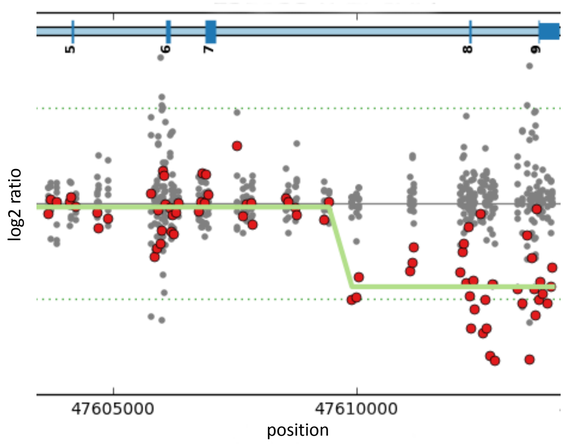
MLPAseq: Assaying genome copy number variation using MLPA paired with high-throughput sequencing

Daniel J Kvitek, Sarah A McCalmon, Sylvia Lin, Ben Hsieh, Nathan McDonald, Kevin B Jacobs, Joshua S Paul, Eric C Olivares
Invitae Corporation, San Francisco, CA, 94107 www.invitae.com

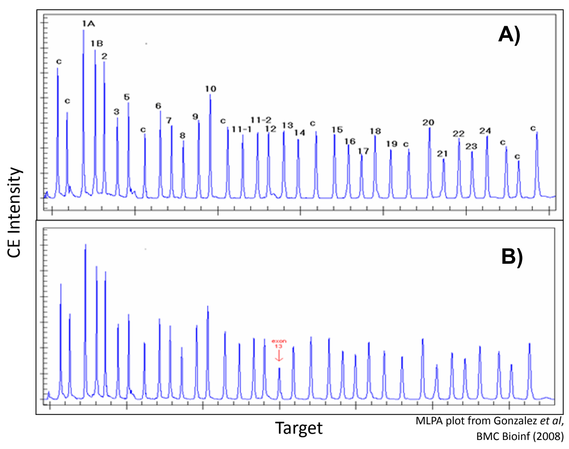
Why MLPAseq?

- Large-scale gene panels are becoming commonplace in clinical genetic diagnostics
- Current clinically accepted techniques for copy number variation (CNV) detection are not ideal due to technical or throughput limitations:



Array-based comparative genomic hybridization (aCGH) data are noisy and suffer from cross-hybridization between similar target sequences

