

# A Novel Next-Generation Sequencing Approach without Donor-Derived Material for Acute Rejection and Infection Monitoring in Solid Organ Transplantation

Bing Wei<sup>1,2,3\*</sup>, Lihong Zeng<sup>4\*</sup>, Di Shao<sup>4,5\*</sup>, Chunting Zheng<sup>4</sup>, Qing Yang<sup>4</sup>, Jibin Zhang<sup>4</sup>, Dong Xiao<sup>1,2,3</sup>, Qihua Deng<sup>6</sup>, Yongping Lin<sup>1,2,3</sup>, Danxia Huang<sup>1,2,3</sup>, Liping Liu<sup>6</sup>, Xin Xu<sup>1,2,3</sup>, Wenhua Liang<sup>1,2,3</sup>, Chunrong Ju<sup>1,2,3</sup>, Jian Wang<sup>7,8</sup>, Karsten Kristiansen<sup>5,7</sup>, Jianxing He<sup>1,2,3#</sup>, Mingzhi Ye<sup>4,5,7,9#</sup>

<sup>1</sup>The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

<sup>2</sup>National Clinical Research Center for Respiratory Disease, Guangzhou, China

<sup>3</sup>State Key Laboratory of Respiratory Disease, Guangzhou, China

<sup>4</sup>BGI-Guangdong, BGI-Shenzhen, Guangzhou, China

<sup>5</sup>Laboratory of Genomics and Molecular Biomedicine, Department of Biology, University of Copenhagen, Copenhagen DK, Denmark

<sup>6</sup>The Translational Medicine Laboratory, The First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China

<sup>7</sup>BGI-Shenzhen, Shenzhen, China

<sup>8</sup>James D. Watson Institute of Genome Sciences, Hangzhou, China

<sup>9</sup>BGI-Guangzhou, Guangzhou Key Laboratory of Cancer Trans-Omics Research, Guangzhou, China

Email: <sup>\*</sup>yemingzhi@genomics.cn, <sup>#</sup>hejx@vip.163.com

**How to cite this paper:** Wei, B., Zeng, L.H., Shao, D., Zheng, C.T., Yang, Q., Zhang, J.B., Xiao, D., Deng, Q.H., Lin, Y.P., Huang, D.X., Liu, L.P., Xu, X., Liang, W.H., Ju, C.R., Wang, J., Kristiansen, K., He, J.X. and Ye, M.Z. (2018) A Novel Next-Generation Sequencing Approach without Donor-Derived Material for Acute Rejection and Infection Monitoring in Solid Organ Transplantation. *Journal of Cancer Therapy*, 9, 623-638. <https://doi.org/10.4236/jct.2018.99054>

**Received:** July 23, 2018

**Accepted:** September 2, 2018

**Published:** September 5, 2018

## Abstract

**Background:** Donor-derived cell free DNA (ddcf DNA) has been reported as a universal noninvasive biomarker for rejection monitoring in heart, kidney, liver, and lung transplantation. Current approaches based on next-generation sequencing for quantification of ddcf DNA, although promising, may be restricted by the requirement for donor material, as donor samples may not be available. **Methods:** We proposed a novel next-generation sequencing approach without donor-derived material and compared the non-donor-derived approach and the donor-derived approach using simulation testing and 69 clinical specimens. We also evaluated the performance for acute rejection and infection monitoring in lung transplantation. **Results:** The non-donor-derived approach reached similar efficacy as the donor-derived approach with a significant linear correlation of  $R^2 = 0.98$ . Subsequent validation in clinical specimens demonstrated significant difference between the acute rejection

\*Bing Wei, Lihong Zeng and Di Shao contributed equally to the work, and all should be considered as first authors.

Copyright © 2018 by authors and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

group ( $4.83\% \pm 2.11\%$ , mean  $\pm$  SD) and the non-rejection group ( $1.61\% \pm 0.63\%$ , mean  $\pm$  SD) ( $P < 0.0001$ , Welch's t test). With the cut-off value of 2.999, our approach had 90.48% sensitivity (95% CI, 69.62% - 98.83%), 100% specificity (95% CI, 91.59% - 100%), and AUC 0.9266 (95% CI, 0.8277 - 1.026). The test also had the ability to simultaneously detect infectious agents, especially cytomegalovirus, as compared with the clinical test. **Conclusion:** The proposed approach without donor-derived material could potentially be used to monitor acute rejection and infection in lung transplantation and may be applied to other types of solid organ transplantation.

## Keywords

Cell-Free DNA, Genome Transplant Dynamics, Acute Rejection, Transplant, Infection

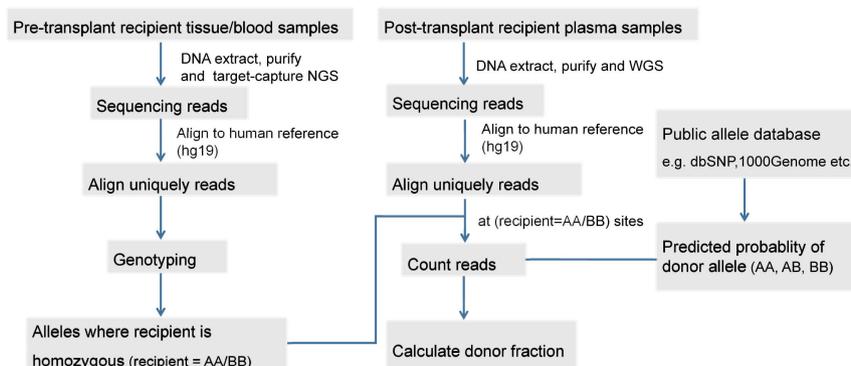
## 1. Introduction

As the respiratory centre, the lungs require strong abilities for environmental adaptation and immuno-protection against microbial infections. For patients with end-stage lung disease, lung transplantation may constitute the only effective approach and may largely increase life expectancy and substantially improve quality of life [1]. However, despite considerable advances and the wide use of immunosuppressant drugs, acute rejection (AR) remains a highly prevalent major complication of transplantations, especially in the first year post-operationally, impacting 50% to 90% of patients [2]. It is also recognized as one of the risk factors for the development of bronchiolitis obliterans syndrome, which ultimately leads to long-term morbidity and mortality after lung transplantation [3]. However, no reliable serum marker is available to monitor AR after lung transplantation [4]. Transbronchial biopsy, the gold standard for diagnosis, is an invasive procedure that may cause side effects and is limited by inter-observer variability in grading [3] [4] [5]. Apart from rejection, lung transplant recipients are also at risk of infections owing to hypoimmunity and susceptibility to immunosuppressants, poor clearance of airway secretions, impaired cough reflex, and impaired blood flow to the lung graft [6]. Differential diagnosis between rejection and infection after lung transplantation has always been difficult for clinicians, as the symptoms are generally too similar to distinguish. Therefore, there is considerable need for simple and noninvasive approaches for early and accurate lung allograft rejection and/or infectious pathogen test methods.

