Supporting Information

Shen et al. 10.1073/pnas.1018981108

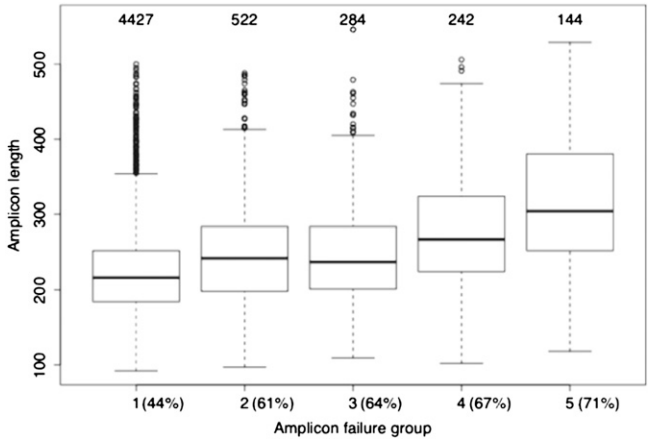


Fig. S1. Capture performance related to amplicon length. We defined five groups of targets (*x* axis) that were captured successfully in all samples (group 1: 4,427 exons) or failed on the basis of our stringent threshold of $R < 9$ in 1–10 samples (group 2: 522 exons), 11–40 samples (group 3: 284 exons), 41–62 samples (group 4: 242 exons), and failed in all samples (group 5: 144 exons). Except for groups 2 vs. 3, we found statistical significant differences of the amplicon length in the five different groups (P value < 0.0001 for 1 vs. 2, 3 vs. 4, and 4 vs. 5) indicating a lower capture performance for longer targets. However, these groups also had a different GC content that correlated with the number of failed sample amplicons. On the basis of our analysis of GC content in comparison with longer (>274 bp; >80th percentile) and shorter (<180 bp; <20th percentile) amplicons (Fig. 2*B*), we concluded that amplification failures are largely caused by higher GC content and to a lesser extend by length.

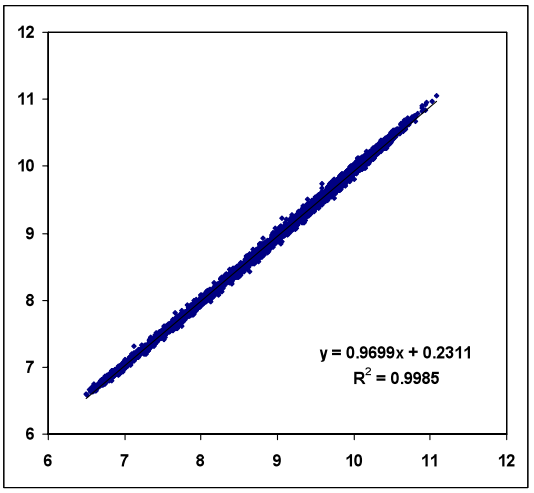


Fig. S2. Concordance of a replicate capture of one sample to two different arrays. Shown is a correlation of log intensities of all array probes (measure *T*) of two different arrays (*x* and *y* axes, respectively) each of which hybridized with an equal aliquot from one sample preparation. The high concordance (R^2 = 0.999) in this technical replicate shows the reproducibility of the resequencing arrays.