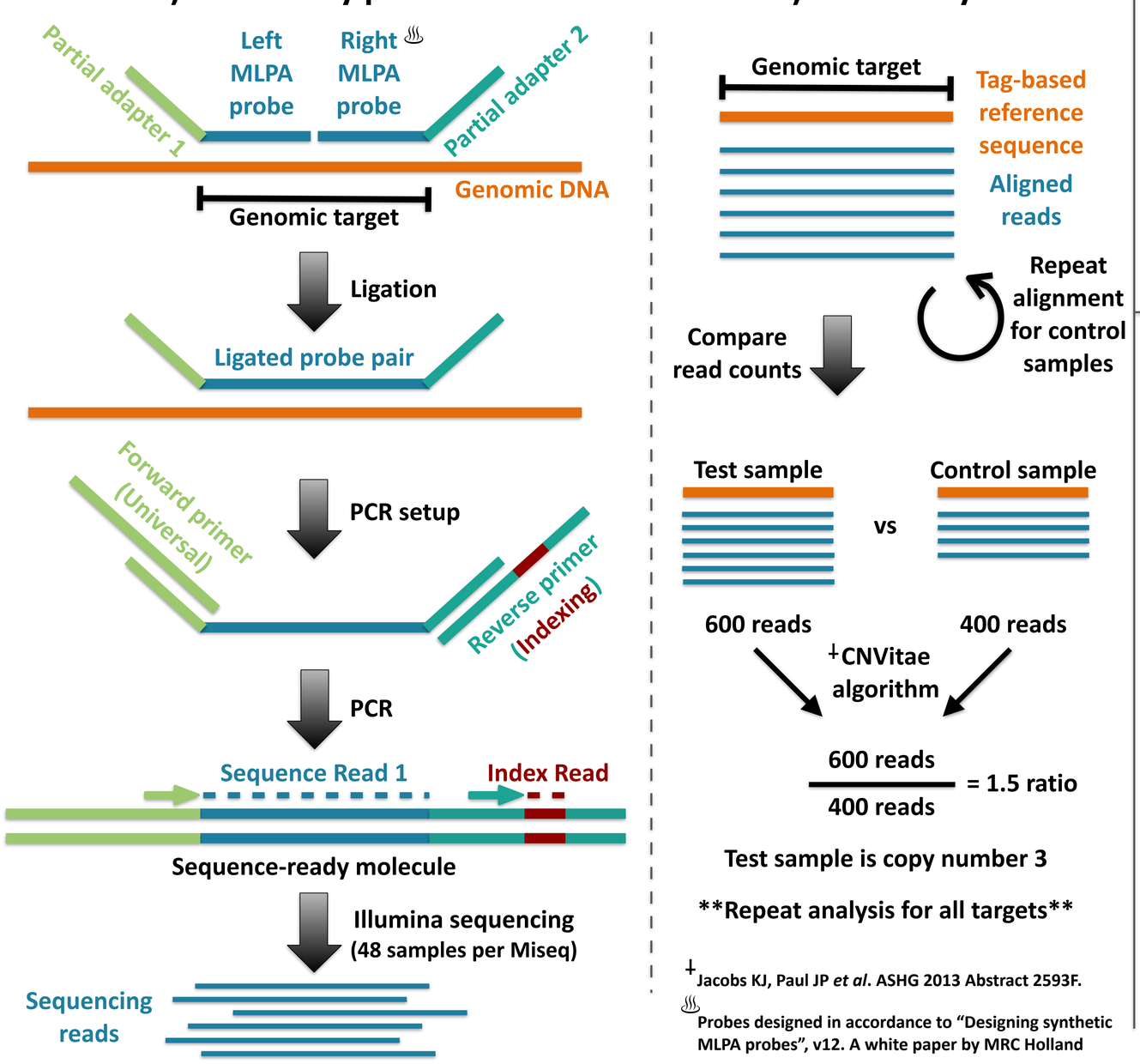
Standard multiplex ligation-dependent probe amplification (MLPA) is less noisy than aCGH, but is low-throughput, limited to 40-50 targets per reaction. Data analysis is not scalable and is difficult to automate.



MLPA plot from Gonzalez *et al*, BMC Bioinf (2008)

- We developed a **1000** target fully-automated MLPAseq assay and ran it on **80** samples with known CNVs, confirmed by hybridization capture and/or aCGH