In 1998, Lo *et al.* found that there were cell-free donor-derived DNA (ddcfDNA) tags existing in the plasma samples of transplant recipients and that these tags might be used for monitoring graft rejection [7]. Since then, methods based on donor-specific chromosome Y, HLA marker, and single nucleotide polymorphism (SNP) sites from plasma DNA [8] [9] [10], with the aid of techniques such as digital droplet PCR coamplification at lower denaturation tempera-

ture-PCR [11] [12], quantitative PCR and next-generation sequencing (NGS), had been used for transplantation rejection monitoring of the liver, kidney, heart, and lung [13] [14] [15] [16] [17]. Current approaches for quantification of ddcfDNA that do not obtain massively parallel signatures and do not use donor-derived material such as digital droplet PCR may lead to instability and inconclusive results. Approaches based on SNPs by plasma sequencing could avoid this shortcoming and have shown great potential for application in solid organ transplantation. One of these methods is the genome transplant dynamics (GTD) approach [10] [14] [17], which used a bead-based system for genotyping from the genomic DNA of pre-transplant donors and recipients to distinguish heterologous SNPs and whole genome sequencing (WGS) from cell-free DNA (cfDNA) of the post-transplant recipient plasma to calculate the donor fraction by a weighted formula. As an essential step of the GTD approach, large-scale use of genotyping arrays would significantly increase the cost of rejection detection [18]. The use of this approach may also be restricted by the need for a donor's genomic DNA information for genotyping, as the policy of privacy protection for donors is strictly enforced, and such donor samples may be lacking in the clinic treating the recipient. Therefore, a NGS-based approach not requiring donor-derived material would greatly enhance transplantation monitoring.

Here, we introduce a non-donor-derived cfDNA transplant dynamics (NDTD) approach that is implemented by genotyping with only genomic DNA from a pre-transplant recipient by targeted capture NGS in a mini-screen SNP array and calculating donor fraction with cell-free DNA from post-transplant recipient samples that contain cfDNA such as plasma and urine by extra-low depth WGS to monitor AR and infection. The scheme of the workflow used to monitor AR by the NDTD approach is shown in **Figure 1**. In the current study, ddcfDNA was first used as a biomarker of transplantation by the NGS approach genotyping without donor-derived materials to solve the differentiation between rejection and infection. A specific cut-off value algorithm was established to calculate acute rejection and non-rejection. Then, clinical specimens were brought in for validation. However, large cohorts should be examined for further validation and study.



**Figure 1.** Scheme of the workflow used to monitor acute rejection by the NDTD approach.

## 2. Materials and Methods

### 2.1. Patients

We recruited 16 patients (see online Supplemental **Table S1**) who had the lung transplantation at The First Affiliated Hospital of Guangzhou Medical University from 2013 to 2016. Multiorgan transplant recipients were excluded. Pre-transplant whole-blood samples were collected from donors and tissues from recipients for genotyping. We collected 69 plasma recipient samples at follow-up time points post-transplant. Clinical tests including 17 trans-bronchial biopsies as conditions of most patients were not suitable for biopsy operation, which is an invasive procedure and may induce severe complications that might be even life-threatening in China, as well as laboratory examinations such as pathogen culture tests, fluorescence quantitative PCR of CMV, blood gas analysis, serum electrolyte, creatinine, and liver-enzyme or the concentration of immunosuppressive agents were also collected. This study was approved by the Institutional Review Board (IRB) of BGI and the First Affiliated Hospital of Guangzhou Medical University (No. BGI-IRB 14079). All patients gave informed consent to the protocol approved by our institutional review board.

### 2.2. DNA Extraction, Library Preparation, and Sequencing

For each time point, 5 ml peripheral blood was collected in an EDTA tube and stored at 4°C immediately after collection. Plasmas were centrifuged within 4 h following a two-step centrifugation procedure: 1) centrifuge the peripheral blood in a Horizontal centrifuge at 1600 g for 10 minutes at 4°C, then transfer the supernatant carefully to new 2 ml microcentrifuge tubes, 2) centrifuge the supernatant in a microcentrifuge at 16000 g for 10 minutes at 4°C, then collect the supernatant carefully to new 2ml microcentrifuge tubes and stored at -80°C. Cell-free DNA was extracted from 0.5 to 1 ml of plasma by using a QIAamp Circulating Nucleic Acid Kit (Qiagen, Hilden, Germany); extracted DNA was then quantified using an Agilent 2100 Bioanalyzer (High Sensitivity DNA Kit). Genomic DNA for genotyping were purified by using a DNeasy blood and tissue kit (Qiagen, Hilden, Germany) and stored at -20°C.

Purified plasma DNA was prepared into a library following the standard library preparation protocol. For genomic DNA used for genotyping, libraries were captured and enriched according to the manufacturer's protocol. After library preparation, library size distribution and quantification were confirmed using the Agilent Bioanalyzer 2100 and sequencing was performed with a BGISEQ-100 (Thermo Fisher, Proton) or HiSeq 2000 (Illumina, San Diego, CA) instrument.

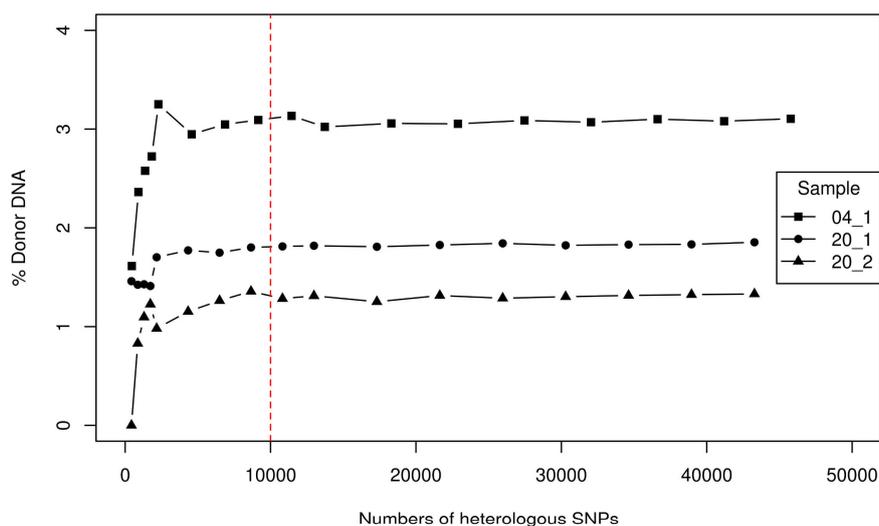
### 2.3. Genotype Array Design

SNP saturation analysis in two lung transplant patients with three plasma samples was carried out by the GTD approach to decide how many heterologous SNP locations should be contained in a SNP array. First, genotyping by ALLi-

