

Short Technical Reports

Comparison of multiplex ligation-dependent probe amplification and real-time PCR accuracy for gene copy number quantification using the β -defensin locus

Andrea Perne^{1*}, XiangHong Zhang^{1*}, Lutz E. Lehmann², Marco Groth³, Frank Stuber², and Malte Book²

¹Department of Anaesthesiology and Intensive Care Medicine, University Hospital Bonn, Bonn, Germany, ²University Department of Anaesthesiology and Pain Therapy, Inselspital, Bern University Department, Bern, Switzerland, and ³Genome Analysis, Leibniz Institute for Age Research—Fritz Lipmann Institute, Jena, Germany

BioTechniques 47:1023-1028 (December 2009) doi 10.2144/000113300

Keywords: β -defensin; copy number variation; real-time PCR; multiplex ligation-dependent probe amplification

Supplementary material for this article is available at www.BioTechniques.com/article/113300.

*A.P. and X.-H.Z. contributed equally to this work.

The reliable quantification of gene copy number variations is a precondition for future investigations regarding their functional relevance. To date, there is no generally accepted gold standard method for copy number quantification, and methods in current use have given inconsistent results in selected cohorts. In this study, we compare two methods for copy number quantification. β -defensin gene copy numbers were determined in parallel in 80 genomic DNA samples by real-time PCR and multiplex ligation-dependent probe amplification (MLPA). The pyrosequencing-based paralog ratio test (PPRT) was used as a standard of comparison in 79 out of 80 samples. Real-time PCR and MPLA results confirmed concordant *DEFB4*, *DEFB103A*, and *DEFB104A* copy numbers within samples. These two methods showed identical results in 32 out of 80 samples; 29 of these 32 samples comprised four or fewer copies. The coefficient of variation of MLPA is lower compared with PCR. In addition, the consistency between MLPA and PPRT is higher than either PCR/MLPA or PCR/PPRT consistency. In summary, these results suggest that MLPA is superior to real-time PCR in β -defensin copy number quantification.

Introduction

Gene copy number variations (CNVs) contribute significantly to interindividual genetic variability (1). One important site of genomic rearrangement is the defensin gene locus located at chromosome 8p23.1

(2). Defensins are antimicrobial and immunomodulating polypeptides that play an important role in the innate immune system of mammals, insects, and plants (3–8). Genes for human β -defensin 2 (hBD-2), β -defensin 3 (hBD-3), and β -defensin 4 (hBD-4) were named *DEFB4*,

DEFB103A, and *DEFB104A*, respectively. Recent studies have shown functional implications of variation in *DEFB4*, *DEFB103A*, and *DEFB104A* gene copy numbers (2,9). Since defensins play an important role in the innate immune response, quantitative variations in the gene copy number might contribute to susceptibility to infectious and inflammatory diseases (10).

In principle, CNVs are detectable using screening techniques such as DNA arrays or sequencing. Redon et al. compared the single-nucleotide polymorphism genotyping array method and the comparative genomic hybridization array method for CNV detection (1), finding that both high-throughput methods detected deletions and duplications qualitatively, but failed to generate definitive gene copy numbers. Recently, parallel sequencing emerged as a method for quantitative identification of genomic structural variance, including deletions, duplications, and rearrangements (11,12). However, this method is limited by its use of random amplification primers and lack of definitive copy number quantification. Although genome-wide CNV mapping has been simplified by these different direct screening methods, none of these approaches currently offers reliable single gene quantification.

Consequently, defensin CNVs are analyzed with methods that offer single gene quantification, such as MLPA, PPRT, fluorescence in situ hybridization (FISH), or real-time PCR. To date, a gold standard method for defensin gene quantification has not been established (2,9,13–15). Hollox et al. reported concordant gene copy numbers of *DEFB4*, *DEFB103A*, and *DEFB104A* by multiplex amplifiable probe hybridization and semiquantitative fluorescence in situ hybridization (2). Our own multiplex ligation-dependent probe amplification (MLPA) results confirmed this concordance (14), whereas PCR analysis detected discordance in intra-individual β -defensin copy numbers (13). Other groups have reported inter-method quantification comparisons that showed inconsistent results for absolute copy number (16,17). Therefore, a thorough reevaluation of different quantification methods is required. The aim of this study was to compare MLPA and real-time PCR for β -defensin gene quantification through the analysis of 80 genomic DNA samples.

