


**METHOD**

# FF-QuantSC: accurate quantification of fetal fraction by a neural network model

Yuying Yuan<sup>1</sup> | Xianghua Chai<sup>1</sup>  | Na Liu<sup>2</sup> | Bida Gu<sup>1</sup> | Shengting Li<sup>3</sup> |  
Ya Gao<sup>4,5</sup> | Lijun Zhou<sup>1</sup> | Qiang Liu<sup>1</sup> | Fan Yang<sup>1</sup> | Jingjuan Liu<sup>1</sup> | Jiao Qiu<sup>1</sup> |  
Jinjin Zhang<sup>1</sup> | Yumei Hou<sup>1</sup> | Miaolan Cen<sup>1</sup> | Zhongming Tian<sup>6</sup> | Weijiang Tang<sup>7</sup> |  
Hongyun Zhang<sup>1</sup> | Fang Chen<sup>3</sup> | Ye Yin<sup>2</sup> | Wei Wang<sup>3</sup>

<sup>1</sup>Clinical laboratory of BGI Health, BGI-Shenzhen, Shenzhen, China

<sup>2</sup>BGI Genomics, BGI-Shenzhen, Shenzhen, China

<sup>3</sup>MGI, BGI-Shenzhen, Shenzhen, China

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

<sup>5</sup>China National GeneBank, BGI-Shenzhen, Shenzhen, China

<sup>6</sup>Tianjin Medical Laboratory, BGI-Tianjin, BGI-Shenzhen, Tianjin, China

<sup>7</sup>BGI-Wuhan Clinical Laboratories, BGI-Shenzhen, Wuhan, China

**Correspondence**

Wei Wang, Building No.11 2, Beishan Industrial Zone, Yantian District, Shenzhen 518083, China.  
Email: wangw@genomics.cn

**Abstract**

**Background:** Noninvasive prenatal testing (NIPT) is one of the most commonly employed clinical measures for screening of fetal aneuploidy. Fetal Fraction (*ff*) has been demonstrated to be one of the key factors affecting the performance of NIPT. Accurate quantification of *ff* plays vital role in NIPT.

**Methods:** In this study, we present a new approach, the accurate Quantification of Fetal Fraction with Shallow-Coverage sequencing of maternal plasma DNA (FF-QuantSC), for the estimation of *ff* in NIPT. The method employs neural network model and utilizes differential genomic patterns between fetal and maternal genomes to quantify *ff*.

**Results:** Our results show that the predicted *ff* by FF-QuantSC exhibit high correlation with the Y chromosome-based method on male pregnancies, and achieves the highest accuracy compared with other *ff* estimation approaches. We also demonstrate that the model generates statistically similar results on both male and female pregnancies.

**Conclusion:** FF-QuantSC achieves high accuracy in *ff* quantification. The method is suitable for application in both male and female pregnancies. Since the method does not require additional information upon NIPT routines, it can be easily incorporated into current NIPT settings without causing extra costs. We believe that FF-QuantSC shall provide valuable additions to NIPT.

**KEYWORDS**

female pregnancy, fetal fraction, FF-QuantSC, neural network, NIPT, shallow-coverage sequencing

Yuying Yuan, Xianghua Chai, and Na Liu should be considered joint first author.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals, Inc.

## 1 | INTRODUCTION

In late 20th century, use of plasma DNA in molecular diagnosis has been demonstrated as a valuable potential (Sorenson, Pribish, Valone, & Memoli, 1994; D. Lo et al., 1997). One of the interests led to the discovery of the presence of fetal circulating cell-free DNA (ccfDNA) in maternal plasma during pregnancy (D. Lo et al., 1997), which revolutionized the clinical application of noninvasive prenatal testing (NIPT) of fetal chromosomal aneuploidies and microdeletion/microduplication syndromes. Statistics shows that chromosomal abnormalities occur approximately 1 in 150 live births and Down syndrome (trisomy 21) occurs approximately 1 in 800 live births (Gregg et al., 2016; Mercer, 2016). The risk even increases as the mother ages (Mercer, 2016). As a result, screening of common fetal chromosomal abnormalities during pregnancy has been required by governments in many countries. The proportion of fetal ccfDNA in maternal plasma, known as the Fetal Fraction ( $ff$ ), has been shown to be one of the fundamental factors affecting the performance of NIPT (Canick, Palomaki, Kloza, Lambert-Messerlian, & Haddow, 2013). Previous study has revealed that  $ff$  is closely related to maternal weight and gestational age (Kinnings et al., 2015; Wang et al., 2013; Wataganara, Peter, Messerlian, Borgatta, & Bianchi, 2004), and the value of  $ff$  typically ranges from 3% to over 30% throughout the pregnancy (Canick et al., 2013). Our previous study concluded that  $ff$  should attain 3.5% for a reliable NIPT performance, and proceeding of NIPT with lower  $ff$  was associated with elevated false-positive and false-negative rates (Zhang et al., 2015).