nONE array (BGI, Shenzhen, China) with a target region size of approximately 180 megabases (Mb) including the whole exome (44 Mb), a population representative tagSNP region (132 Mb), and the major histocompatibility complex region (4.9 Mb) for a specific population was performed to select the complete set of heterologous SNPs. Then, we selected particular random subsets of those SNPs for 100 repetitions each by increasing the ratio from 0.01 to 0.05 with graduations of 0.01, from 0.05 to 0.3 by 0.05, and from 0.3 to 0.9 by 0.1, and re-calculated the average of the donor fraction (**Figure 2**). The value clearly decreased simultaneously when heterologous SNPs were less than 10,000 and especially when below 5000. This indicated that a mini-screen array used for target capture NGS for genotyping should contain no less than 10 thousand heterologous SNPs. A sliding window of 50 kb across each chromosome was applied. 54,571 (97.36%) SNPs were selected as one SNP per window from the HumanOmniZhongHua-8 Beadchip (Illumina, San Diego, CA). 235 (0.42%) windows that were not located were filled with locations from the HumanCNV370-Duo Beadchip (Illumina, San Diego, CA). The remaining 1243 windows (2.22%) were filled with SNPs with highest minor allele frequency from the dbSNP [19] database. Finally, 56,049 target SNPs were selected and extended to 100 bp on both sides for the oligonucleotide probe design and capturing.

#### 2.4. Quantification of ddcfDNA

High quality reads were firstly aligned to the human reference genome (UCSC hg19), using BWA or TAMPtools (for BGISEQ-100 sequencing data) with default parameters and then PCR duplications were removed by using SAMtools rmdup or BamDuplications tools with default parameters. Next, genomic DNA sequencing reads from pre-transplantation recipient samples were genotyped by



**Figure 2.** Simulation tests by randomly decreasing numbers of informative and control SNPs for sample 04\_1, 20\_1, and 20\_2. Donor percentages remained stable when the number of informative SNPs was no less than 10,000 (red dashed line). The x-axis indicates the number of informative SNPs; y-axis indicates the value of the donor percentage.

SNP calling using SAMtoolsmpileup (-C 50 -E -g -u -I -m 2) or Torrent Variant Caller™ tools with target-seq germline low stringency’s parameters. SNP locations that were not genotyped by the above tools were genotyped autonomously after describing the base-pair information at each chromosomal position by SAMtools mpileup (total depth cut-off 6), based on variant allele frequency (25% - 95% as heterozygote, >95% as homozygote). For sequencing reads of plasma samples, only unique mapping reads were reserved; sequencing information at SNP positions that corresponded to features in genotyping was collected using SAM tools mpileup.

Without the requirement of genotyping the pre-transplant donor genomic DNA, the predicted probability of a population allele such as reference homozygous  $P_{db}(AA)$ , allele homozygous  $P_{db}(BB)$ , and heterozygous  $P_{db}(AB)$  genotype frequencies were calculated in the East Asian population from the 1000 Genomes Project database [20] assuming Hardy-Weinberg equilibrium. For example, if the Reference allele A frequency is 0.6, the other allele B frequency is 0.4, then the genotype frequency of the reference homozygous  $P_{db}(AA)$ , allele homozygous  $P_{db}(BB)$ , and heterozygous  $P_{db}(AB)$  are 0.36, 0.16, and 0.48 respectively. For post-transplant recipient plasma reads at SNP locations (recipient = AA/BB), four conditions of recipient-donor genotype combinations usable for calculating donor signal are considered (assuming no sequencing errors). The predicted probability of the donor is equal to the probability of the population allele when conditions 1 and 4 are [ $P(AA) = P_{db}(AA)$ ;  $P(AB) = P_{db}(AB)$ ;  $P(BB) = P_{db}(BB)$ ] and greater than the probability of the population allele owing to the absence of the homozygous allele, which is the same as in the recipient, when conditions 2 and 3 are [ $P(AB) = P_{db}(AB)/(P(AB) + P(BB))$ ;  $P(BB) = P_{db}(BB)/(P(AB) + P(BB))$ ;  $P(AA) = P_{db}(AA)/(P(AB) + P(AA))$ ] (Table 1). The predicted base call accuracy rate of each plasma read at the target sites, Q, is calculated from the Phred score (Qs) [ $Q = 1 - 10^{-(Qs-33)/10}$ ] when considering sequencing errors in reality. Finally, the ddcfDNA fraction could be calculated using the weighted formula that summarized the particular probability of the heterozygous and homozygous alleles differing between donor and recipient per read at SNP locations (recipient = AA/BB) in each plasma sample [Donor Fraction =  $(2\Sigma P(AB)Q(B) + \Sigma P(BB)Q(B))/(\Sigma P(AB)Q(A) + \Sigma P(AB)Q(B) + \Sigma P(BB)Q(A) + \Sigma P(BB)Q(B))$ ].

**Table 1.** Four conditions of recipient-donor genotype combinations in post-transplant recipient plasma.

| Condition | Reference allele | Recipient genotype | Plasma base | Predicted donor Genotype | Predicted probability of Donor                              |
|-----------|------------------|--------------------|-------------|--------------------------|---|
| 1         | A                | AA                 | A           | AA; AB; BB               | $P_{db}(AA)$ ; $P_{db}(AB)$ ; $P_{db}(BB)$                  |
| 2         | A                | AA                 | B           | AB; BB                   | $P_{db}(AB)/(P(AB) + P(BB))$ ; $P_{db}(BB)/(P(AB) + P(BB))$ |
| 3         | A                | BB                 | A           | AB; AA                   | $P_{db}(BB)/(P(AB) + P(BB))$ ; $P_{db}(AA)/(P(AB) + P(AA))$ |
| 4         | A                | BB                 | B           | BB; AB; AA               | $P_{db}(BB)$ ; $P_{db}(AB)$ ; $P_{db}(AA)$                  |

## 2.5. Data Analysis of Pathogenic Agents

High quality reads of the sequencing data were primarily aligned using BWA mem tools (-k 32 -M -t 10) to the human reference genome (UCSC hg19). The remaining reads (usually less than 5%) that were unable to map to the human genome were secondarily aligned to the human-related microbe genomics database encompassing viruses, bacteria, fungi, and protozoa, that were mainly collected from National Center for Biotechnology Information (NCBI) genome database autonomously using BWA mem tools (-k 32 -M -t 10). The normalization value of one pathogenic abundance, abu, was calculated according to the formula, [abu = total reads of one pathogenic agent/(millions of mapped reads of all pathogenic agents in the same kingdom × kilobases of pathogenic agent genomic sequence)]. Then, the species taxonomy and gene information identifier was annotated from the NCBI database. Finally, infection event for each recipient was determined with elevated levels of relative abundance, abu, in time-points dynamic monitoring instead of pathogen-specific thresholds to discriminate between colonization, infection, and disease.