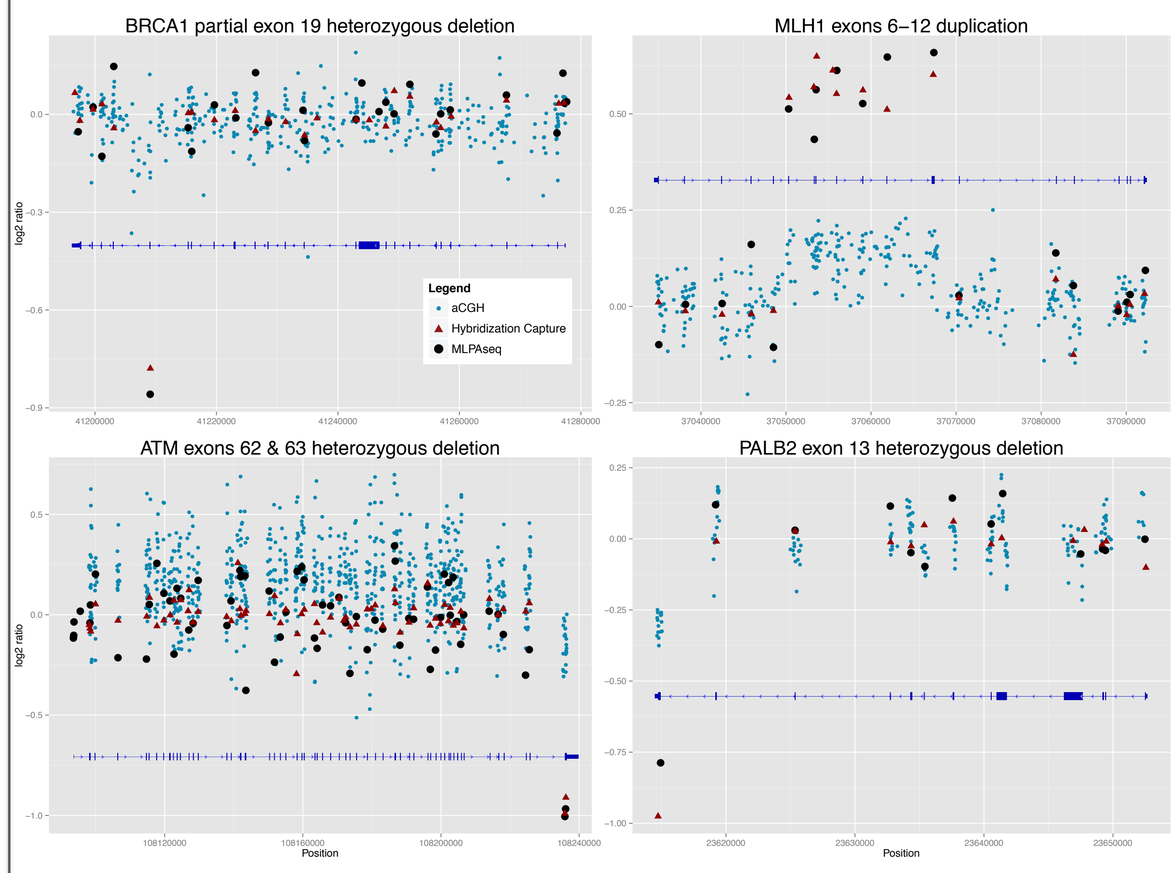
MLPAseq: The Method



† Jacobs KJ, Paul JP *et al*. ASHG 2013 Abstract 2593F.

♨ Probes designed in accordance to "Designing synthetic MLPA probes", v12. A white paper by MRC Holland

Comparing Technologies



Data Summary

- Genes tested: *BRCA1*, *PMP22*, *MSH2*, *FANCA*, *NPHP1*, *BRCA2*, *ATM*, *MLH1*, *CFTR*, *TP53*, *DMD*, *SMAD4*, *RAD51C*, *PLP1*, *SMN1*, *PALB2*, *NBN*, *MEN1*, *ASPA*, *APC*
- 79 of 80 events confirmed. The exception is *SMN1* exon 8. A probe re-design is necessary.

Gene	# samples	Events	Confirmed?
BRCA1	**11**	**e6 dup; e20 del; e1-3 del; e10-14 del; e12 dup; e12-13 dup; e13-15 dup; e13-19 del; e19 del**	**11/11**
PMP22	**9**	**e1-5 del; e1-5 dup**	**9/9**
MSH2	**8**	**e1 del; e1-2 del; e1-6 del; e5-6 del; e7 del; e8 del; e9-10 del; e16 del**	**8/8**
FANCA	**7**	**e1 del; e1-3 dup; e3-8 dup; e17-21 del; e22-28 del; e22-29 del; e22-31 del**	**7/7**

Genes with the most events

Conclusions

- MLPAseq accurately detects CNVs in a variety of different genes and ploidies
- The limitations of traditional CNV detection technologies are overcome by MLPAseq. MLPAseq is:
  - infinitely scalable — add more probes & sequence deeper
  - easily automated — both lab process & data analysis
  - cheap — oligo synthesis & sequencing are biggest costs
- MLPAseq is ideal for primary CNV detection or secondary CNV confirmation of large gene panels in the clinical genetic diagnostics laboratory