Materials and methods

Genomic DNA

DNA samples from cell lines and whole blood samples were included. 42 genomic DNA samples, derived from immor-

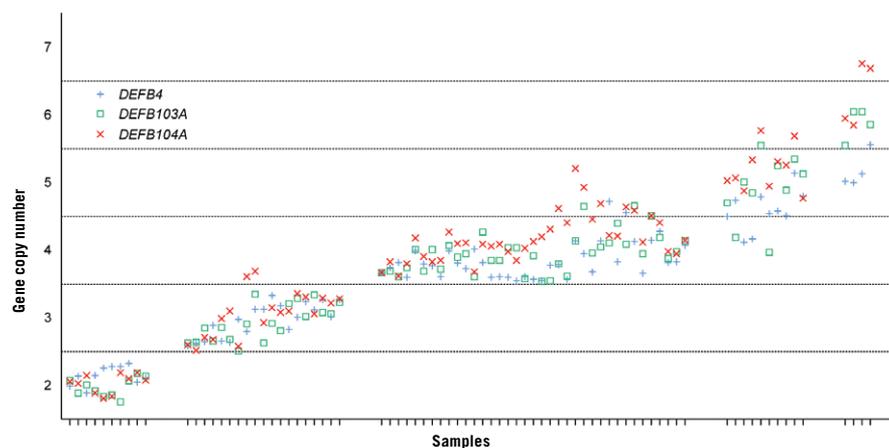


Figure 1. Real-time PCR β -defensin copy number determination. Indicated are the raw, unrounded copy number values quantified by real-time PCR of DEF B4 (+), DEF B103A (\square) and DEF B104A (x). The samples were grouped by the means of the intra-sample unrounded copy numbers of DEF B4, DEF B103A, and DEF B104A (means are provided in Supplementary Table S1). The distribution of gene copy numbers of all three genes passed the Kolmogorov-Smirnov normality test ($P > 0.05$).

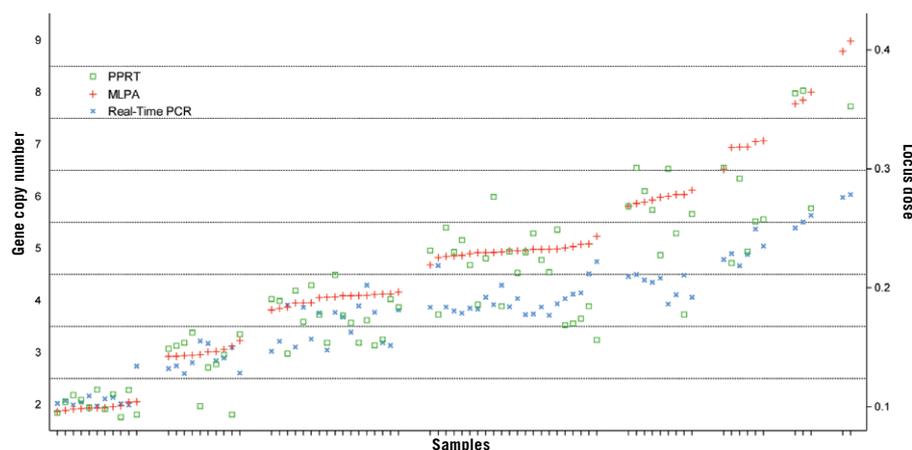


Figure 2. Inter-method differences in β -defensin locus quantification. DEF B4, DEF B103A, and DEF B104A copy numbers were determined by real-time PCR, MLPA, and PPRT. Raw data of PCR and PPRT were processed to unrounded copy number values (left y-axis). Raw data of the MLPA method are presented as unrounded locus dose (right y-axis) as described previously (14).

talized B-lymphoblastoid cells lines, were purchased from the Coriell Institute for Medical Research (Camden, NJ, USA) and the European Collection of Cell Cultures (ECACC; Wiltshire, UK). Samples were named NAxxxxx (Coriell) and Cxxxx (ECACC) and have been designated with a superscripted 'HM' if they were used in the