## 2.6. Statistical Analysis

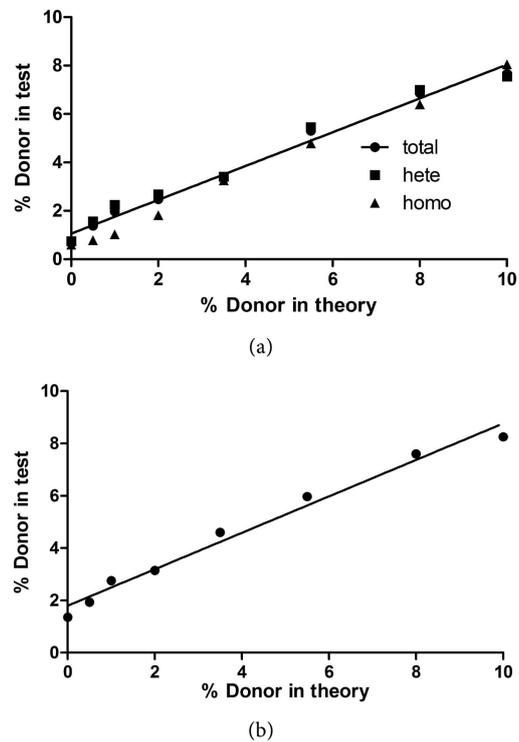
Coefficients of determination (R squared) were performed using Excel (Microsoft). Kolmogorov-Smirnov test and Welch's t test were performed in R 2.15.1. A *P* value of < 0.05 was considered statistically significant. ROC analyses were performed using GraphPad Prism 5.

## 3. Results

### 3.1. Evaluation Testing of the NDTD Approach

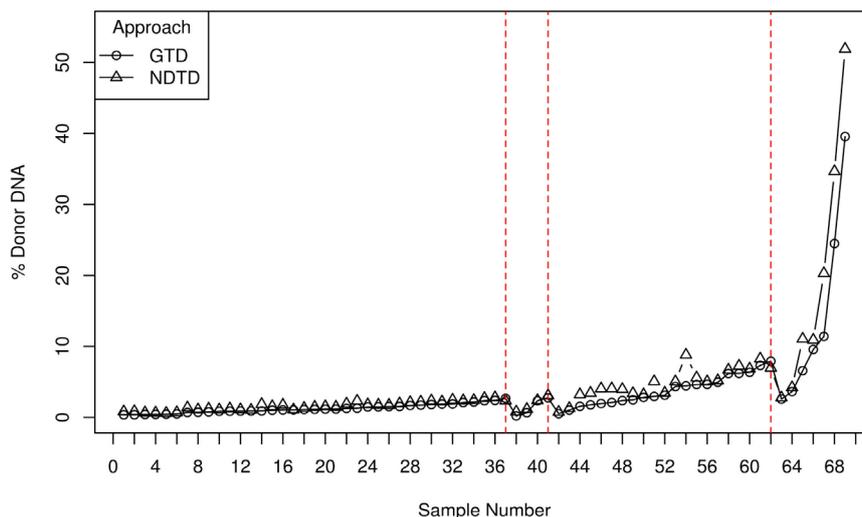
To check the availability of the mini-screen target capture array, genomic DNA from two healthy volunteers (to simulate pre-transplant recipient and donor, respectively) was extracted and sequenced with mean depth = 1.7 gigabases (Gb), representing 110-fold coverage per sample (see online Supplemental **Table S2**), by targeted capture NGS in the selected mini-screen SNP array and then 14,804 (which is greater than 10,000) heterologous SNPs were detected by the GTD approach. Synchronously, we defined "0% donor" as the negative control and mixed cell-free DNA of the volunteer "donor" into the "recipient" with the donor DNA fraction varying from 0.5% to 10% to simulate the post-transplant recipient plasma samples. We sequenced eight simulation samples by extra-low depth WGS: mean depth = 1.59 Gb, 0.5-fold mean coverage per sample (see online Supplemental **Table S3**), with 8192 reads average, located at heterologous SNP sites. Finally, the donor fraction was calculated, showing a significant linear correlation ( $R^2 = 0.99$ , **Figure 3(a)**) between the calculated donor fraction in the test and the donor percentage in theory, indicating sufficiency to measure organ transplant rejection by the mini-screen target capture array.

Next, the donor fraction was re-calculated from the simulation data abandoning donor genotyping information by the NDTD approach as pre-transplant



**Figure 3.** Result of % donor DNA in test calculated by the GTD ((a):  $y = 0.7515x + 0.8986$ ,  $R^2 = 0.9917$ ,  $P < 0.001$ ) and NDTD ((b):  $y = 0.6967x + 1.7955$ ,  $R^2 = 0.9816$ ,  $P < 0.001$ ) approach compared with % donor DNA in theory from the mock sequencing libraries in the evaluation assay. “Total”, “hete” and “homo” represent the calculated % donor DNA in testing using all donor SNPs including both heterozygous and homozygous donor SNPs, using just the heterozygous donor SNPs, and using just the homozygous donor SNPs, respectively.

donor DNA information is likely particularly lacking in long-term and severely affected patients. This demonstrated a significant linear correlation between % Donor in the library and % Donor DNA ( $R^2 = 0.98$ , **Figure 3(b)**  $y = 0.6967x + 1.7955$ ,  $R^2 = 0.9816$   $P < 0.001$ ). Next, the ddcfDNA levels from 69 plasma samples of lung-transplanted recipients that had undergone genotyping by targeted capture NGS in the mini-screen SNP array were calculated respectively by the GTD approach, which requires donor genotyping information, and by the NDTD approach, which is without the need of donor genomic DNA sample, in the same sample cohort (**Figure 4**). Overall, % donor cfDNA calculated by the NDTD approach was coincident with that calculated by the previous GTD approach ( $P = 0.2477$ , Kolmogorov-Smirnov test), which may indicate that the ddcfDNA levels could be quantified by our NDTD approach. Moreover, the slight variation of the two approaches, especially in the group with signs of acute rejection and plasma samples collected within 14 days, may imply possible deviation in SNPs between the public allele database and one particular individual in the amplified signal under the condition of higher ddcfDNA concentrations. It is shown that levels of ddcfDNA increased during the first 14 days post-transplant although this is a period absent of rejection.