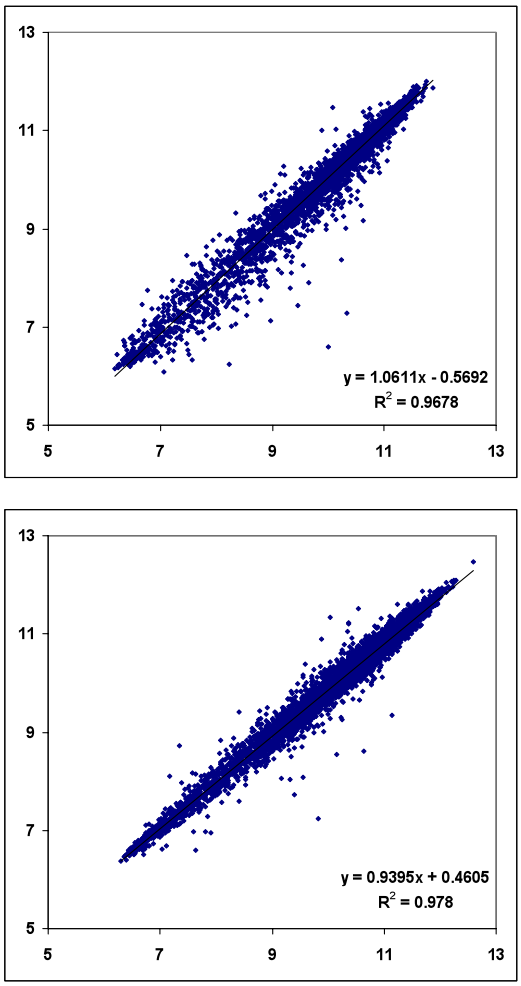
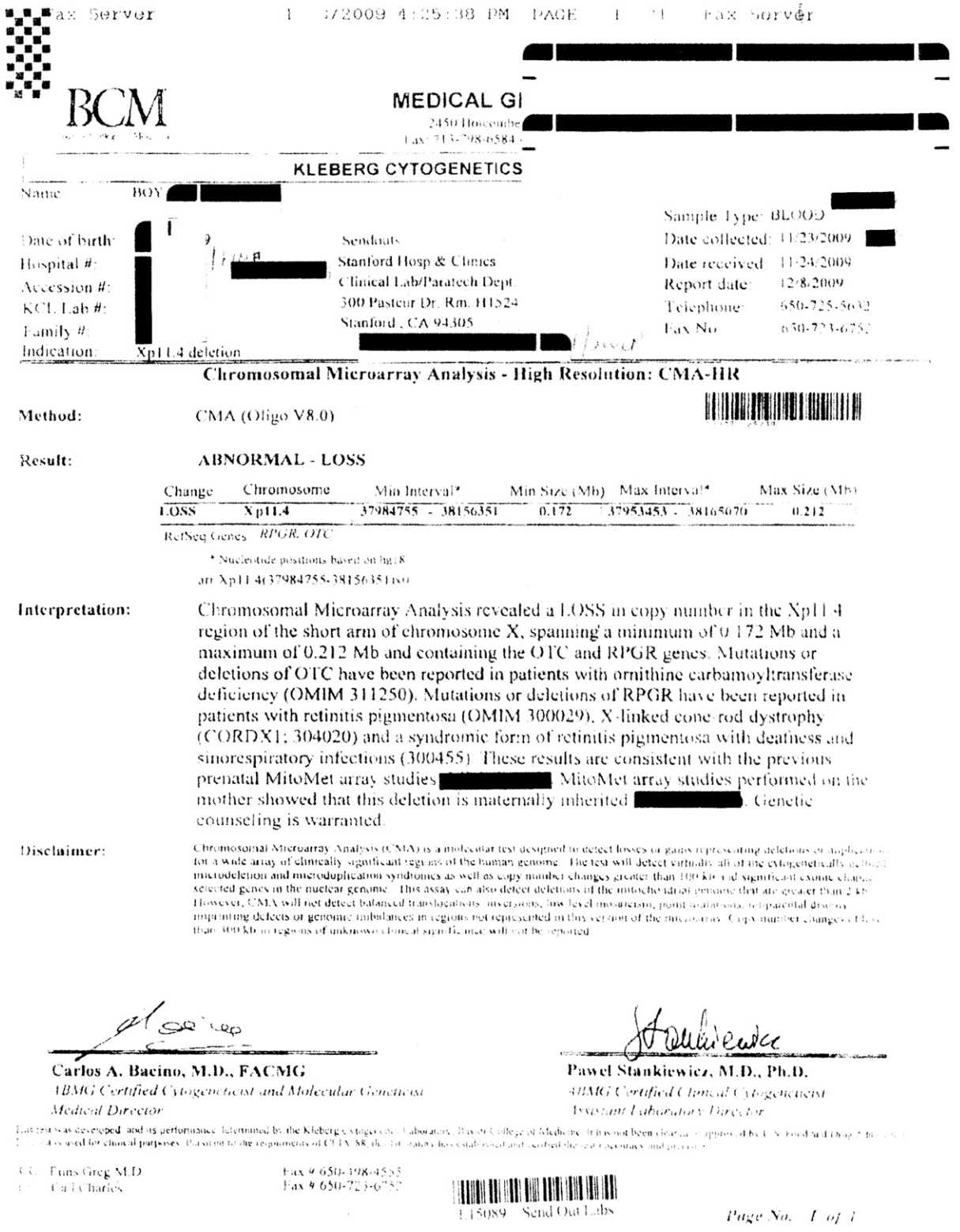