HapMap Project (18). In addition, 38 DNA samples (code RC) from a white population extracted from whole blood were also analyzed. Written informed consent from healthy volunteers was obtained, as requested by the local ethics committee. This procedure complied with all relevant laws, guidelines, and policies. For simplicity, all specimens will

Table 1. Inter-method agreement as assessed by kappa coefficients

Comparison	Kappa coefficient	95% CI
PCR vs. MLPA	0.47	(0.38–0.56)
PCR vs. PPRT	0.49	(0.39–0.60)
MLPA vs. PPRT	0.77	(0.69–0.85)

be referred to hereafter as “DNA samples.” The concentration and purity of DNA was determined using a NanoDrop device (NanoDrop Technologies, Wilmington, DE, USA). MLPA and PPRT data from the Coriell and ECACC samples were part of a previous report (14).

Real-time PCR assay

The primers and hybridization probes were synthesized according to our previously published approach (13). The albumin gene (*ALB*) served as a single copy per chromosome reference gene in this study. The method of real-time quantitative PCR using LightCycler Relative Quantification Software 1.0 (Roche, Mannheim, Germany) for analysis requires an appropriate calibrator. Genomic DNA sample NA18608^{HM}, with one copy of target gene per haploid genome as determined by different methods (14,19) and one copy of reference gene per haploid genome, was used as calibrator. The PCR reaction was performed according to the protocol established by our group (13). In brief, this method of relative real-time quantitative PCR requires a coefficient file from which the efficiency of PCR is calculated by the software. To create a coefficient file for the *ALB* gene and the *DEF B4*, *DEF B103A*, or *DEF B104A* gene, an unknown genomic DNA sample was prepared in a 10-step dilution series. Then, for each dilution, the β -defensin and the *ALB* genes were amplified in a single capillary to acquire the relative quantification standard curves for both of these genes. Using the LightCycler software, a coefficient file can be prepared from the two relative quantification standard curves. Data analysis of the real-time PCR method has been described in detail by our group in a previous publication (13).

MLPA and analysis

MLPA was carried out using the MLPA Kit P139 (MRC Holland, The Netherlands) as described in our previous publication (14). This MLPA assay uses a total of 43 probes. Ten of these probes hybridize to single-copy genes not located on chromosome eight. The remaining 33 probes are mapped exclusively to chromosome eight, which carries the α - and β -defensin gene cluster. Of these 33 probes, 10 cover the β -defensin cluster, 10 hybridize to genes within the α -defensin cluster, and 13 are specific for genes flanking



CALIBRATION SERVICES LLC
Committed to Quality Pipette Service

At MEU Calibration Services quality and precision is the cornerstone of our work and customer satisfaction is our commitment. We repair and calibrate all brands of pipettes including, but not limited to Rainin, Eppendorf, Gilson and Brand.

Visit us at meucalibrationservices.com or call us at 609-556-8660

the α - or β -defensin cluster. Concordance analysis was performed using a correlation matrix, and absolute β -defensin copy numbers for each individual were determined using the relative locus dose calculation (14).

Pyrosequencing-based paralog ratio test

Of the 80 samples, 79 were successfully analyzed by PPRT (one sample was excluded due to failed analysis). The PPRT method is based on the paralog ratio test (PRT) technique described by Armour et al. (17), which targets paralogous gene loci of *HSPDP3*. The detailed procedures have been described by our group (14). In brief, a first PCR was carried out using unlabeled forward and reverse primers as reported previously (17). Subsequently, a second PCR amplification was performed using the product from the first PCR as a template to obtain biotin-labeled PCR product. This second PCR was set up using the previous forward primer (but 5'-biotin-labeled) and a second reverse primer. The resulting PCR product was used for the pyrosequencing reaction, with the second PCR reverse primer used as the sequencing primer. Two paralog sequence variations (PSV) were used to quantify chromosome 8 and chromosome 5 paralogs in only one pyrosequencing reaction. Pyrosequencing results were rejected if the deviation between the two PSVs were >5%. The copy number (cn) of the chromosome 8 paralog was calculated subsequently using the following function: $cn = 2 \times AV8 / (100 - AV8)$, where AV8 describes the average percentage of the chromosome 8 PSV alleles. The calculated values were rounded to the nearest integer.