$ff$  may be estimated by contrasting various differential patterns between the maternal genome and the fetal genome, such as the differential methylation pattern of specific epigenetic markers (Nygren et al., 2010), the differential genotype of selected SNPs (Sparks, Struble, Wang, Song, & Oliphant, 2012; Y. M. D. Lo et al., 2010; Zimmermann et al., 2012), the size differences between fetal and maternal DNA fragments (Yu et al., 2014), and the relative coverage of the Y chromosome in male pregnancies (Lun et al., 2008). Each of these approaches has limitations yet to be overcome. The methylation-based method requires whole-genome bisulfite sequencing and the size-based method requires paired-end next-generation sequencing. Both sequencing requirements add further cost to routine NIPT. The SNP-based method, although accurate, utilizes extra parental SNP information, which may not be readily available. Finally, the Y chromosome method is not applicable for female fetuses.

In this report, we present a new  $ff$  estimation method to circumvent these limitations. The method, named the accurate Quantification of  $ff$  with Shallow-Coverage sequencing of maternal plasma DNA (FF-QuantSC), employs an artificial neural network model. It is developed upon a two-step hypothesis: the distributions of sequencing data between maternal and fetal ccfDNA are different, and the difference is

related to  $ff$ . While evidence from previous study provides support to the hypothesis (Kim et al., 2015), we here demonstrate a further prove. Our result shows that FF-QuantSC achieves accurate and cost-effective estimation of  $ff$  on routine NIPT settings. More importantly, FF-QuantSC does not rely on information from fetal Y chromosome and thus also permits estimation of  $ff$  of female fetuses.

## 2 | MATERIALS AND METHODS

### 2.1 | Editorial policies and ethical considerations

This study was approved by BGI-IRB (BGI's institutional review board on bioethics and biosafety) and the sequencing data of all samples were obtained under BGI-IRB approved protocols BGI-IRB 17,048 on the BGISEQ-1000 platform.

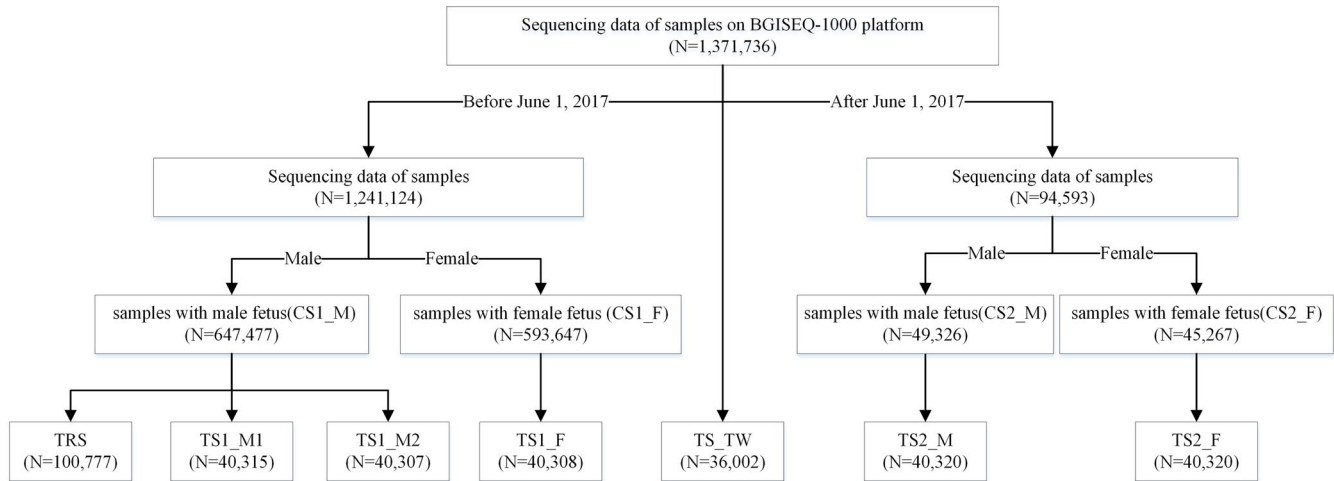
The pregnant women, whose samples used in this study, all signed the informed consents before sampling and agreed that the sequencing data could be used for research after anonymization. All women were of 12–24 weeks ( $17.62 \pm 4.07$ ) after gestation at the time of sample collection. Gender information of fetuses was obtained by following-up visit of testing subjects.

