
4p & 11p focal amp
BE
TP53 p.R213*+17p UPD
18qA*
LGD1
HGD1
FHIT
7q*
CDKN2A p.Q50*
+ 9p comp.
LGD2
15q focal amp
6*,15qA*,18qB*
HGD2
2q*,13q*,15qA*
TP53 p.R249M
+ 17p del
ARID1A p.V2118fs
EAC
10q focal amp
Patient 7
TP53 p.P152L
SMARCA4 p.R1192L
COLME
BE1
BE2
LGD1
HGD2
FHIT
CDKN2A p.L16fs
+ 9p comp.
HGD1
17q*,18*
TP53 p.H233fs +17p del
LGD2
ERBB2 focal amp
EAC
GATA6 focal amp
Multiple/intermediate whole-genome duplication
Patient 1
BE
COLME
ARID1A p.G2012-L2017fs
ARID1B p.A935fs
LGD
FHIT
HGD
TP53 p.A159V
1p focal amp
IMEAC2
17pUPD
IMEAC3
1p*,5p*,6p*,8p*,15q*,16p*
KMT2D p.R2635*
EAC1
ARID1A p.delG95-A107fs
+ segmental del.
1p, 5p, 16p, GATA6 focal amp
8p comp. +GATA4 focal amp
Patient 3
BE
HGD1
FHIT
4q*,13q*,7q*,14q*
CDKN2A p.A76 fs + 9p del
TP53 p.R273C + 17p del
HGD2
22q*
VEGFA, 14q & 22q focal amp
IMEAC
IGF1R focal amp

Oncogenic high level amplifications are present in the most closely related area of dysplasia

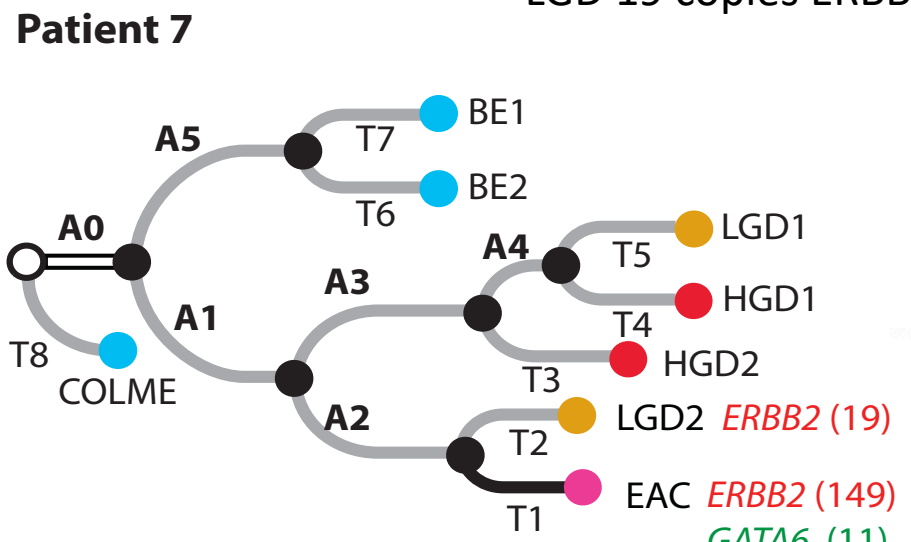


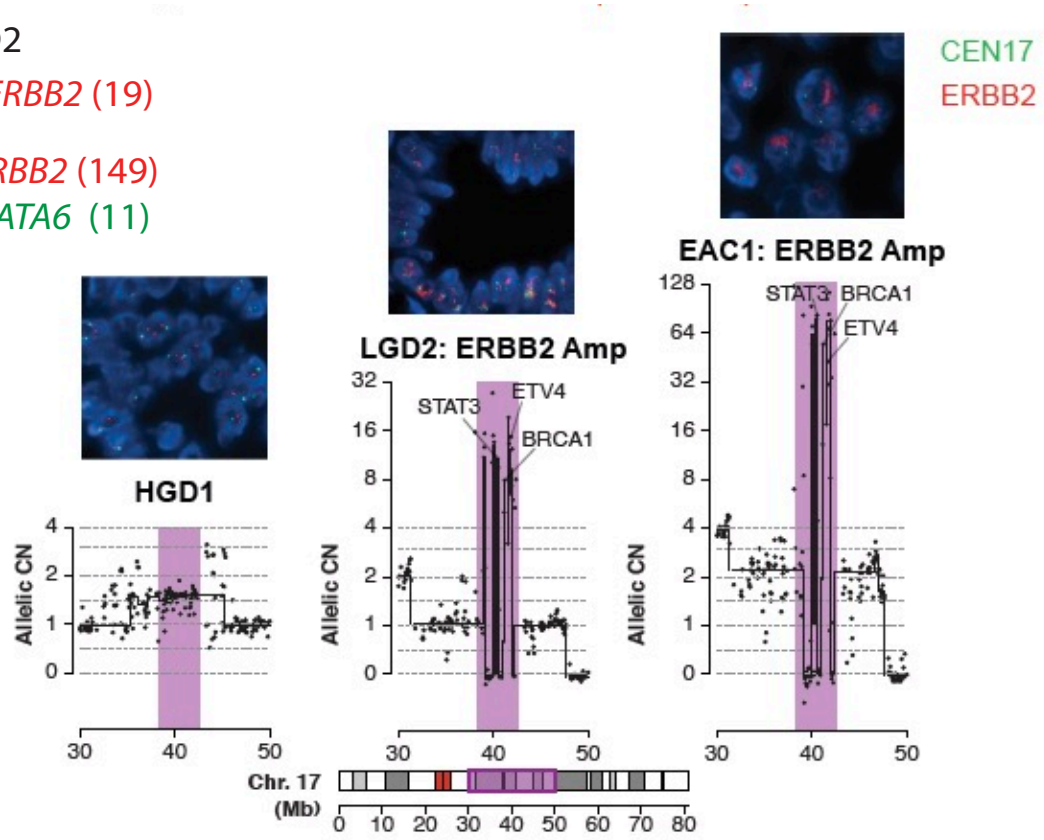


- 10/15 Pts focal amp involving at least one oncogene was shared between cancer and most closely related dysplasia
- Commonly genes encoding receptor tyrosine kinases (RTKs), including EGFR, FGFR2, ERBB2, and other oncogenes including MYB, CDK6, MYC, GATA4 and GATA6
- **4/5 Pts with multiple dysplasia, only found in most closely related to cancer**