Statistical methods

The analysis of the intra-method variances was performed by one-way analysis of variance with Bonferroni's multiple comparison test. A p -value <0.05 was regarded as statistically significant. Bland-Altman plots were used to visually assess agreement between the three methods (20). The linearly weighted kappa statistic was calculated to estimate agreement between the methods. Confidence intervals were calculated using bias-corrected bootstrap estimations.

Results and discussion

Identification of five gene copy number clusters by real-time PCR

β -defensin gene (*DEFB4*, *DEFB103A*, and *DEFB104A*) quantification was performed in 80 genomic DNA samples by real-time PCR; Figure 1 shows the raw, unrounded copy numbers obtained by this method. The samples were grouped by the mean of the intra-sample normalized ratios for *DEFB4*, *DEFB103A*, and *DEFB104A* (ratios are provided in Supplementary Table S1). This clustering enabled the final copy number identification and resulted in five clusters with two to six gene copies, respectively. The coefficient of variation (CV) in the two-copy cluster was 7.5% and increased to a maximum of 10% in the six-copy cluster. In 27 out of 29 samples with two or three *DEFB4* gene copies, the *DEFB103A* and *DEFB104A* gene copy numbers were concordant with *DEFB4*. In contrast, in the samples with four or more *DEFB4* copies, *DEFB4*, *DEFB103A*, and *DEFB104A* gene copy number concordance was detected in only 33 out of 51 samples. Overall, 60 out of 80 samples showed concordant results. In the 20 discordant samples, 19 showed an absolute difference of one copy number, whereas one sample (NA18502^{HM}) showed a two-copy number difference (Supplementary Table S1).

Falsely detected discordance at high copy numbers is a limitation of real-time PCR copy number quantification. Similar limitations in the high-copy number range have been previously described for MLPA by our group (14), and in PRT by Armour and et al. (17).

Comparison of copy number quantification methods

The inter-method comparison was performed by using the unrounded β -defensin gene copy numbers detected by PCR and the locus dose-based copy number determination by MLPA. In the absence of a gold standard method, unrounded copy numbers determined by PPRT were used to evaluate the real-time PCR and MLPA results.

Figure 2 shows the β -defensin gene copy number determinations (the numerical values are available in Supplementary Table S2). Real-time PCR and MLPA quantification gave consistent results for copy number for 32 out of 80 samples. Also, by real-time PCR quantification, 30 of these 32 consistent samples showed fully concordant copy numbers for *DEFB4*, *DEFB103A*, and *DEFB104A*. In 29 of the 32 samples with consistent results between the two methods, there were four or fewer gene copies. This emphasizes that consistent MLPA and PCR results tend to occur in the lower-copy number range and in samples with high intra-sample concordance among the three genes in real-time PCR results.

In the 32 samples with consistent real-time PCR/MLPA results, PPRT confirmed the quantifications in 24. Of note, 51 out of 79 samples showed consistent MLPA and PPRT copy number quantification results, but consistent real-time PCR and PPRT quantification results were detectable in only 35 out of 79 samples.

The CV of the raw data was calculated to evaluate the performance of the three methods. The CVs of real-time PCR, MLPA, and PPRT were calculated and grouped by copy numbers (Supplementary Figure S1). One-way analysis of variance with Bonferroni multiple comparison test showed significantly greater CV in the real-time PCR ($P < 0.05$) and PPRT ($P < 0.05$) methods compared with MLPA (Supplementary


Setting standards
in analytical science



AGOWA genomics

Next generation sequencing

applying the

Roche GS FLX TITANIUM Technology

Service features:

- *De novo* sequencing of prokaryotic and small eukaryotic genomes
- Finishing of *de novo* sequencing
- Analysis of metagenomes, transcriptomes / normalised cDNA, methylation patterns, pools of tagged fosmids and BACs
- Targeted re-sequencing
- ChIP and ncRNA sequencing.

YEAR END OFFER 2009
Please check our webpage
www.lgc.co.uk/genomics
for more information.