### 2.2 | Training data and testing data

Degree of freedom of the neural network model was calculated by the Vapnik–Chervonenkis (VC) dimension (Cherkassky, 1997). Since a rule of thumb suggested an optimal training data set 10 times of the model degree of freedom, we randomly selected 100,777 samples tested before June 1, 2017 with male fetuses as the training set (TRS). To test the robustness of FF-QuantSC, we designed six testing sets and performed individual model testing on each set. Two testing sets (TS1\_M1 and TS1\_M2) were randomly selected from the remaining samples with male fetuses that were tested before June 1, 2017. A set containing female pregnancies tested before June 1, 2017 was also selected as TS1\_F. Another two testing sets (TS2\_M and TS2\_F) were randomly generated from male and female pregnancies tested after June 1, 2017. A last set of twin pregnancies was randomly selected regardless of the testing time. Each testing set contains approximately 40,000 samples (Figure 1).

### 2.3 | Sequencing data distribution between maternal and fetal ccfDNA

To confirm the two-step hypothesis, we conducted sequencing of plasma samples from 24 randomly selected male



**FIGURE 1** Flowchart of training set and testing set partition. Samples of singleton pregnancies were first partitioned by the time of NIPT analysis and further divided by gender. Training and testing sets were then randomly selected from each group as indicated. Samples of twin pregnancies were selected independently from the original data pool

pregnancies with a conventional protocol and a fetal-fraction enriched protocol in parallel. The conventional procedures included plasma separation, DNA extraction, sequencing library construction, library quality control, and sequencing. Based on Lo and his colleagues' work (Y. M. D. Lo et al., 2010), fetal-fraction enrichment was achieved by segment selection of DNA fragments of 100–150 bp in length. A median of 8.25 (3.41–12.3) million uniquely aligned sequencing reads were obtained for each sample for subsequent analysis. These reads were partitioned into genomic regions with 60 kb windows according to their mapping positions, from which the total number of mapped reads within each window was calculated. T-test was then applied onto each window to identify differences in read counts between the two groups of samples.

## 2.4 | General flow of FF-QuantSC

The overall analysis flow of FF-QuantSC is summarized in Figure 2. All DNA sequencing was conducted on BGISEQ-1000 platform to produce 28bp single-end reads. Raw sequencing data were aligned to the human reference genome (GRCh37) by BWA (V0.7.7-r441) (Li & Durbin, 2010). Reads mapped with mismatches or more than one hit were removed. The retained effective reads were partitioned into continuous genomic windows of 60kb in length. We removed windows constantly showing no coverage and applied PCA for feature selection. Normalization was performed by dividing data of each selected feature by the sum of data of all features. The resulted data were further standardized by within-sample z-score transformation. This generated a final feature matrix, with each row representing a sample and each column representing a selected feature. Finally, testing

results from all six testing sets were compared to evaluate the predictive capability of the trained model.