Genomic evolution from Dysplasia to EAC can still occur

LGD 19 copies ERBB2 > EAC 100 copies



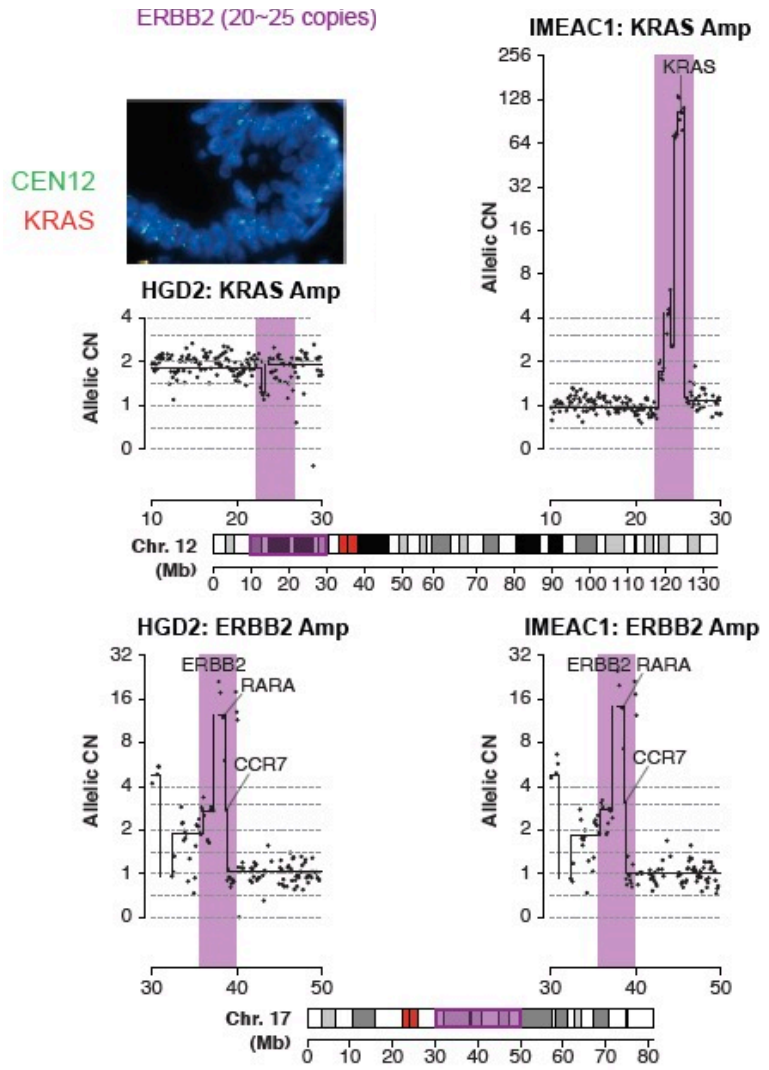


- 7/15 pts clear evidence of evolution between the cancer sample and most closely related dysplasia
- 3/7 increase copy number
- 4/7 gain of new amplification

Genomic evolution from Dysplasia to EAC can still occur

HGD ERBB2, CDK6 amp > EAC ERBB2, CDK6, KRAS amp

Patient 10

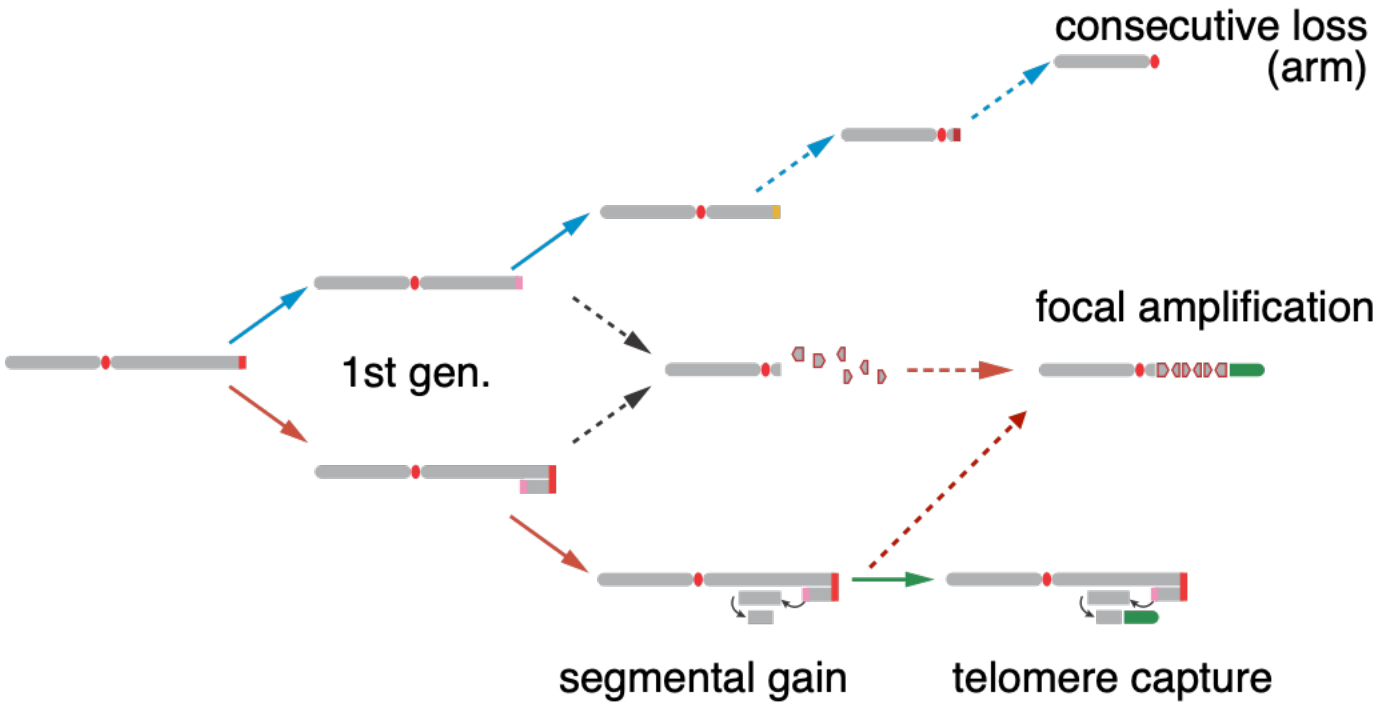


- 7/15 pts clear evidence of evolution between the cancer sample and most closely related dysplasia
- 3/7 increase copy number
- 4/7 gain of new amplification

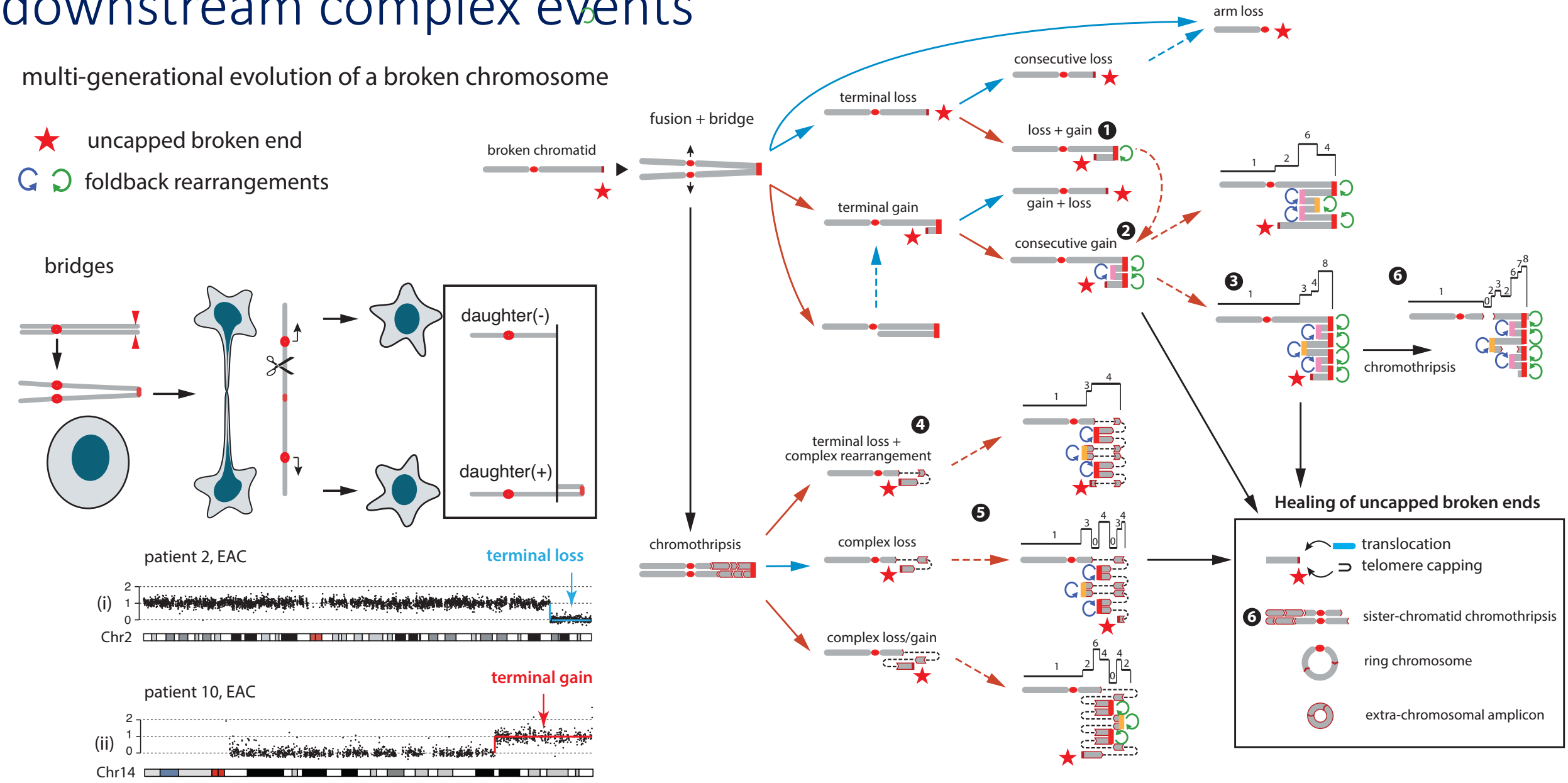
Early genomic evolution may explain BE evolution and transformation into EAC