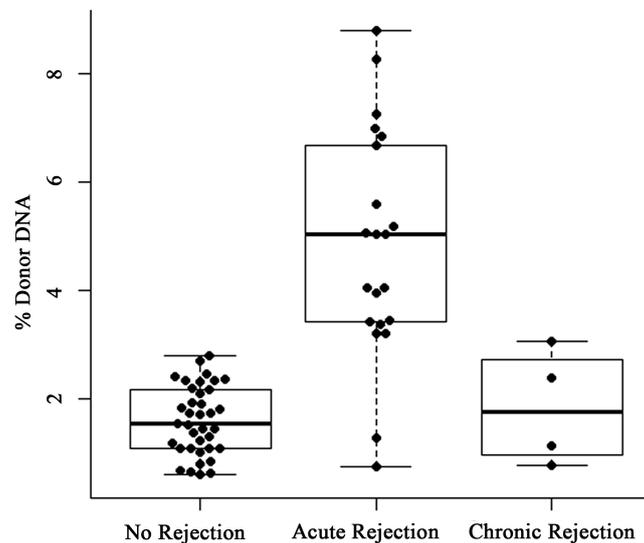


**Figure 4.** ddcfDNA levels from 69 plasma samples in lung transplant recipients. Line graph by the GTD and NDTD approach between cumulative sample numbers (x-axis) and sort ascending % ddcfDNA (y-axis) is shown. 69 samples were divided into 4 groups (red dashed line): 1) No rejection group of plasma sample collected > 14 days included biopsy score equal to A0 (quiescence), no treatment for rejection and no clinical signs of rejection (n = 37). 2) Chronic rejection group of plasma sample collected > 14 days with biopsy proven (n = 4), which excluded the AR group. 3) AR group of plasma sample collected > 14 days included biopsy score  $\geq$  A1 (minimal-to-severe rejection), treatment for AR (steroid pulse therapy), and clinical signs of AR (n = 21). 4) Plasma samples were collected during the first 14 days (n = 7).

### 3.2. Differentiation of Lung Transplant Rejection by the NDTD Approach

We performed quantification of ddcfDNA to monitor acute rejection and detection of the infectious agents simultaneously by the NDTD approach and compared the results with clinical examination in a cohort of 69 recipient plasma samples collected from 16 lung transplantation patients. For rejection surveillance, samples collected during the first 14 days post-transplant, which is a period absent of rejection and may exhibit elevated levels of ddcfDNA, were excluded. The ddcfDNA levels (n = 62) were significantly different between the AR group ( $4.83 \pm 2.11$  %, mean  $\pm$  SD) and the non-rejection group ( $1.61 \pm 0.63$  %, mean  $\pm$  SD) ( $P < 0.0001$ , Welch's t test). Findings were validated by biopsies (n = 17) and clinical indications (n = 45) (Figure 5). With the cut-off value of 2.999, this method exhibited 90.48% sensitivity (95% CI, 69.62% - 98.83%), 100% specificity (95% CI, 91.59% - 100%), and AUC 0.9266 (95% CI, 0.8277 - 1.026) in lung transplantation (see online Supplemental Table S4). However, the difference was not significant between the non-rejection group and the chronic rejection group ( $P = 0.9340$ , Welch's t test), implying that additional chronic rejection events should be observed in further studies (Figure 5). According to these results, we may find that ddcfDNA levels from lung allograft recipients increase when rejection events occur, especially during acute rejection.

For detection of infectious agents, whole genome sequencing reads were used



**Figure 5.** Scatter plot for % ddcfDNA of lung post-transplantation plasma samples. Plasma samples were collected > 14 days ( $n = 62$ ) post-transplant and divided into 3 groups: 1) No rejection ( $n = 37$ ). 2) AR group ( $n = 21$ ). 3) Chronic rejection ( $n = 4$ ). Generally, the mean % ddcfDNA in the acute rejection group was higher than that in the non-rejection group and that in the chronic rejection group.

to evaluate the virus, bacteria, and fungus infection concurrently after removing host reads of human sequence. We found positive infection status that was validated by the clinical tests, including cytomegalovirus (CMV) in 5 recipients (patient No. 03, 04, 06, 13 and 15); bacterial agents, such as *Acinetobacterbaumannii*, *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* in 6 recipients (patient No. 03, 04, 05, 08, 13 and 16); and the fungus *Aspergillus fumigatus* in 1 recipient (patient No. 05). No reads matched the CMV genomic sequence, with zero abundance in 4 CMV-negative recipients (*i.e.*, concentration less than 1000 copies/ml) (patient No. 10, 12, 14 and 16). We also found additional agents including adenovirus in 6 recipient (patient No. 01, 03, 04, 11, 12 and 13) and CMV in 3 recipients (patient No. 01, 05 and 07) (see online Supplemental **Table S4**).

#### 4. Discussion

The distinction between rejection and infection after solid organ transplantation has always presented a problem for clinical therapy because the clinical symptoms are sometimes similar. There is no reliable marker for AR monitoring, which is limited to detecting restricted pathogen species in the clinic; detection of rejection and infection only using the same data from blood samples thus presents an exciting prospect. Our results demonstrate that the NDTD approach without donor-derived material has the ability to monitor acute rejection by quantification of ddcfDNA and to detect the infectious agents simultaneously. The approach was composed of two processes: genotyping of recipient pre-transplantation and ddcfDNA detection of recipient post-transplantation, which can be carried out on different sequencing platforms with automated data analysis. The geno-

typing procedure implemented by targeted-capture NGS in a mini-screen SNP array requires less sequencing data and would reduce cost. The whole procedure including genotyping required 3 days, needing only 1.5 days if the genotyping step had been done in advance. Our approach showed high consistency with the previous GTD approach as shown in the validation step, which included both simulation tests and detection of events in lung transplantation. Additionally, the slight variation of the two approaches may imply an individual difference; thus, a more comprehensive and convincing public allele frequency database such as dbSNP may be required in the future. Subsequently, verification of the lung transplant cohort indicated that the differentiation of ddcfDNA between no rejection and rejection groups was obvious, especially in the case of acute rejection.

Notably, the subsequent sequencing data annotation was also indicative of pathogenic agents such as virus, bacteria, and fungus agents. Out of all the screened infectious agents, this approach delivered an advantage in virus testing, especially for CMV infection, which posed the most common threat for infectious complications after lung transplantation. To build up a more reasonable rejection-infection differentiated model, potential modifications include: 1) increasing the sequencing depth of plasma samples to capture more pathogen materials, although the current sequencing depth is sufficient for rejection monitoring; 2) building pathogen-specific thresholds to discriminate between colonization, infection, and disease; and 3) expanding sample types such as sputum, bronchoalveolar lavage fluid, nasopharyngeal swabs, and plasma samples.