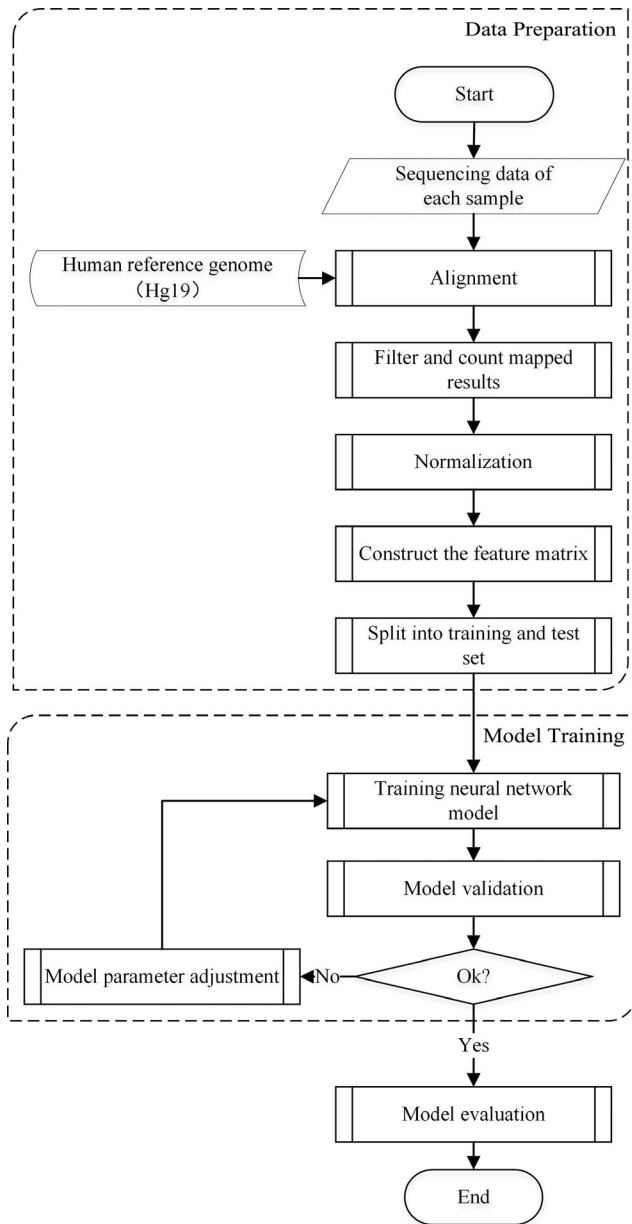
## 2.5 | Model training

By comparing the fitting capability of different network structures on the training dataset, a fully connected neural network with a single hidden layer and 128 neurons was selected. Several model parameters were to be finely tuned for various purposes. The batch size was empirically set to be identical as the number of Graphics Processing Unit (GPU) cores to give the best GPU accelerating performance. Learning rate of approaches for gradient diffusion elimination was set to achieve fast model convergence. Finally, in order to balance the fitting and generalization capabilities as well as to improve the predictive accuracy of the network, two penalty factors (Goeman, 2010) and the dropout rate were autotuned by *Hyperopt* (Bergstra, Yamins, & Cox, 2013). Please refer to S1 Text and Figure S1 for more details.

## 3 | RESULTS

### 3.1 | Sequencing data distribution between maternal and fetal ccfDNA

Differential analysis of the 24 samples was summarized in Figure 3. Among all sequenced genomic regions, 8,107 of 46,981 windows were tested to show significant difference ( $p < .05$ ) in sequencing read counts between the untreated ccfDNA samples and the fetal-fraction enriched ccfDNA samples. This makes up roughly 17% of the entire genome (N-regions excluded). Interestingly, instead of a sparsely



**FIGURE 2** Flowchart of FF-QuantSC method. (1) Data preparation. Sequencing reads were mapped to reference sequence (GRCh37). Mapped reads were then filtered and counted to construct feature matrixes of different datasets. (2) Model training. The network model was trained by planning procedures. (3) Model evaluation. After model training, predictive results of all test datasets were assessed to evaluate the performance of FF-QuantSC

random spread, these windows clustered into small groups containing up to dozens of continuous or extremely close windows, resulting in notable differential regions (*SD-Region*) on the genome. As contrasted in red from the insignificant regions (*NSD-Region*) in the figure, most of these differential regions were evenly distributed throughout the genome. However, we also noted that certain chromosomal regions harvested a higher density of the differences, such as 3p, 9q, 11q, and 19p. Since the enrichment protocol was

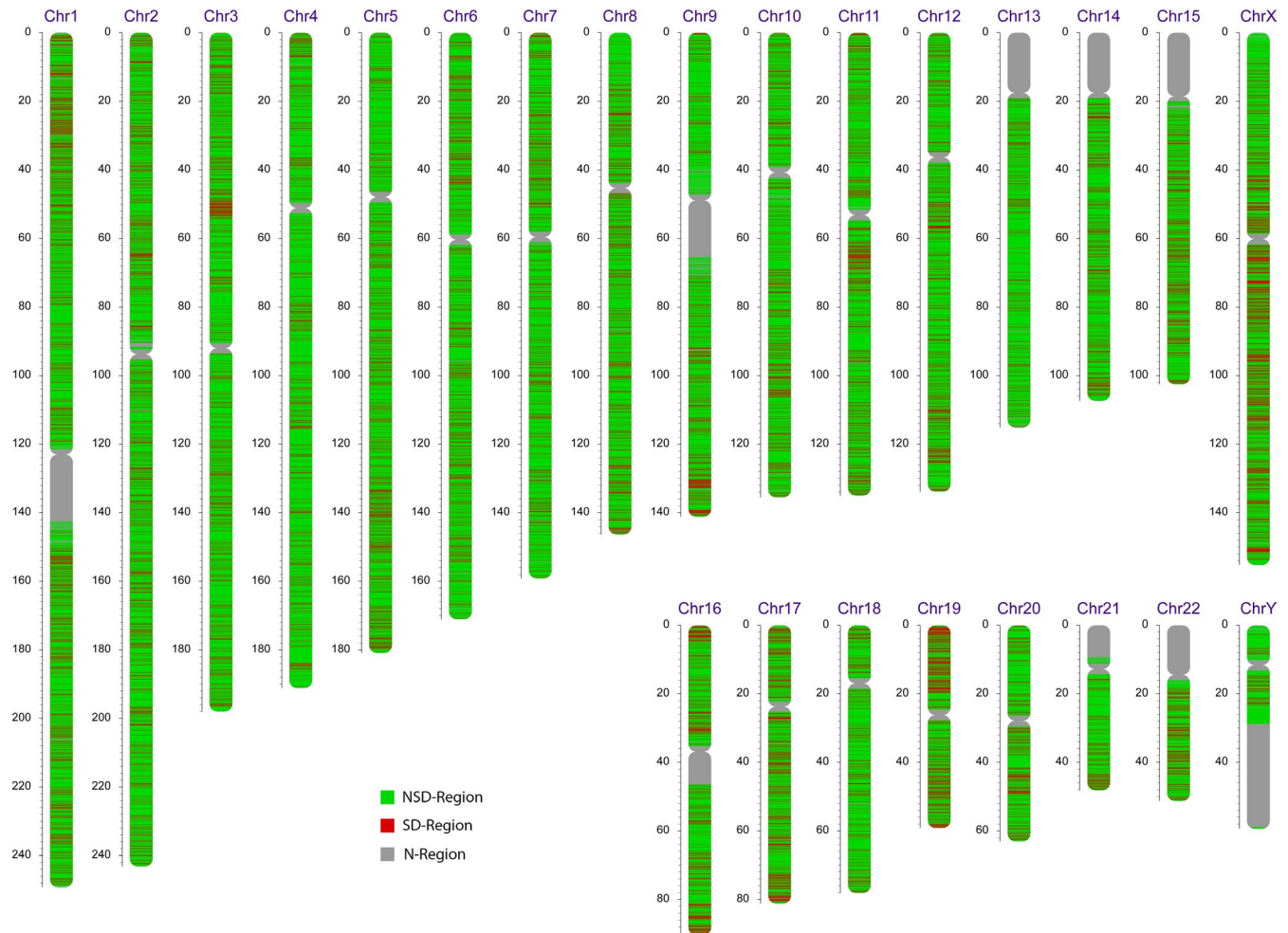
purposed to increase the proportion of fetal cfDNA, the observed difference should be largely attributed to elevated *ff*. This confirmed our hypothesis that there is a noticeable difference in the distribution of sequencing data between maternal and fetal cfDNA, and this difference is related to *ff*. In turn, these differential features should provide insight for *ff* estimation in the neural network.