AGOWA GmbH (part of LGC) • Berlin • Germany
Email: ngs@agowa.de • Web: www.lgc.co.uk/genomics

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or any retrieval system, without the written permission of the copyright holder. © LGC Limited, 2009. All rights reserved.
Picture: Museum Wiesbaden and Kevin Clarke. The Invisible Body, Wiesbaden 1999 (Portrait of Miguel Alemán, page 28)

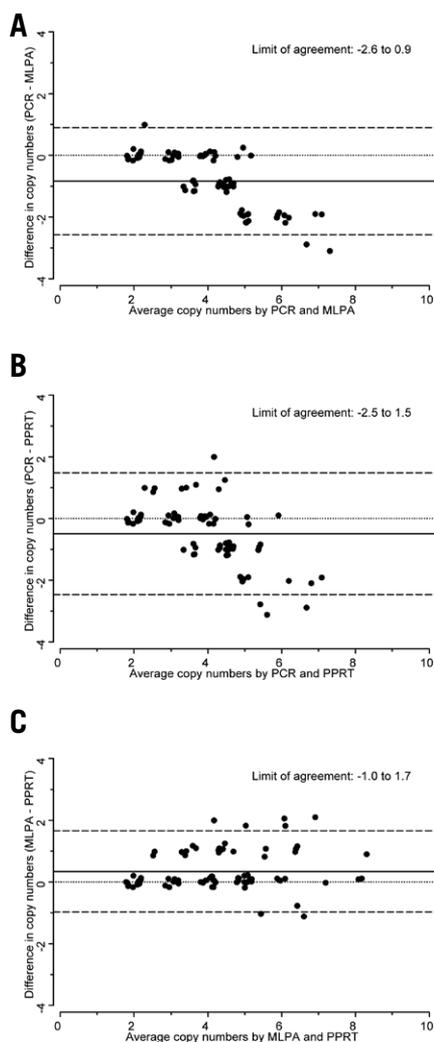


Figure 3. Bland-Altman plots of inter-method comparisons. The plots show the difference between the methods against their mean. The points were offset slightly to allow individual measurements to be better distinguished. The solid line represents the mean of all samples and the dashed line is two times the standard deviation (limit of agreement). (A) Comparison of real-time PCR vs. MLPA. (B) Comparison of real-time PCR vs. PPRT. (C) Comparison of MLPA vs. PPRT.

Figure S1). In addition, the inter-assay CV was calculated for four samples in five replicates. The highest CVs were 9.61% for PPRT, 6.25% for real-time PCR, and 2.88% for MLPA. Together, these data indicate differences in the performance of the different methods.

Template amplification is a crucial step in MLPA and PCR. In the PCR method, different primers are necessary for different genes, including the reference gene. Although PCR copy number quantification is corrected for amplification efficiencies that diverge from the expected value of 2, poor amplification might occur exclusively in the target or reference gene. Such an effect might be attributable to a suboptimal template to

enzyme ratio in the range of high copy numbers. A modified amount of template in the high-copy number range might shift the start of the log-linear target gene amplification phase to higher cycle numbers.

In MLPA, hybridization probes hybridize specifically to the target sequence and ligate with the corresponding probe prior to the amplification procedure. The discriminatory power (concordant versus discordant) of the probes used in the present investigation was confirmed by probe to probe correlation coefficients, which were previously published by our group (14). Moreover, incomplete hybridization would not be expected to significantly influence the copy number quantification because a reference gene signal based on the averaged signal intensity of five nearest-neighbor probes is used, and the target gene signal represents the average of 10 gene-specific probes within the β -defensin locus. MLPA amplification of all target and reference genes was performed in parallel by a single pair of primers. Consequently, compromised reaction efficiency would bias the amplification efficiency of all target sequences equally. This fact might contribute to the lower coefficient of variation of MLPA as well as to its superior discriminatory power in the high-copy number range.