Fig. S3. Concordance of capture for genomic and WGA DNA sample preparation. Shown is a correlation of log intensities of all array probes (measure *T*) of two resequencing arrays hybridized with two different genomic DNA samples (*Upper*) and two arrays hybridized with two different WGA DNA samples (*Lower*). This analysis demonstrates the comparable performance of genomic and WGA DNAs in different sample preparations, which is critical in comparative sequence analysis.

Supporting Information for confirmation of the OTC deletion in the male child using aCGH.



Fax Server 1 /2009 4:25:38 PM PAGE 1 Fax Server

BCM

MEDICAL GI
2450 Holcombe
Fax 713-798-6584

KLEBERG CYTOGENETICS

Name: BOY [redacted]

Date of birth: [redacted]
Hospital #: [redacted]
Accession #: [redacted]
KCL Lab #: [redacted]
Family #: [redacted]
Indication: Xp11.4 deletion

Semdaats
Stanford Hosp & Clinics
Clinical Lab/Paratech Dept.
300 Pasteur Dr. Rm. H1524
Stanford, CA 94305

Sample Type: BLOOD
Date collected: 11/23/2009
Date received: 11-24-2009
Report date: 12/8/2009
Telephone: 650-725-5632
Fax No: 650-723-6782

Chromosomal Microarray Analysis - High Resolution: CMA-HR

Method: CMA (Oligo V8.0)

Result: **ABNORMAL - LOSS**

Change	Chromosome	Min Interval*	Min Size (Mb)	Max Interval*	Max Size (Mb)
LOSS	Xp11.4	37984755 - 38156351	0.172	37953453 - 38165070	0.212

RefSeq Genes *RPGR, OTC*

* Nucleotide positions based on hg18

arr Xp11.4(37984755-38156351)x0

Interpretation: Chromosomal Microarray Analysis revealed a LOSS in copy number in the Xp11.4 region of the short arm of chromosome X, spanning a minimum of 0.172 Mb and a maximum of 0.212 Mb and containing the OTC and RPGR genes. Mutations or deletions of OTC have been reported in patients with ornithine carbamoyltransferase deficiency (OMIM 311250). Mutations or deletions of RPGR have been reported in patients with retinitis pigmentosa (OMIM 300029), X-linked cone-rod dystrophy (CORDX1; 304020) and a syndromic form of retinitis pigmentosa with deafness and sinorespiratory infections (300455). These results are consistent with the previous prenatal MitoMet array studies [redacted]. MitoMet array studies performed on the mother showed that this deletion is maternally inherited [redacted]. Genetic counseling is warranted.

Disclaimer: Chromosomal Microarray Analysis (CMA) is a molecular test designed to detect losses or gains representing deletions or duplications for a wide array of clinically significant regions of the human genome. The test will detect virtually all of the cytogenetically detected microdeletion and microduplication syndromes as well as copy number changes greater than 100 Kb and significant exonic changes in selected genes in the nuclear genome. This assay can also detect deletions of the mitochondrial genome that are greater than 2 Kb. However, CMA will not detect balanced translocations, inversions, low level mosaicism, point mutations, uniparental disomy, imprinting defects or genomic imbalances in regions not represented in this version of the microarray. Copy number changes of less than 100 Kb in regions of unknown clinical significance will not be reported.