### 3.2 | Estimation of *ff* using FF-QuantSC

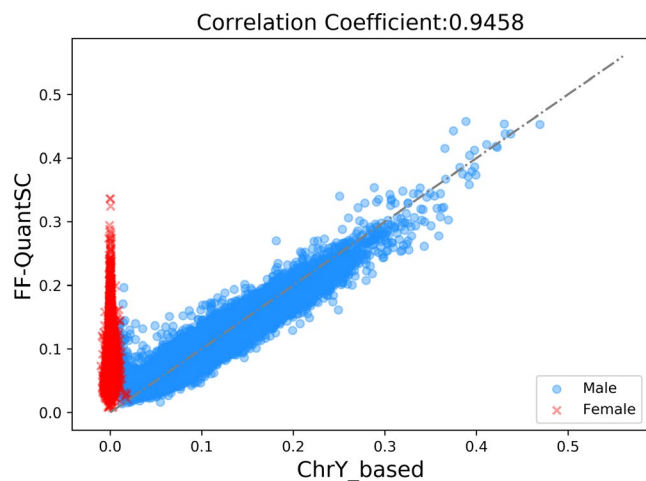
Following the confirmation of our two-step hypothesis, we constructed and tested the neural network model. Predicted *ff* from the neural network model was compared with that from the Y chromosomal-based method (Figure 4). As shown by male fetal samples (TS1\_M1 set), *ff* results from the two methods highly correlated with each other, with a Pearson's Correlation Coefficient (PCC) of 0.9458 ( $p < .0001$ ). We also compared the performance of FF-QuantSC with reported performance of some of the other previously described *ff* estimation techniques (Table 1). Among all compared methods, FF-QuantSC showed highest correlation with the Y chromosomal-based method. We thus concluded that FF-QuantSC was able to estimate *ff* of male pregnancies with high accuracy. As also shown in Figure 4, *ff* of female fetuses (TS1\_F set) estimated by FF-QuantSC exhibited similar distribution to that of male fetuses. This suggested that it may also be feasible to apply FF-QuantSC on female fetuses. To further demonstrate this feasibility, we compared distribution of predicted *ff* of various testing sets by FF-QuantSC (Figure 5). As expected, among all five testing sets with singleton pregnancies, regardless of the gender of the fetus, no statistically significant differences ( $p < .0001$ ) was detected in the distributions of estimated *ff*. This confirmed similar finding claimed by a previous study that exploited SNP-based method (Kim et al., 2015) and suggested that FF-QuantSC was robust and should provide equally accurate estimation of *ff* of both male and female fetuses. Finally, it is also worth noting that the twin pregnancies (TS\_TW set) showed an overall elevation of the value of estimated *ff* than singleton pregnancies.