Umbreit NT, Zhang CZ, Lynch LD, et al. **Mechanisms generating cancer genome complexity from a single cell division error.** *Science*. 2020;368(6488):eaba0712. doi:10.1126/science.aba0712

- Recent paper suggests a single CNV event such as deletion of the chromosomal end can cause a BFB and lead to multiple downstream complex CNVs in only a couple of cell divisions
  - Followed damaged cells in tissue culture
  - Showed through BFB simple gains or losses **can** quickly evolve into more catastrophic events
  - A single terminal deletion can initiate the cascade

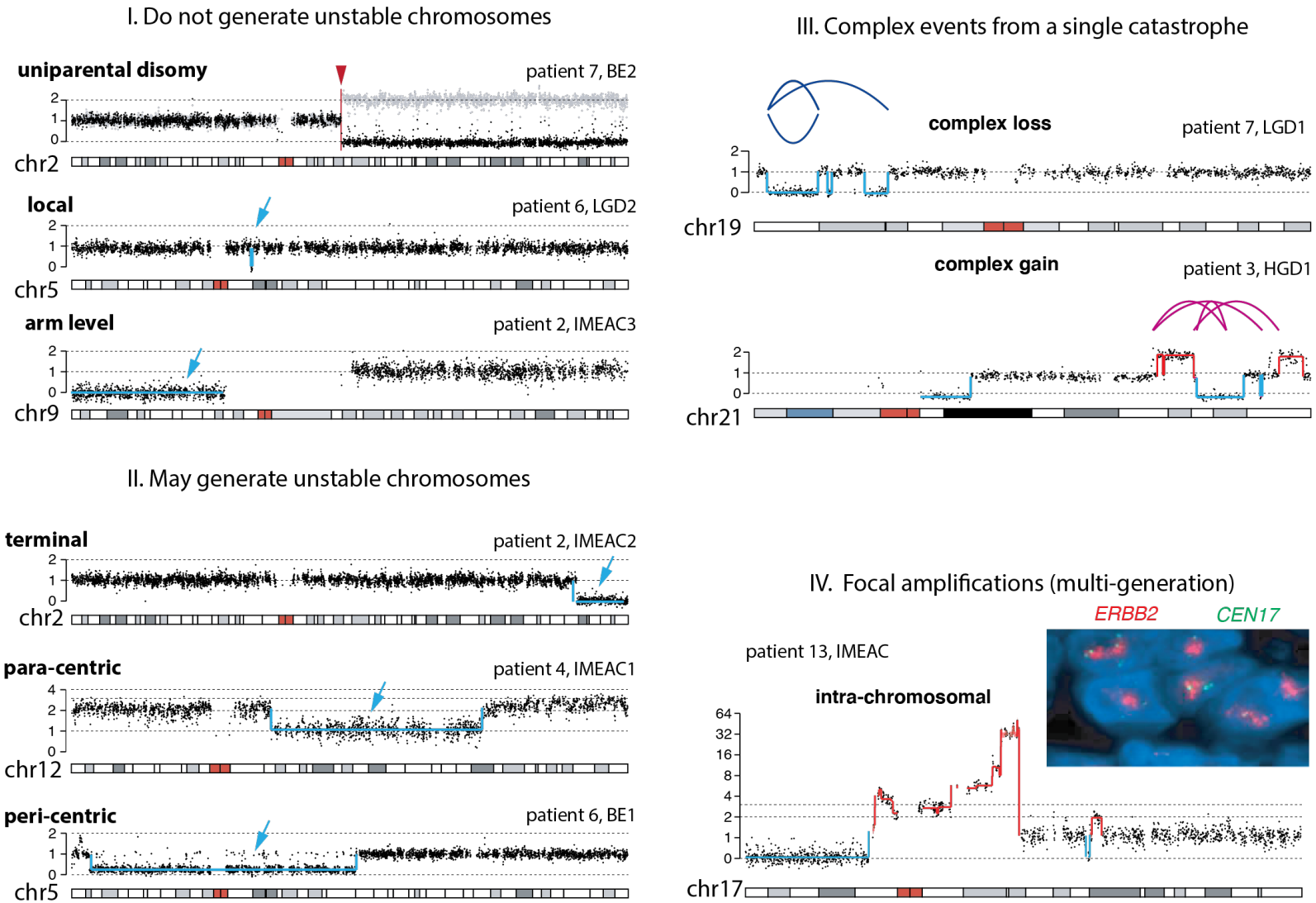


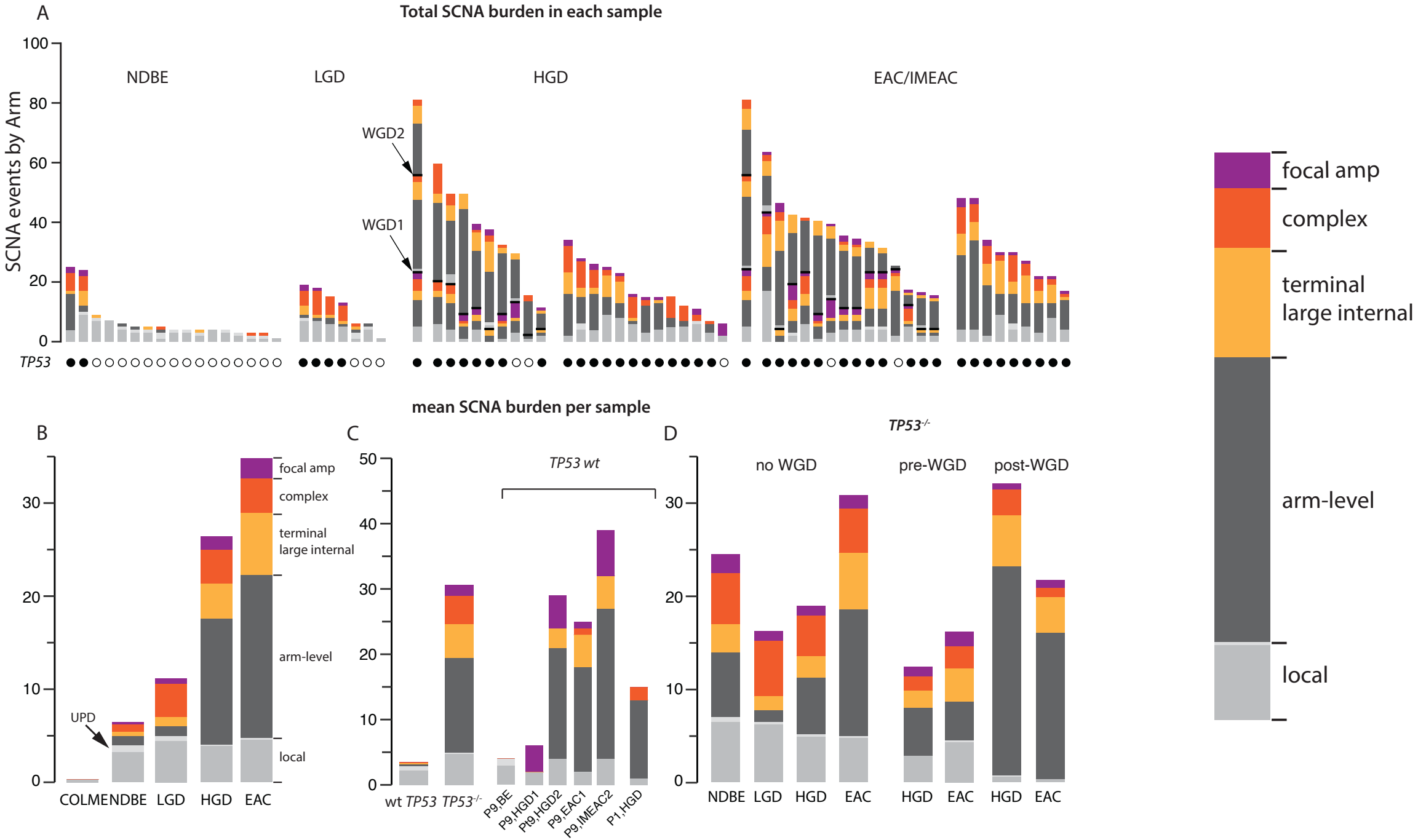
A single, simple copy number alteration can lead to more downstream complex events