## 5. Conclusion

In conclusion, these findings suggest that the NDTD approach has the ability of diagnosis and discrimination between rejection and infection post-transplant in lung transplantation and may be applied to other types of solid organ transplantation (such as heart, kidney, and liver) where ddcfDNA may also exist in the recipient's plasma. It demonstrates a cost-effective and noninvasive sequencing approach without the requirement of donor-derived genotyping, which will better satisfy the needs of clinical situations and show a wider range of clinical application to accelerate the development of precautionary molecular diagnosis in solid organ transplantation.

## Acknowledgements

This project is supported by Guangzhou Key Laboratory of Cancer Trans-Omics Research (GZ2012, NO348), Guangdong Science and Technology Project (2014B020212006, 201400000001-2), and Shenzhen Engineering Laboratory for Clinical Molecular Diagnostic. We would like to acknowledge Shengjian Yuan, Mingjie Huang, Xiyan Xiang, Wei Wei, and Rongqing Deng for their technical assistance in sequencing of DNA.

## Authors' Disclosures

The authors declare no conflict of interest.

## References

- [1] Castleberry, A.W., Worni, M., Kuchibhatla, M., Lin, S.S., Snyder, L.D., Shofer, S.L., *et al.* (2013) A Comparative Analysis of Bronchial Stricture after Lung Transplantation in Recipients with and without Early Acute Rejection. *Annals of Thoracic Surgery*, **96**, 1008-1017. <https://doi.org/10.1016/j.athoracsur.2013.01.104>
- [2] Christie, J.D., Edwards, L.B., Aurora, P., Dobbels, F., Kirk, R., Rahmel, A.O., *et al.* (2009) The Registry of the International Society for Heart and Lung Transplantation: Twenty-Sixth Official Adult Lung and Heart-Lung Transplantation Report-2009. *Journal of Heart and Lung Transplantation*, **28**, 1031-1049. <https://doi.org/10.1016/j.healun.2009.08.004>
- [3] Bhorade, S.M., Husain, A.N., Liao, C., Li, L.C., Ahya, V.N., Baz, M.A., *et al.* (2013) Interobserver Variability in Grading Transbronchial Lung Biopsy Specimens after Lung Transplantation. *Chest*, **143**, 1717-1724. <https://doi.org/10.1378/chest.12-2107>
- [4] Hachem, R.R. (2009) Lung Allograft Rejection: Diagnosis and Management. *Current Opinion in Organ Transplantation*, **14**, 477-482. <https://doi.org/10.1097/MOT.0b013e32832fb981>
- [5] Gordon, I.O., Bhorade, S., Vigneswaran, W.T., Garrity, E.R. and Husain, A.N. (2012) SaLUTaRy: Survey of Lung Transplant Rejection. *Journal of Heart and Lung Transplantation*, **31**, 972-979. <https://doi.org/10.1016/j.healun.2012.05.014>
- [6] Borewicz, K., Pragman, A.A., Kim, H.B., Hertz, M., Wendt, C. and Isaacson, R.E. (2013) Longitudinal Analysis of the Lung Microbiome in Lung Transplantation. *FEMS Microbiology Letters*, **339**, 57-65. <https://doi.org/10.1111/1574-6968.12053>
- [7] Lo, Y.M., Tein, M.S., Pang, C.C., Yeung, C.K., Tong, K.L. and Hjelm, N.M. (1998) Presence of Donor-Specific DNA in Plasma of Kidney and Liver-Transplant Recipients. *Lancet*, **351**, 1329-1330. [https://doi.org/10.1016/S0140-6736\(05\)79055-3](https://doi.org/10.1016/S0140-6736(05)79055-3)
- [8] Garcia Moreira, V., Prieto García, B., Baltar Martín, J.M., Ortega Suárez, F. and Alvarez, F.V. (2009) Cell-Free DNA as a Noninvasive Acute Rejection Marker in Renal Transplantation. *Clinical Chemistry*, **55**, 1958-1966. <https://doi.org/10.1373/clinchem.2009.129072>
- [9] Gadi, V.K., Nelson, J.L., Boespflug, N.D., Guthrie, K.A. and Kuhr, C.S. (2006) Soluble Donor DNA Concentrations in Recipient Serum Correlate with Pancreas-Kidney Rejection. *Clinical Chemistry*, **52**, 379-382. <https://doi.org/10.1373/clinchem.2005.058974>
- [10] Snyder, T.M., Khush, K.K., Valantine, H.A. and Quake, S.R. (2011) Universal Non-invasive Detection of Solid Organ Transplant Rejection. *Proceedings of the National Academy of Sciences of USA*, **108**, 6229-6234. <https://doi.org/10.1073/pnas.1013924108>
- [11] Beck, J., Bierau, S., Balzer, S., Andag, R., Kanzow, P., Schmitz, J., *et al.* (2013) PCR for Rapid Quantification of Donor DNA in the Circulation of Transplant Recipients as a Potential Universal Biomarker of Graft Injury. *Clinical Chemistry*, **59**, 1732-1741. <https://doi.org/10.1373/clinchem.2013.210328>
- [12] Romanov, V., Whyard, T.C., Waltzer, W.C. and Darras, F.S. (2014) A Noninvasive Assay for Monitoring Renal Allograft Status. *Journal of Biological Methods*, **1**, e2. <https://doi.org/10.14440/jbm.2014.19>
- [13] Adamek, M., Opelz, G., Klein, K., Morath, C. and Tran, T.H. (2016) A Fast and

Simple Method for Detecting and Quantifying Donor-Derived Cell-Free DNA in Sera of Solid Organ Transplant Recipients as a Biomarker for Graft Function. *Clinical Chemistry and Laboratory Medicine*, **54**, 1147-1155.