### 3.3 | Comparison of FF-QuantSC and SeqFF on Female and Twin Pregnancies

Due to lack of gold standard for *ff* estimation nonmale fetal samples, to further evaluate the performance of FF-QuantSC, we compared FF-QuantSC with a reasonably accurate algorithm, SeqFF, on 40,000 female and 10,326 twin pregnancies. With a PCC of 0.9201 ( $p < .0001$ ) for female fetuses (Figure S2) and 0.8531 ( $p < .0001$ ) for twin



**FIGURE 3** Distributional difference between maternal and fetal genomes. Red and green labels indicate regions of difference and no difference, respectively. Grey labels stand for gaps in human genome

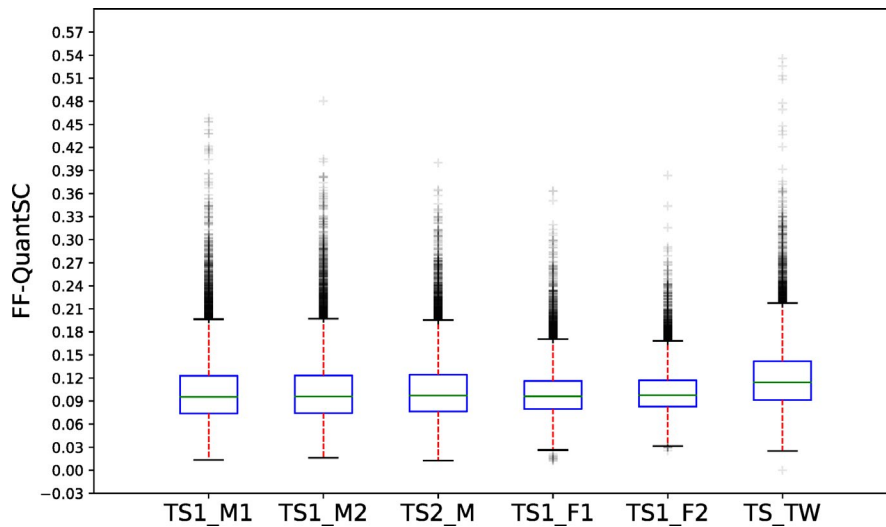


**FIGURE 4** Estimated *ff* of FF-QuantSC and Chromosome Y-based method. X-axis represents Chromosome Y-based *ff* estimation. Y-axis represents FF-QuantSC *ff* estimation. Samples with male fetuses (TS1\_M1 set) are plotted by blue round dots and samples with female fetuses (TS1\_F set) are plotted by red crosses. Dashed line stands for the expected correlation to be achieved

**TABLE 1** Comparison of FF-QuantSC with reported methods

Methods	PCC with Y chr-based method	P value
FF-QuantSC	0.945	<.0001
Size-based method (Yu et al., 2014)	0.827	<.0001
SNP-based method (Liao et al., 2011)	0.932	NA
Fetal Methylation Marker-Based method (Chan et al., 2006; Lun et al., 2013; Nygren et al., 2010)	0.85	<.001
Cell-Free DNA Nucleosome Track-Based method (Straver, Oudejans, Sistermans, & Reinders, 2016)	0.636	$1.61 \times 10^{-18}$
SeqFF <sup>†</sup> (Kim et al., 2015)	0.938	NA

<sup>†</sup>Note: SeqFF is a multivariate regression model and developed by Kim SK et al.



**FIGURE 5** Distribution of FF-QuantSC  $ff$  estimation of all testing sets. X-axis represents test sets. TS1\_M1, TS1\_M2, TS2\_M, TS1\_F1, and TS1\_F2 are sets of singleton pregnancies and TS\_TW contains twin pregnancies. Y-axis represents FF-QuantSC  $ff$  estimation. Green lines indicate the medians of predicted  $ff$ s

fetuses (Figure S3), the two algorithms showed overall high concordance.

## 4 | DISCUSSION

The predicted result of TS1\_M1 showed that the accuracy of FF-QuantSC for male pregnancies was better than the other compared methods. It is worth mentioning that this is under the premise that the Y chromosomal-based method was the gold standard for the estimation of  $ff$  of male fetuses. Although a correlation analysis between FF-QuantSC and SeqFF on non-male fetuses was performed, samples with discordant  $ff$  estimates did not necessarily imply poor prediction by FF-QuantSC as results from SeqFF may also be inaccurate. Distribution of the estimated  $ff$  by FF-QuantSC also indicated that at extreme  $ff$ , error rate of  $ff$  prediction was much higher (Figure 4). This was possibly due to the fact that there were few samples with extreme  $ff$  in the training set. Therefore, FF-QuantSC tends to learn the characteristics of samples with moderate  $ff$  during artificial neural network training. This provided potential direction for future model improvements. The predicted  $ff$  results of TS\_TW showed that the median of twins is about 1.2 times that of singletons, which may also provide valuable information for future work in estimation of twin  $ff$ .

FF-QuantSC requires no additional sequencing data or experimental steps to estimate  $ff$  in male and female pregnancies. As a result, it could be easily integrated to the routine NIPT analysis workflow without increasing the current NIPT cost. FF-QuantSC can also be handily adjusted for different platforms. Once FF-QuantSC is trained on a platform or sequencing mode, the model can be readily applied on any type of samples, such as pregnancies with male, female fetuses, and even potentially twins. Although high comprehensiveness and large size of the training sample are optimal in terms of model accuracy, the actual hardware support and training time should also be considered.