Classification of copy-number alterations

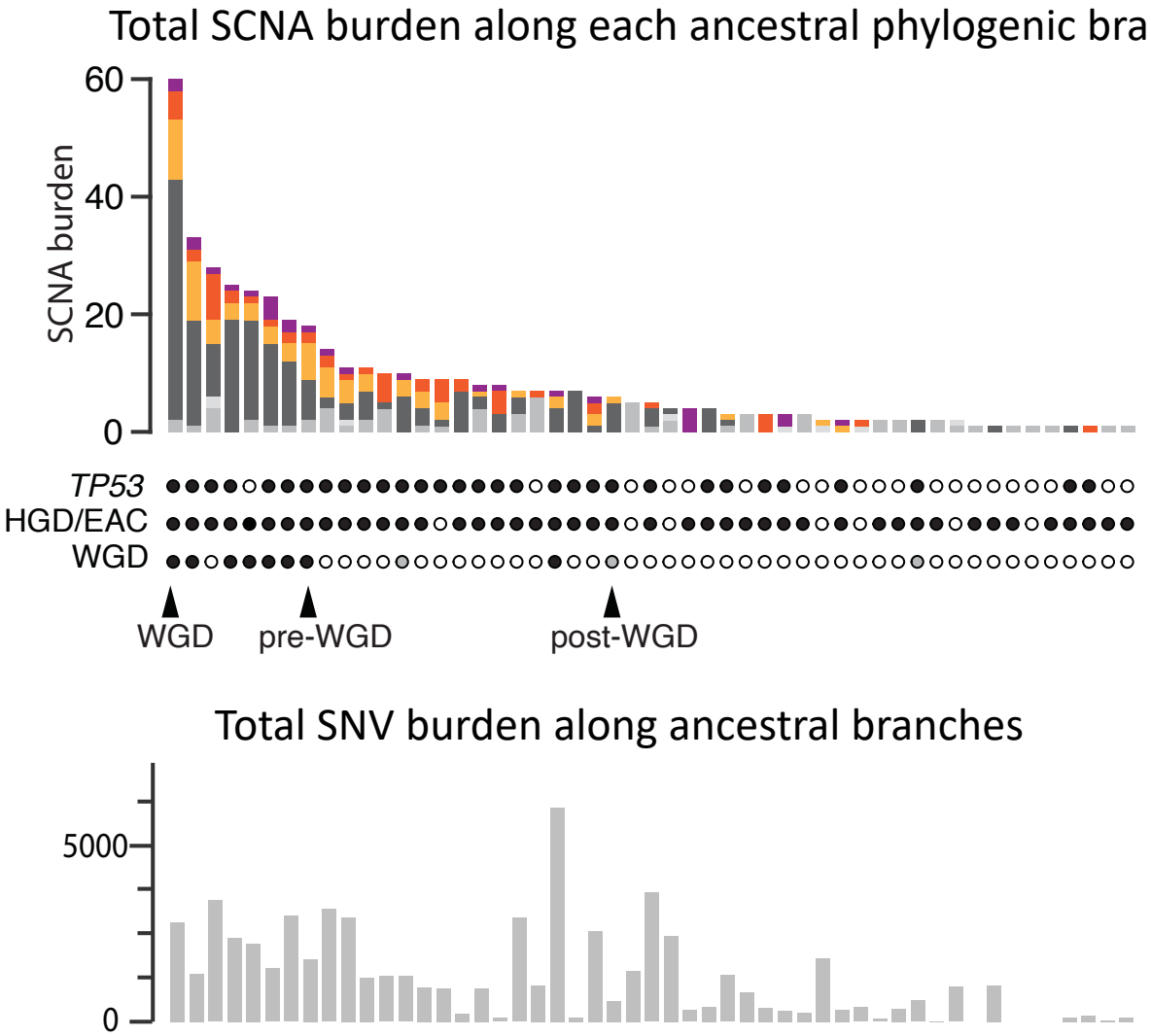
- Previous studies have shown some types of CNVs can lead to further CIN in an in vitro setting
- Unknown how these mechanism may influence genomic progression in BE




A
Total SCNA burden in each sample
SCNA events by Arm
NDBE
LGD
HGD
EAC/IMEAC
WGD2
WGD1
TP53
mean SCNA burden per sample
B
C
D
TP53 wt
TP53-/-
no WGD
pre-WGD
post-WGD
UPD
COL
NDBE
LGD
HGD
EAC
wt TP53
P9,BE
P9,HGD1
P8,HGD2
P9,EAC1
P9,IMEAC2
P1,HGD
focal amp
complex
terminal
large internal
arm-level
local

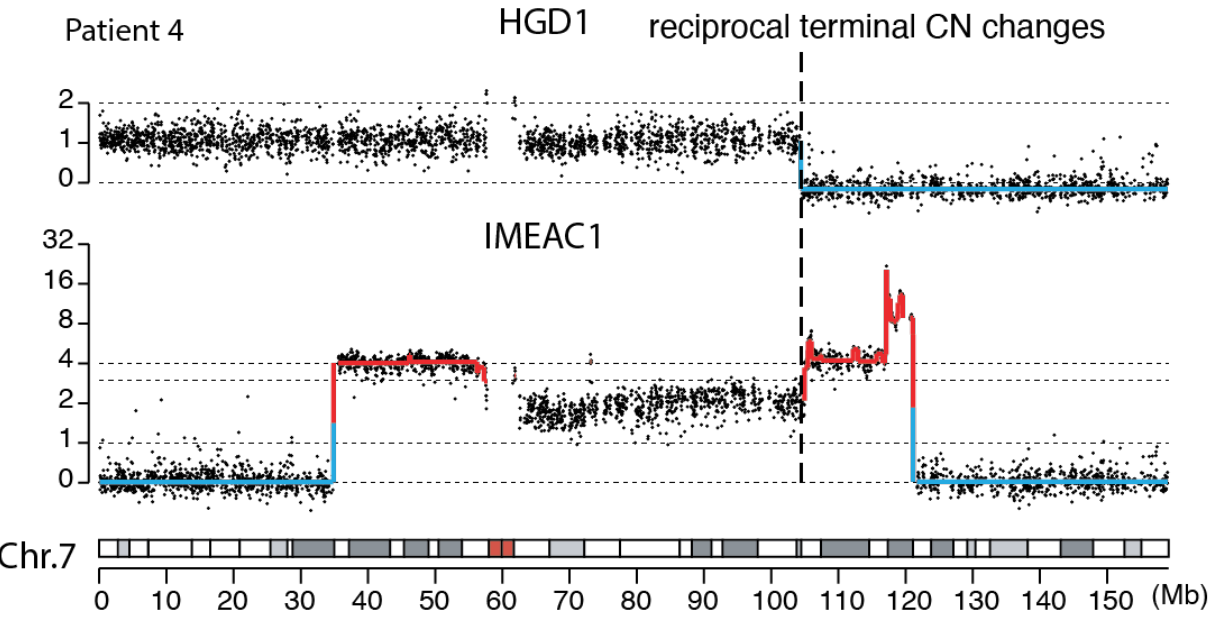
BE genome evolution is driven by both episodic and continuous genomic instability

- Number of CNVs on each Phylogenetic branches does not correlate well with the number of SNVs
- Suggests CNVs may occur in episodic bursts