Evaluation of inter-method consistency

For visualization and analysis of inter-method agreement, we chose Bland-Altman plots comprising the 79 MLPA, PPRT, and PCR quantified samples (Figure 3). These plots display the intra-sample differences against the intra-sample mean, and are based on the assumption that the mean of the two measurements is the best available estimation (20). The limit of agreement was calculated as two times the standard deviation of the differences. The smallest limit of agreement was found in the MLPA/PPRT comparison (range -1.0–1.7; Figure 3C), whereas the comparison of real-time PCR with MLPA (Figure 3A) and PPRT (Figure 3B) resulted in higher ranges (-2.6–0.9 and -2.5–1.5, respectively). Thus, the Bland-Altman plots indicate the limited consistency of real-time PCR with both of the other methods. However, both plots show higher consistency in the low-copy-number range compared with the high-copy number range. This observation is consistent with the decreased concordance seen within the high-copy number PCR results (Figure 1).

Although Bland et al. suggested that methods yielding values within the limits of agreement can be considered clinically interchangeable (20), this conclusion is not applicable to gene copy number quantification, where a reliable quantification is mandatory. However, in the absence of a gold standard method, the assessment of agreement by

these plots is an important contribution to the evaluation of different methods.

Additionally, the estimated agreement between the methods was calculated by linearly weighted kappa coefficients. The kappa coefficients for inter-method agreement are corrected for random agreement and are displayed in Table 1. According to Landis and Koch (21), the range 0.41–0.60 can be interpreted as moderate agreement, 0.61–0.80 as substantial agreement, and 0.81–1.00 as almost perfect agreement. These data confirmed the results seen in the Bland-Altman plots, in that there is limited consistency between real-time PCR and both alternative methods.

In the end, there is increasing evidence for the functional relevance of copy number variations and, moreover, positive associations with diseases have been reported (2,10,22). Reliable quantification is the basic precondition for further research in the rapidly growing field of genetic variability and predisposition in complex diseases. In this study, we found that two quantification methods—real-time PCR and MLPA—are affected similarly by decreased reliability in the discrimination of high copy numbers. This makes intuitive sense because, presuming optimal conditions, the difference between 2 and 4 gene copies causes an increase of 100% signal amplitude, while the change from 8 to 10 copies results in only a 25% increase in amplitude. Despite this limitation, the presented results indicate the advantage of MLPA compared with real-time PCR in gene copy number quantification.

Acknowledgments

The authors thank Sven Trelle from the Clinical Trial Unit of the Inselspital, Bern University Hospital for the excellent statistical advice. This work was supported by a grant from the “Deutsche Forschungsgemeinschaft” (Bo 1929/1-2).

Competing interests

The authors declare no competing interests.

References

1. Redon, R., S. Ishikawa, K.R. Fitch, L. Feuk, G.H. Perry, T.D. Andrews, H. Fiegler, M.H. Shapero, et al. 2006. Global variation in copy number in the human genome. *Nature* 444:444–454.
2. Hollox, E.J., J.A. Armour, and J.C. Barber. 2003. Extensive normal copy number variation of a beta-defensin antimicrobial-gene cluster. *Am. J. Hum. Genet.* 73:591–600.
3. Lehrer, R.I. and T. Ganz. 2002. Defensins of vertebrate animals. *Curr. Opin. Immunol.* 14:96–102.