<https://doi.org/10.1515/cclm-2015-0622>

- [14] De Vlaminck, I., Valantine, H.A., Snyder, T.M., Strehl, C., Cohen, G., Luikart, H., *et al.* (2014) Circulating Cell-Free DNA Enables Noninvasive Diagnosis of Heart Transplant Rejection. *Science Translational Medicine*, **6**, 241ra77. <https://doi.org/10.1126/scitranslmed.3007803>
- [15] Oellerich, M., Schütz, E., Kanzow, P., Schmitz, J., Beck, J., Kollmar, O., *et al.* (2014) Use of Graft-Derived Cell-Free DNA as an Organ Integrity Biomarker to Reexamine Effective Tacrolimus trough Concentrations after Liver Transplantation. *Therapeutic Drug Monitoring*, **36**, 136-140. <https://doi.org/10.1097/FTD.0000000000000044>
- [16] Sigdel, T.K., Vitalone, M.J., Tran, T.Q., Dai, H., Hsieh, S.C., Salvatierra, O., *et al.* (2013) A Rapid Noninvasive Assay for the Detection of Renal Transplant Injury. *Transplantation*, **96**, 97-101. <https://doi.org/10.1097/TP.0b013e318295ee5a>
- [17] De Vlaminck, I., Martin, L., Kertesz, M., Patel, K., Kowarsky, M., Strehl, C., *et al.* (2015) Noninvasive Monitoring of Infection and Rejection after Lung Transplantation. *Proceedings of the National Academy of Sciences of USA*, **112**, 13336-13341. <https://doi.org/10.1073/pnas.1517494112>
- [18] Lo, Y.M. (2011) Transplantation Monitoring by Plasma DNA Sequencing. *Clinical Chemistry*, **57**, 941-942. <https://doi.org/10.1373/clinchem.2011.166686>
- [19] Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M. and Sirotkin, K. (2001) dbSNP: The NCBI Database of Genetic Variation. *Nucleic Acids Research*, **29**, 308-311. <https://doi.org/10.1093/nar/29.1.308>
- [20] The 1000 Genomes Project Consortium (2015) A Global Reference for Human Genetic Variation. *Nature*, **526**, 68-74. <https://doi.org/10.1038/nature15393>

## Supplemental

**Table S1.** Lung transplant recipient demographic characteristics.

| Characteristic                               |             |
|--|-------------|
| Adult recipients, n                          | 16          |
| Age at time of transplant                    |             |
| Mean $\pm$ SD                                | 56 $\pm$ 15 |
| Male sex, no. (%)                            | 14 (88)     |
| Type of lung transplanted, no. (%)           |             |
| Both   | 12 (75)     |
| Left   | 2 (12.5)    |
| Right  | 2 (12.5)    |
| Indication for lung transplantation, no. (%) |             |
| Fibrosis                                     | 7 (43.75)   |
| Chronic pulmonary obstruction                | 4 (25)      |
| Inflammation                                 | 6 (37.5)    |
| Silicon lung disease                         | 1 (6.25)    |
| Other  | 1 (6.25)    |
| Maintenance immunosuppression, no. (%)       |             |
| Cyclosporine                                 | 1 (6.25)    |
| Tacrolimus                                   | 16 (100)    |
| Hospitalization status, no. (%)              |             |
| Inpatient                                    | 11 (68.75)  |
| Outpatient                                   | 5 (31.25)   |
| Transbronchial biopsy, no. (%)               |             |
| Yes  | 12 (75)     |
| No   | 4 (25)      |

**Table S2.** Performance of the target capture sequencing quality.

| Sample   | Recipient  | Donor      |
|--|------------|------------|
| Initial bases on target                              | 7,705,014  | 7,705,014  |
| Total effective reads                                | 12,789,800 | 11,387,756 |
| Total effective yield (Mb)                           | 1705.48    | 1772.44    |
| Average read length (bp)                             | 133.35     | 155.64     |
| Effective sequences on target (Mb)                   | 912.85     | 844.49     |
| Fraction of effective bases on target                | 53.50%     | 47.60%     |
| Average sequencing depth on target                   | 118.48     | 109.6      |
| Base covered on target                               | 7,413,741  | 7,621,021  |
| Coverage of target region                            | 96.20%     | 98.90%     |
| Fraction of target covered with at least 20 $\times$ | 85.00%     | 94.00%     |
| Mapping rate   | 98.85%     | 99.02%     |
| Duplicate rate                                       | 19.86%     | 20%        |

**Table S3.** Statistics of the simulation plasma cell-free DNA libraries.

| <b>% Donor</b>        | <b>0</b>   | <b>0.5</b> | <b>1</b>  | <b>2</b>   | <b>3.5</b> | <b>5.5</b> | <b>8</b>   | <b>10</b>  |
|-----------------------|------------|------------|-----------|------------|------------|------------|------------|------------|
| Total reads           | 11,602,242 | 9,742,563  | 9,598,491 | 10,676,827 | 10,060,264 | 11,595,993 | 11,862,646 | 16,041,446 |
| Aligned               | 11,512,905 | 9,663,648  | 9,524,583 | 10,598,886 | 9,992,860  | 11,518,300 | 11,780,794 | 13,655,526 |
| Unique                | 10,195,569 | 8,520,064  | 8,433,245 | 9,555,528  | 9,025,429  | 10,540,346 | 10,533,601 | 13,655,526 |
| Total reads with SNPs | 9111       | 7299       | 7350      | 8071       | 7580       | 8600       | 8826       | 8703       |
| Heterozygous SNPs     |            |            |           |            |            |            |            |            |
| Total reads           | 6951       | 5636       | 5703      | 6201       | 5800       | 6609       | 6819       | 6702       |
| Recipient reads       | 6914       | 5584       | 5632      | 6111       | 5697       | 6419       | 6567       | 6424       |
| Donor reads           | 26         | 44         | 64        | 83         | 99         | 180        | 238        | 252        |
| Errors                | 11         | 8          | 7         | 7          | 4          | 10         | 14         | 26         |
| Homozygous SNPs       |            |            |           |            |            |            |            |            |
| Total reads           | 2160       | 1663       | 1647      | 1870       | 1780       | 1991       | 2007       | 2001       |
| Recipient reads       | 2143       | 1645       | 1629      | 1832       | 1720       | 1890       | 1873       | 1826       |
| Donor reads           | 13         | 13         | 17        | 34         | 58         | 95         | 128        | 160        |
| Errors                | 4          | 5          | 1         | 4          | 2          | 6          | 6          | 15         |

**Table S4.** Sensitivity and specificity performance by the NDTD approach in lung transplantation in ROC curve analysis.