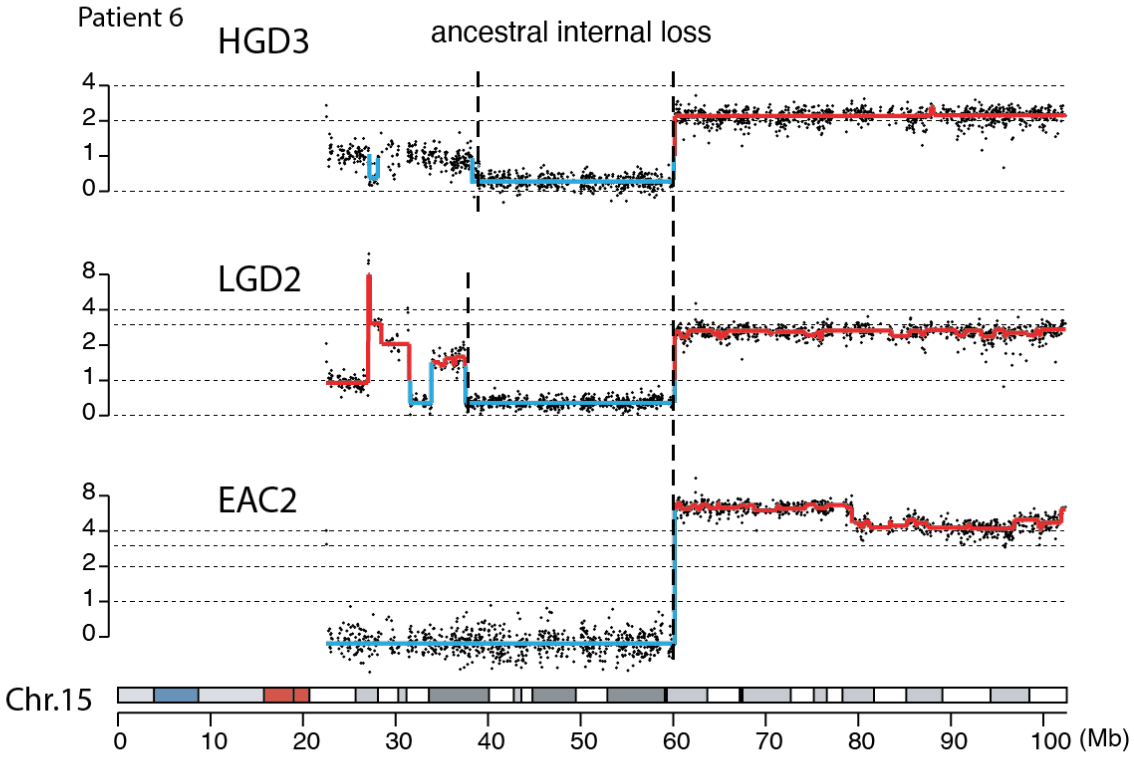
BE genome evolution is driven by both episodic and continuous genomic instability

Evidence 1: Divergent copy-number alterations present in related dysplastic lesions that can be traced to a single ancestral unstable chromosome



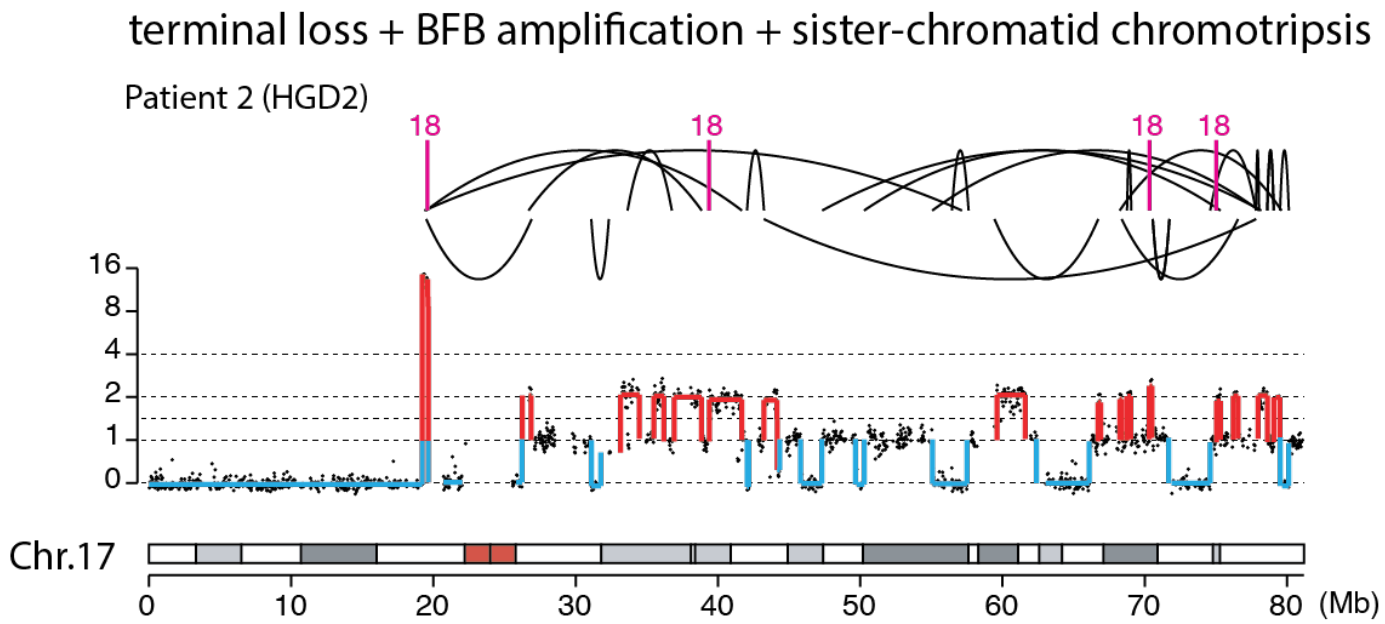
Evidence 2: Resolution of chromosome bridges can also generate more complex copy-number changes including chromothripsis

Evidence 3: Evidence for complex alterations arising downstream of ancestral chromosome breaks also came from examples of divergent genome evolution with complex genomic alterations seen in one progeny clone that shared terminal or internal copy-number changes with other clones



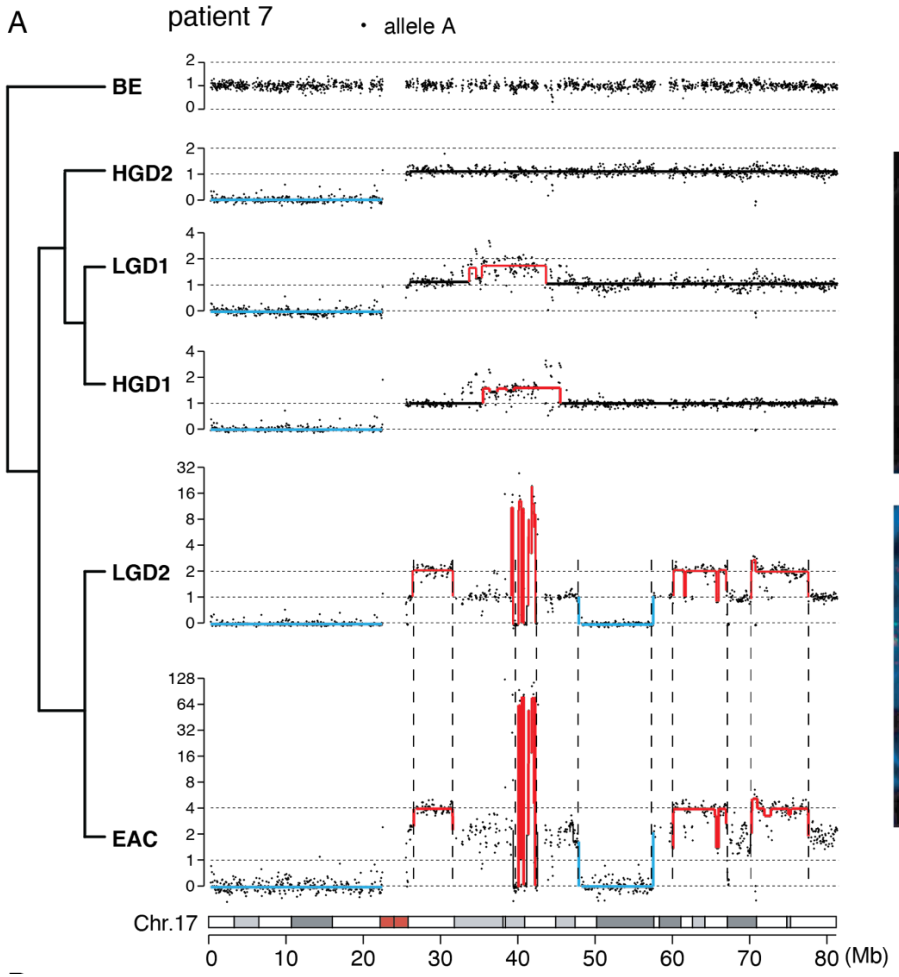
BE genome evolution is driven by both episodic and continuous genomic instability