## 5 | CONCLUSION

In this study, by comparing the performance of FF-QuantSC with that of other  $ff$  estimating methods, we demonstrate that, regardless of the fetal gender, FF-QuantSC is capable of estimating  $ff$  with high accuracy. The fact that FF-QuantSC requires no additional experimental procedures on current NIPT routines enables inexpensive implementation of the model across various platforms. In summary, FF-QuantSC shall stand as a valuable tool to the accurate estimation of  $ff$  in NIPT.

### CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

### ACKNOWLEDGMENTS

The authors thank the members of BGI-Shenzhen genome sequencing team for help with sequencing.

### AUTHORSHIP

Wei Wang and Yuying Yuan designed research; Xianghua Chai participated in research design, performed research, provide methodology, and software and wrote the study; Bida Gu reviewed the work, wrote the study and participated in submission; Lijun Zhou and Qiang Liu performed research and analyzed data; Miaolan Cen, Zhongming Tian, and Weijiang Tang provided data; Fan Yang, Jingjuan Liu, Jiao Qiu, Jinjin Zhang, and Yumei Hou provided validation; Ya Gao and Shengting Li reviewed the study; Na Liu, Hongyun Zhang, Fang Chen, and Ye Yin provided conceptualization.

### DATA AVAILABILITY STATEMENT

Research data are not shared.