Carlos A. Bacino, M.D., FACMG
ABMG Certified Cytogeneticist and Molecular Geneticist
Medical Director

Pawel Stankiewicz, M.D., Ph.D.
ABMG Certified Clinical Cytogeneticist
Assistant Laboratory Director

This test was developed and its performance determined by the Kleberg Cytogenetics Laboratory, Baylor College of Medicine. It has not been cleared or approved by U.S. Food and Drug Administration. This test is used for clinical purposes. Pursuant to the requirements of CLIA '88, this laboratory has established and verified the test accuracy and precision.

cc: Funs Greg M.D. Fax # 650-498-4555
Ga Charles Fax # 650-723-6752

L15089 Send Out Labs

Page No. 1 of 1

Fig. S4. Confirmation of the OTC gene deletion using aCGH. In this report for patient P44 (Dataset S3), Chromosomal Microarray Analysis revealed a loss in copy number in the Xp11.4 region of the short arm of chromosome X, spanning a minimum of 0.172 Mb and a maximum of 0.212 Mb and containing the OTC and RPGR genes. Microarray analysis performed on the mother showed that this deletion is maternally inherited, which is consistent with our findings using target capture and array-based resequencing.

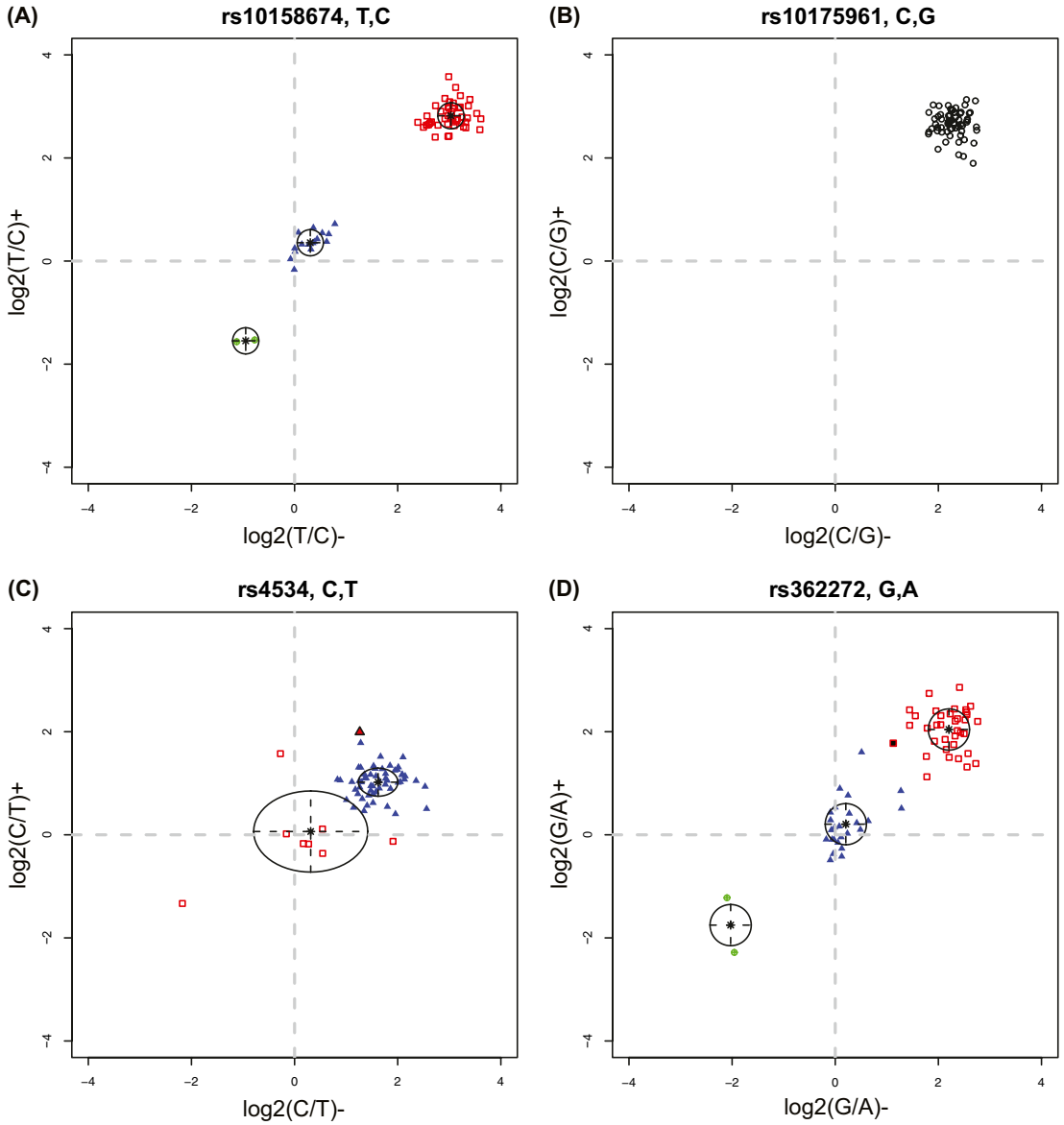