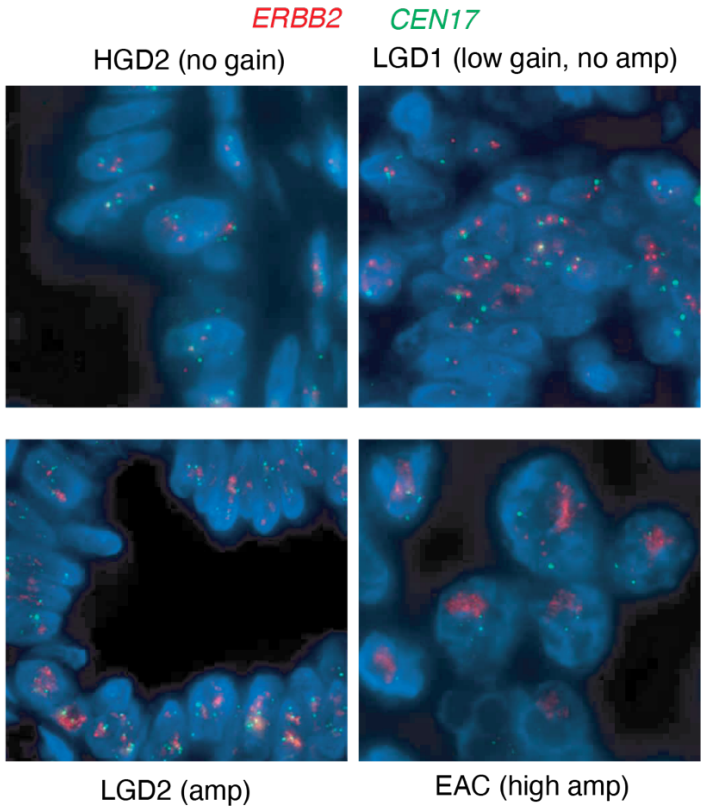
293 chromosomes with complex copy-number patterns (defined as more than 2 copy-number changes), 58% also had preceding terminal copy-number changes.



Evolution and polyclonal expansion of dysplastic BE lesions

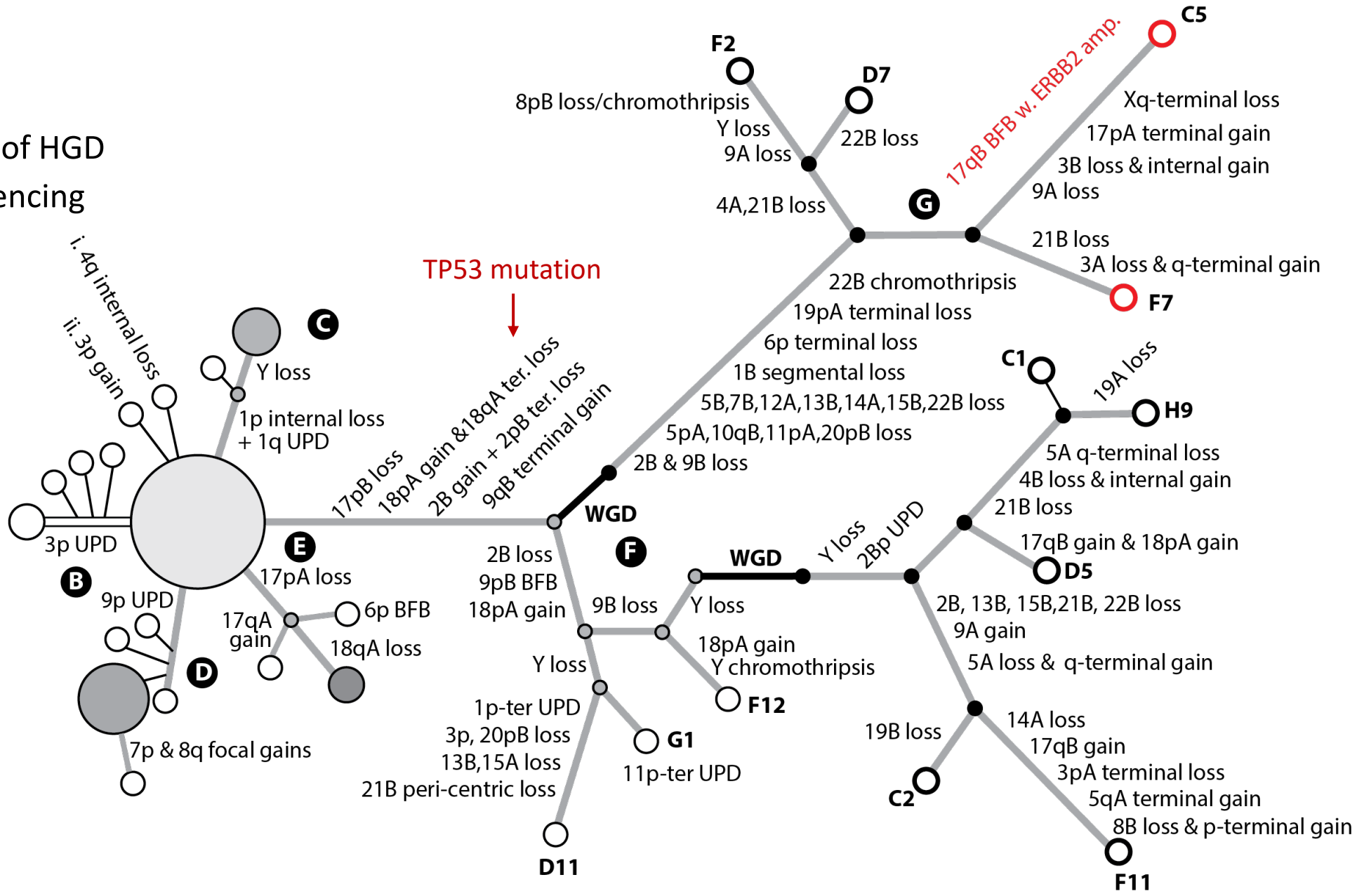
- Through episodic and more continuous/multi-step processes we see a picture of continuous branching evolution which creates complex genomic heterogeneity throughout the patient's entire field of BE