4. Ganz, T. 2003. Defensins: antimicrobial peptides of innate immunity. *Nat. Rev. Immunol.* 3:710-720.
5. Ashitani, J., H. Mukae, T. Hiratsuka, M. Nakazato, K. Kumamoto, and S. Matsukura. 2001. Plasma and BAL fluid concentrations of antimicrobial peptides in patients with *Mycobacterium avium*-intracellulare infection. *Chest* 119:1131-1137.
6. Wehkamp, J., J. Harder, M. Weichenthal, O. Mueller, K.R. Herrlinger, K. Fellermann, J.M. Schroeder, and E.F. Stange. 2003. Inducible and constitutive beta-defensins are differentially expressed in Crohn's disease and ulcerative colitis. *Inflamm. Bowel Dis.* 9:215-223.
7. Yang, D., O. Chertov, S.N. Bykovskaia, Q. Chen, M.J. Buffo, J. Shogan, M. Anderson, J.M. Schroder, et al. 1999. Beta-defensins: linking innate and adaptive immunity through dendritic and T cell CCR6. *Science* 286:525-528.
8. Yang, D., A. Biragyn, L.W. Kwak, and J.J. Oppenheim. 2002. Mammalian defensins in immunity: more than just microbicidal. *Trends Immunol.* 23:291-296.
9. Linzmeier, R.M. and T. Ganz. 2005. Human defensin gene copy number polymorphisms: comprehensive analysis of independent variation in alpha- and beta-defensin regions at 8p22-p23. *Genomics* 86:423-430.
10. Fellermann, K., D.E. Stange, E. Schaeffeler, H. Schmalzl, J. Wehkamp, C.L. Bevens, W. Reinisch, A. Teml, et al. 2006. A chromosome 8 gene-cluster polymorphism with low human beta-defensin 2 gene copy number predisposes to Crohn disease of the colon. *Am. J. Hum. Genet.* 79:439-448.
11. Campbell, P.J., P.J. Stephens, E.D. Pleasance, S. O'Meara, H. Li, T. Santarius, L.A. Stebbings, C. Leroy, et al. 2008. Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nat. Genet.* 40:722-729.
12. Chiang, D.Y., G. Getz, D.B. Jaffe, M.J. O'Kelly, X. Zhao, S.L. Carter, C. Russ, C. Nusbaum, et al. 2009. High-resolution mapping of copy-number alterations with massively parallel sequencing. *Nat. Methods* 6:99-103.
13. Chen, Q., M. Book, X. Fang, A. Hoefft, and F. Stuber. 2006. Screening of copy number polymorphisms in human beta-defensin genes using modified real-time quantitative PCR. *J. Immunol. Methods* 308:231-240.
14. Groth, M., K. Szafranski, S. Taudien, K. Huse, O. Mueller, P. Rosenstiel, A.O. Nygren, S. Schreiber, et al. 2008. High-resolution mapping of the 8p23.1 beta-defensin cluster reveals strictly concordant copy number variation of all genes. *Hum. Mutat.* 29:1247-1254.
15. Nuytten, H., I. Wlodarska, K. Nackaerts, S. Vermeire, J. Vermeesch, J.J. Cassiman, and H. Cuppens. 2009. Accurate determination of copy number variations (CNVs): application to the alpha- and beta-defensin CNVs. *J. Immunol. Methods* 344:35-44.
16. Hollox, E.J., U. Huffmeier, P.L. Zeeuwen, R. Palla, J. Lascorz, D. Rodijk-Olthuis, P.C. van de Kerkhof, H. Traupe, et al. 2008. Psoriasis is associated with increased beta-defensin genomic copy number. *Nat. Genet.* 40:23-25.
17. Armour, J.A., R. Palla, P.L. Zeeuwen, M. den Heijer, J. Schalkwijk, and E.J. Hollox. 2007. Accurate, high-throughput typing of copy number variation using paralogue ratios from dispersed repeats. *Nucleic Acids Res.* 35:e19.
18. International HapMap Consortium. 2005. A haplotype map of the human genome. *Nature* 437:1299-1320.
19. Huse, K., S. Taudien, M. Groth, P. Rosenstiel, K. Szafranski, M. Hiller, J. Hampe, K. Junker, et al. 2008. Genetic variants of the copy number polymorphic beta-defensin locus are associated with sporadic prostate cancer. *Tumour Biol.* 29:83-92.
20. Bland, J.M. and D.G. Altman. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1:307-310.
21. Landis, J.R. and G.G. Koch. 1977. The measurement of observer agreement for categorical data. *Biometrics* 33:159-174.
22. Ionita-Laza, I., A.J. Rogers, C. Lange, B.A. Raby, and C. Lee. 2008. Genetic association analysis of copy-number variation (CNV) in human disease pathogenesis. *Genomics*.

Received 6 July 2009; accepted 9 October 2009.

Address correspondence to Malte Book, University Department of Anaesthesiology and Pain Therapy, Inselspital, CH-3010 Bern, Switzerland. email: malte.book@dkf.unibe.ch

LOTS OF SAMPLES TO HOMOGENIZE?

We've got the solutions you've been looking for

Multi-Sample Processing • No Cross-Contamination



Find the best homogenizer for your specific application at

www.omni-inc.com/apps

Call now to find out about our special year-end discounts!



1-800-776-4431
www.omni-inc.com