Fig. S5. Base calling of selected positions in resequencing array data analysis. Each plot shows the δ's: $\theta_{RM} - \theta_{AM}$, for sense (*y* axis) and antisense (*x* axis) strand, for all 63 samples with each data point representing one sample. (*A*) True positive position rs10158674. The clustering analysis correctly identified three genotypes for all samples. (*B*) True negative position rs10175961. Clustering identified only one cluster, which indicates all samples are nonvariant references. (*C*) Position rs4534 in sample NA18507 (red triangle) was detected as a homozygous reference C, whereas dbSNP listed this sample position as homozygous variant T. (*D*) False negative position rs362272. Clustering called this position in NA11840 (red square filled black) as reference, but this base call was removed because of a low *q* score. This sample position is a true heterozygous GA.

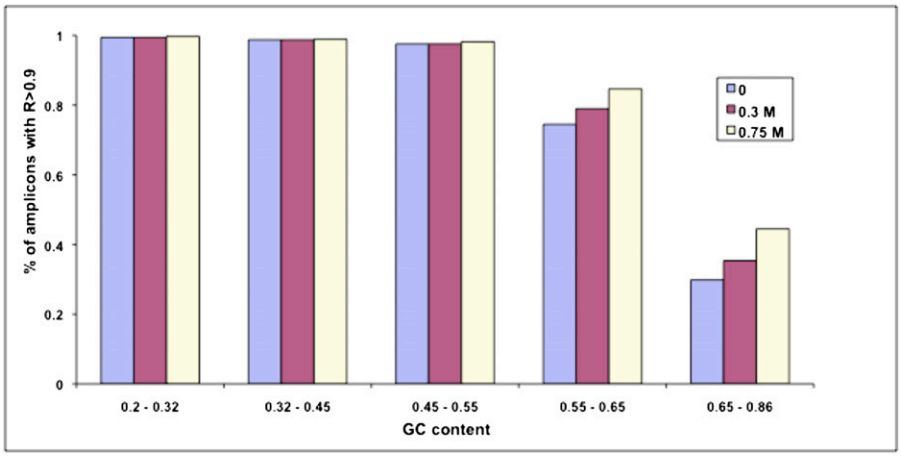


Fig. S6. Improved capture performance using betaine supplementation. Shown is the DNA sequence capture performance for all 5,619 targeted exons based on our stringent threshold of $R \geq 0.9$. The amplification success was improved by ~1.5-fold (~50% increase of amplicons that passed $R \geq 0.9$) after supplementing betaine at higher concentration (0.75 M) to the capture pools and in particular for exons with a GC content of 55% and higher.

Other Supporting Information Files

Dataset S1 (XLS)
Dataset S2 (XLS)
Dataset S3 (XLS)