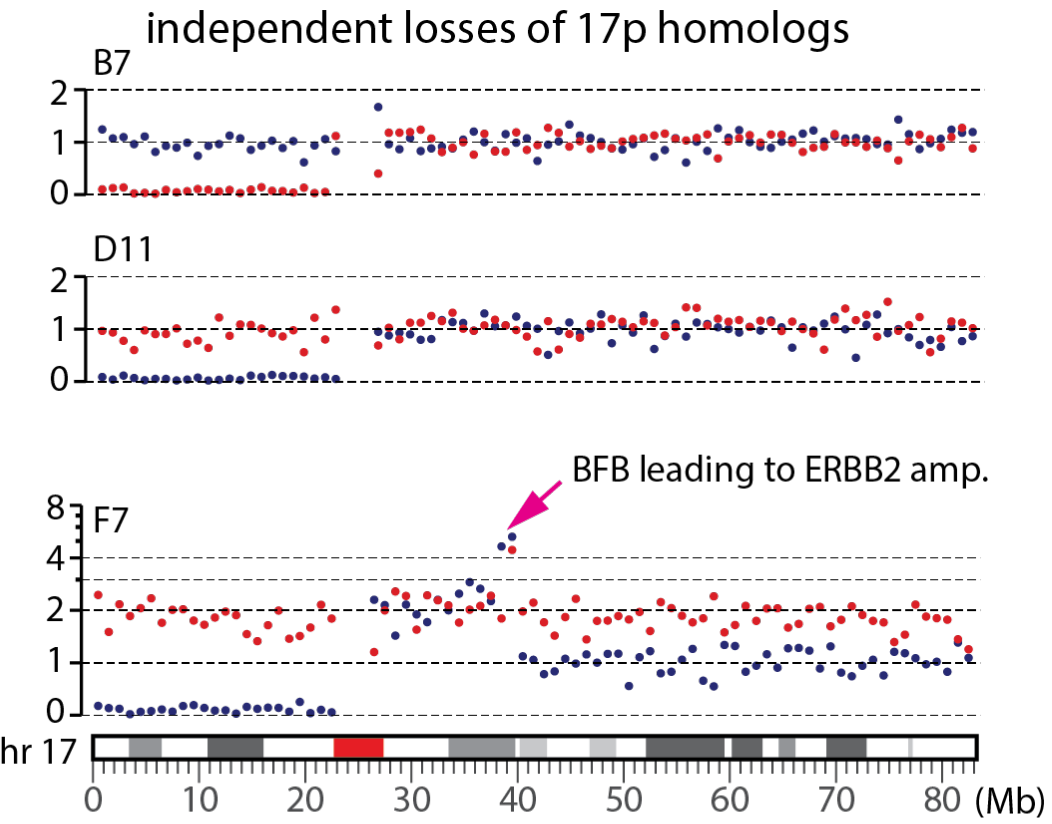
Evolution and polyclonal expansion of dysplastic BE lesions

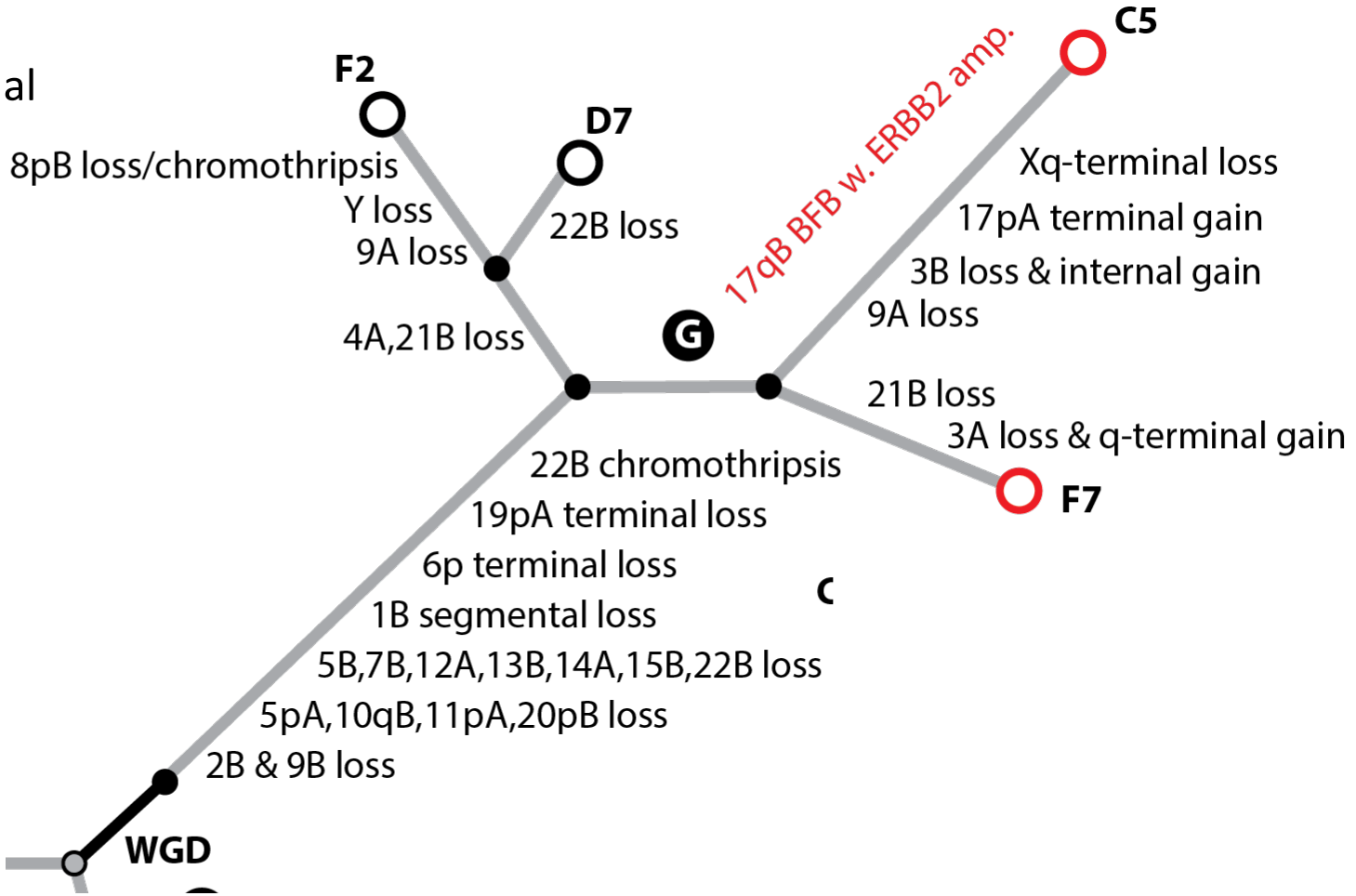
- Single pt with HGD
  - Endoscopic brushing of HGD
  - Single cell DNA sequencing
  - ~1x WGS



Evolution and polyclonal expansion of dysplastic BE lesions

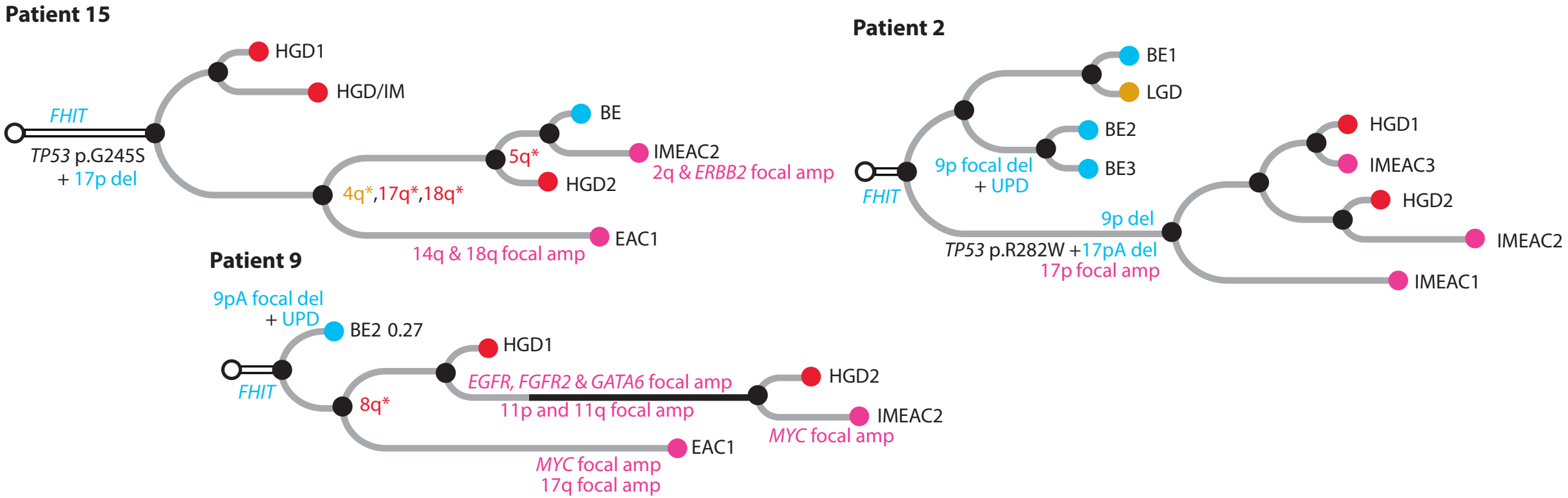
- Genomic alterations highly heterogeneous (especially after loss of TP53
- Able to see individual cells with same chromosomal patterns as other 15 patients