### ORCID

Xianghua Chai  <https://orcid.org/0000-0001-9476-6426>

## REFERENCES

- Bergstra, J. S., Yamins, D., & Cox, D. D. (2013). Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In *12th Python in science conference (SCIPY 2013)*. <https://doi.org/10.1088/1749-4699/8/1/014008>
- Canick, J. A., Palomaki, G. E., Kloza, E. M., Lambert-Messerlian, G. M., & Haddow, J. E. (2013). The impact of maternal plasma DNA fetal fraction on next generation sequencing tests for common fetal aneuploidies. *Prenatal Diagnosis*, *33*(7), 667–674. <https://doi.org/10.1002/pd.4126>
- Chan, K. C. A., Ding, C., Gerovassili, A., Yeung, S. W., Chiu, R. W. K., Leung, T. N., ... Lo, Y. M. D. (2006). Hypermethylated RASSF1A in maternal plasma: A universal fetal DNA marker that improves the reliability of noninvasive prenatal diagnosis. *Clinical Chemistry*, *52*(12), 2211–2218. <https://doi.org/10.1373/clinchem.2006.074997>
- Cherkassky, V. (1997). The Nature Of Statistical Learning Theory~. *IEEE Transactions on Neural Networks*, *8*(6), 1564–1564. <https://doi.org/10.1109/TNN.1997.641482>
- Goeman, J. J. (2010). L1 penalized estimation in the Cox proportional hazards model. *Biometrical Journal*, *52*(1):70–84. <https://doi.org/10.1002/bimj.200900028>
- Gregg, A. R., Skotko, B. G., Benkendorf, J. L., Monaghan, K. G., Bajaj, K., Best, R. G., ... Watson, M. S. (2016). Noninvasive prenatal screening for fetal aneuploidy, 2016 update: A position statement of the American College of Medical Genetics and Genomics. *Genetics in Medicine*, *18*(10), 1056–1065. <https://doi.org/10.1038/gim.2016.97>
- Kim, S. K., Hannum, G., Geis, J., Tynan, J., Hogg, G., Zhao, C., ... Deciu, C. (2015). Determination of fetal DNA fraction from the plasma of pregnant women using sequence read counts. *Prenatal Diagnosis*, *35*(8), 810–815. <https://doi.org/10.1002/pd.4615>
- Kinnings, S. L., Geis, J. A., Almasri, E., Wang, H., Guan, X., McCullough, R. M., ... Deciu, C. (2015). Factors affecting levels of circulating cell-free fetal DNA in maternal plasma and their implications for noninvasive prenatal testing. *Prenatal Diagnosis*, *35*(8), 816–822. <https://doi.org/10.1002/pd.4625>
- Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, *26*(5), 589–595. <https://doi.org/10.1093/bioinformatics/btp698>
- Liao, G. J. W., Lun, F. M. F., Zheng, Y. W. L., Chan, K. C. A., Leung, T. Y., Lau, T. K., ... Lo, Y. M. D. (2011). Targeted massively parallel sequencing of maternal plasma DNA permits efficient and unbiased detection of fetal alleles. *Clinical Chemistry*, *57*(1), 92–101. <https://doi.org/10.1373/clinchem.2010.154336>
- Lo, D., Sargent, I. L., Redman, C. W., Lo, Y. M. D., Corbetta, N., Chamberlain, P. F., ... Redman, C. W. G. (1997). Presence of Fetal DNA in Maternal Plasma and Serum Early report Presence of fetal DNA in maternal plasma and serum. *The Lancet*, *6736*(August 2016), 485–487. [https://doi.org/10.1016/S0140-6736\(97\)02174-0](https://doi.org/10.1016/S0140-6736(97)02174-0)
- Lo, Y. M. D., Chan, K. C. A., Sun, H., Chen, E. Z., Jiang, P., Lun, F. M. F., ... Chiu, R. W. K. (2010). Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. *Science Translational Medicine*, *2*(61), 61ra91–61ra91. <https://doi.org/10.1126/scitranslmed.3001720>
- Lun, F. M. F., Chiu, R. W. K., Chan, K. C. A., Tak, Y. L., Tze, K. L., & Lo, Y. M. D. (2008). Microfluidics digital PCR reveals a higher than expected fraction of fetal DNA in maternal plasma. *Clinical Chemistry*, *54*(10), 1664–1672.
- Lun, F. M. F., Chiu, R. W. K., Sun, K., Leung, T. Y., Jiang, P., Chan, K. C. A., ... Lo, Y. M. D. (2013). Noninvasive prenatal methylomic analysis by genomewide bisulfite sequencing of maternal plasma DNA. *Clinical Chemistry*, *59*(11), 1583–1594. <https://doi.org/10.1373/clinchem.2013.212274>
- Mercer, B. M. (2016). PB 163 Aneuploidy Screening, 1–15.
- Nygren, A. O. H., Dean, J., Jensen, T. J., Kruse, S., Kwong, W., Van Den Boom, D., & Ehrich, M. (2010). Quantification of fetal DNA by use of methylation-based DNA discrimination. *Clinical Chemistry*, *56*(10), 1627–1635. <https://doi.org/10.1373/clinchem.2010.146290>
- Sorenson, G. D., Pribish, D. M., Valone, F. H., & Memoli, V. A. (1994). Soluble Normal and Mutated DNA Sequences Genes in Human Blood” from. *Prevention*, *3*(February), 67–71.
- Sparks, A. B., Struble, C. A., Wang, E. T., Song, K., & Oliphant, A. (2012). Noninvasive prenatal detection and selective analysis of cell-free DNA obtained from maternal blood: Evaluation for trisomy 21 and trisomy 18. *American Journal of Obstetrics and Gynecology*, *206*(4), 319.e1–319.e9. <https://doi.org/10.1016/j.ajog.2012.01.030>
- Straver, R., Oudejans, C. B. M., Sistermans, E. A., & Reinders, M. J. T. (2016). Calculating the fetal fraction for noninvasive prenatal testing based on genome-wide nucleosome profiles. *Prenatal Diagnosis*, *36*(7), 614–621. <https://doi.org/10.1002/pd.4816>
- Wang, E., Batey, A., Struble, C., Musci, T., Song, K., & Oliphant, A. (2013). Gestational age and maternal weight effects on fetal cell-free DNA in maternal plasma. *Prenatal Diagnosis*, *33*(7), 662–666. <https://doi.org/10.1002/pd.4119>
- Watanagana, T., Peter, I., Messerlian, G. M., Borgatta, L., & Bianchi, D. W. (2004). Inverse correlation between maternal weight and second trimester circulating cell-free fetal DNA levels. *Obstetrics and Gynecology*, *104*(3), 545–550. <https://doi.org/10.1097/01.AOG.0000137352.93110.15>
- Yu, S. C. Y., Chan, K. C. A., Zheng, Y. W. L., Jiang, P., Liao, G. J. W., Sun, H., ... Lo, Y. M. D. (2014). Size-based molecular diagnostics using plasma DNA for noninvasive prenatal testing. *Proceedings of the National Academy of Sciences*, *111*(23), 8583–8588. <https://doi.org/10.1073/pnas.1406103111>
- Zhang, H., Gao, Y., Jiang, F., Fu, M., Yuan, Y., Guo, Y., ... Wang, W. (2015). Non-invasive prenatal testing for trisomies 21, 18 and 13: Clinical experience from 146 958 pregnancies. *Ultrasound in Obstetrics and Gynecology*, *45*(5), 530–538. <https://doi.org/10.1002/uog.14792>
- Zimmermann, B., Hill, M., Gemelos, G., Demko, Z., Banjevic, M., Baner, J., ... Rabinowitz, M. (2012). Noninvasive prenatal aneuploidy testing of chromosomes 13, 18, 21, X, and Y, using targeted sequencing of polymorphic loci. *Prenatal Diagnosis*, *32*(13), 1233–1241. <https://doi.org/10.1002/pd.3993>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

**How to cite this article:** Yuan Y, Chai X, Liu N, et al. FF-QuantSC: accurate quantification of fetal fraction by a neural network model. *Mol Genet Genomic Med*. 2020;8:e1232. <https://doi.org/10.1002/mgg3.1232>