| <b>Cutoff</b> | <b>Sensitivity%</b> | <b>95% CI</b>    | <b>Specificity%</b> | <b>95% CI</b>      | <b>Likelihood ratio</b> |
|---------------|---------------------|------------------|---------------------|--------------------|-------------------------|
| >0.6213       | 100                 | 83.89% to 100.0% | 2.381               | 0.06026% to 12.57% | 1.02                    |
| >0.6490       | 100                 | 83.89% to 100.0% | 4.762               | 0.5820% to 16.16%  | 1.05                    |
| >0.6714       | 100                 | 83.89% to 100.0% | 7.143               | 1.498% to 19.48%   | 1.08                    |
| >0.7136       | 100                 | 83.89% to 100.0% | 9.524               | 2.656% to 22.62%   | 1.11                    |
| >0.7757       | 95.24               | 76.18% to 99.88% | 9.524               | 2.656% to 22.62%   | 1.05                    |
| >0.8251       | 95.24               | 76.18% to 99.88% | 11.9                | 3.981% to 25.63%   | 1.08                    |
| >0.9344       | 95.24               | 76.18% to 99.88% | 16.67               | 6.974% to 31.36%   | 1.14                    |
| >1.053        | 95.24               | 76.18% to 99.88% | 19.05               | 8.601% to 34.12%   | 1.18                    |
| >1.079        | 95.24               | 76.18% to 99.88% | 21.43               | 10.30% to 36.81%   | 1.21                    |
| >1.086        | 95.24               | 76.18% to 99.88% | 23.81               | 12.05% to 39.45%   | 1.25                    |
| >1.096        | 95.24               | 76.18% to 99.88% | 26.19               | 13.86% to 42.04%   | 1.29                    |
| >1.149        | 95.24               | 76.18% to 99.88% | 28.57               | 15.72% to 44.58%   | 1.33                    |
| >1.210        | 95.24               | 76.18% to 99.88% | 30.95               | 17.62% to 47.09%   | 1.38                    |
| >1.250        | 95.24               | 76.18% to 99.88% | 33.33               | 19.57% to 49.55%   | 1.43                    |
| >1.288        | 90.48               | 69.62% to 98.83% | 33.33               | 19.57% to 49.55%   | 1.36                    |
| >1.333        | 90.48               | 69.62% to 98.83% | 35.71               | 21.55% to 51.97%   | 1.41                    |
| >1.411        | 90.48               | 69.62% to 98.83% | 40.48               | 25.63% to 56.72%   | 1.52                    |
| >1.457        | 90.48               | 69.62% to 98.83% | 42.86               | 27.72% to 59.04%   | 1.58                    |
| >1.491        | 90.48               | 69.62% to 98.83% | 45.24               | 29.85% to 61.33%   | 1.65                    |
| >1.531        | 90.48               | 69.62% to 98.83% | 47.62               | 32.00% to 63.58%   | 1.73                    |

## Continued

|        |       |                   |       |                  |       |
|--------|-------|-------------------|-------|------------------|-------|
| >1.630 | 90.48 | 69.62% to 98.83%  | 50    | 34.19% to 65.81% | 1.81  |
| >1.732 | 90.48 | 69.62% to 98.83%  | 52.38 | 36.42% to 68.00% | 1.9   |
| >1.744 | 90.48 | 69.62% to 98.83%  | 54.76 | 38.67% to 70.15% | 2     |
| >1.773 | 90.48 | 69.62% to 98.83%  | 57.14 | 40.96% to 72.28% | 2.11  |
| >1.820 | 90.48 | 69.62% to 98.83%  | 61.9  | 45.64% to 76.43% | 2.38  |
| >1.871 | 90.48 | 69.62% to 98.83%  | 64.29 | 48.03% to 78.45% | 2.53  |
| >1.918 | 90.48 | 69.62% to 98.83%  | 66.67 | 50.45% to 80.43% | 2.71  |
| >2.021 | 90.48 | 69.62% to 98.83%  | 71.43 | 55.42% to 84.28% | 3.17  |
| >2.145 | 90.48 | 69.62% to 98.83%  | 73.81 | 57.96% to 86.14% | 3.45  |
| >2.193 | 90.48 | 69.62% to 98.83%  | 76.19 | 60.55% to 87.95% | 3.8   |
| >2.256 | 90.48 | 69.62% to 98.83%  | 80.95 | 65.88% to 91.40% | 4.75  |
| >2.324 | 90.48 | 69.62% to 98.83%  | 83.33 | 68.64% to 93.03% | 5.43  |
| >2.345 | 90.48 | 69.62% to 98.83%  | 85.71 | 71.46% to 94.57% | 6.33  |
| >2.355 | 90.48 | 69.62% to 98.83%  | 88.1  | 74.37% to 96.02% | 7.6   |
| >2.382 | 90.48 | 69.62% to 98.83%  | 90.48 | 77.38% to 97.34% | 9.5   |
| >2.429 | 90.48 | 69.62% to 98.83%  | 92.86 | 80.52% to 98.50% | 12.67 |
| >2.576 | 90.48 | 69.62% to 98.83%  | 95.24 | 83.84% to 99.42% | 19    |
| >2.746 | 90.48 | 69.62% to 98.83%  | 97.62 | 87.43% to 99.94% | 38    |
| >2.999 | 90.48 | 69.62% to 98.83%  | 100   | 91.59% to 100.0% |       |
| >3.207 | 85.71 | 63.66% to 96.95%  | 100   | 91.59% to 100.0% |       |
| >3.293 | 80.95 | 58.09% to 94.55%  | 100   | 91.59% to 100.0% |       |
| >3.399 | 76.19 | 52.83% to 91.78%  | 100   | 91.59% to 100.0% |       |
| >3.431 | 71.43 | 47.82% to 88.72%  | 100   | 91.59% to 100.0% |       |
| >3.697 | 66.67 | 43.03% to 85.41%  | 100   | 91.59% to 100.0% |       |
| >3.999 | 61.9  | 38.44% to 81.89%  | 100   | 91.59% to 100.0% |       |
| >4.048 | 57.14 | 34.02% to 78.18%  | 100   | 91.59% to 100.0% |       |
| >4.546 | 52.38 | 29.78% to 74.29%  | 100   | 91.59% to 100.0% |       |
| >5.041 | 47.62 | 25.71% to 70.22%  | 100   | 91.59% to 100.0% |       |
| >5.049 | 42.86 | 21.82% to 65.98%  | 100   | 91.59% to 100.0% |       |
| >5.115 | 38.1  | 18.11% to 61.56%  | 100   | 91.59% to 100.0% |       |
| >5.377 | 33.33 | 14.59% to 56.97%  | 100   | 91.59% to 100.0% |       |
| >6.126 | 28.57 | 11.28% to 52.18%  | 100   | 91.59% to 100.0% |       |
| >6.754 | 23.81 | 8.218% to 47.17%  | 100   | 91.59% to 100.0% |       |
| >6.901 | 19.05 | 5.446% to 41.91%  | 100   | 91.59% to 100.0% |       |
| >7.105 | 14.29 | 3.049% to 36.34%  | 100   | 91.59% to 100.0% |       |
| >7.744 | 9.524 | 1.175% to 30.38%  | 100   | 91.59% to 100.0% |       |
| >8.521 | 4.762 | 0.1205% to 23.82% | 100   | 91.59% to 100.0% |       |