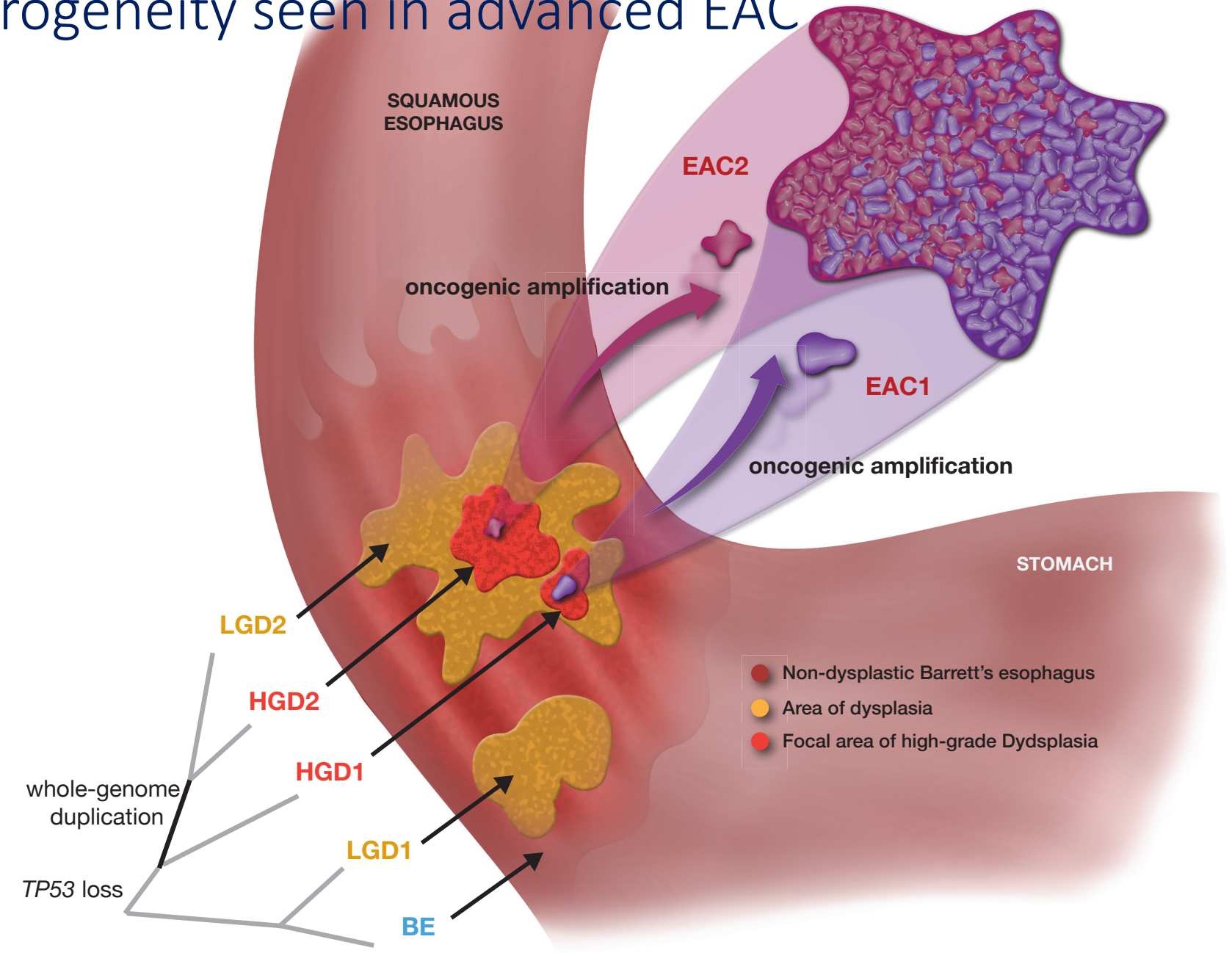
Evolution and polyclonal expansion of dysplastic BE lesions leads to multiple independent transformations to EAC

- In five patients with multiple areas of intramucosal (IMEAC) or early adenocarcinoma, the EAC lesions showed significant genomic divergence and were inferred to be in separate evolutionary branches.
- 4/5 patients had pre-cancerous samples more closely related to a cancer sample then the cancer samples were to each other.
- Strongly suggest that the progression from dysplasia to adenocarcinoma occurred independently within one or more fields of dysplasia

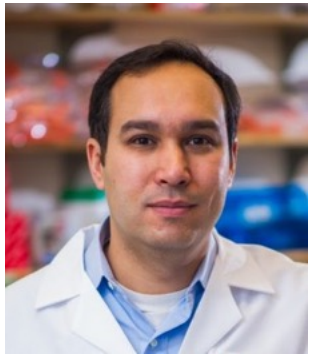


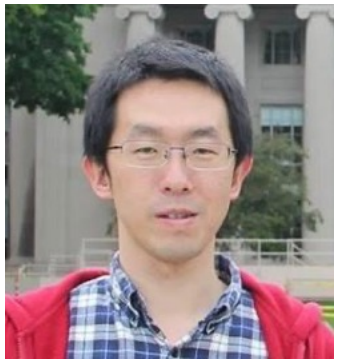
Early genomic heterogeneity and multiple transformations help explain oncogenic driver heterogeneity seen in advanced EAC

			Primary	Metastasis	
PANGEA 1	*TP53*	NGS	R175H	R175H	
	EGFR	MS	POS		
	MET	NGS	NEG	AMP	
PANGEA 2	*TP53*	NGS	R175H	R175H	
	KRAS	NGS	AMP	AMP	
	ERBB2	NGS	NEG	AMP	
	ERBB2	IHC/FISH	NEG	POS	
PANGEA 3	*TP53*	NGS	R175H	R175H	
	ERBB2	NGS	AMP	NEG	
	ERBB2	IHC/FISH	POS	NEG	NEG
	EGFR	NGS	AMP	AMP	
	EGFR	IHC/FISH	POS		POS
PANGEA 4	*TP53*	NGS	R273C	R273C	
	EGFR	NGS		AMP	
PANGEA 5	*TP53*	NGS	R282G	R282G	
	ERBB2	NGS	NEG	AMP	AMP
	ERBB2	IHC/FISH	NEG	AMP	AMP
PANGEA 6	*TP53*	NGS	C176F		
	FGFR2	NGS	NEG		
PANGEA 7	*TP53*	NGS	N239D	N239D	
	ERBB2	NGS	AMP	NEG	
	ERBB2	IHC/FISH	POS	NEG	
	FGFR2	NGS	NEG	AMP	
PANGEA 8	*EGFR*	NGS	POS	NEG	
	EGFR	MS		POS	
PANGEA 9	*ARID1A*	NGS	R1335insQ	R1335insQ	
	TP53	NGS	R342*	NEG	
	EGFR	NGS	AMP	NEG	
	EGFR	NGS	FUSION	NEG	
	KRAS	NGS	NEG	G12D	
PANGEA 10	*TP53*	NGS	R306*	R306*	
	FGFR2	NGS	FUSION	NEG	



A **truly** collaborative effort: Thanks!


Adam Bass


Chunyang Bao


Cheng-Zhong Zhang

- Bass lab (DFCI Oncology)
  - **Adam Bass**
  - Fahire Goknur Akarca
  - **Chunyang Bao**
- **Cheng-Zhong Zhang** (BROAD/DFCI)
- Getz Lab (BROAD Institute)
- GI Pathology (BWH GI Pathology)
  - Robert Odze (Now CDx diagnostics)
  - Amitabh Srivastava
  - Tony Agoston
- Sequencing:
  - Center for Cancer Genome Discovery (DFCI)
  - BROAD Institute (Harvard/MIT)

Tissue:

- **Kenneth Wang (Mayo)**
- **Jon Davison and Katie Nason (UPMC)**
- Mark Redston (BWH)
- Jacques Bergman (AUMC)

Funding:
NIDDK K08
Doris Duke Charitable Foundation
Barrett's Esophagus Translational Research Network (BETRNet) Seed Grant
NHGRI/BROAD
BROAD SPARC, Next10
NIH/Harvard Cancer Center GI SPORE YIA
BWH Pathology/Harvard Genetics T32