High frequency of Exon 20 S768I *EGFR* mutation detected in malignant pleural effusions: a poor prognosticator of NSCLC?

George D'Souza[1,2*], Chirag Dhar[1,2,3*], Vishal Kyalanoor[3], Lokendra Yadav[3], Mugdha Sharmra[1,4], Mohammad Nawaz S[2], Sweta Srivastava[3]

[1]St. John's Research Institute, Bangalore, India

[2]Department of Pulmonary Medicine, St. John's Medical College and Hospital, Bangalore, India

[3]School of Medicine, University of California, San Diego, La Jolla, U.S.A.

[4]Department of Transfusion Medicine and Immunohaematology, St. John's Medical College and Hospital, Bangalore, India

[5]Department of General Medicine, St. John's Medical College and Hospital, Bangalore, India

* Equal contribution of authors

Corresponding authors: Dr. Sweta Srivastava- sweta.s@stjohns.in, Dr. George D'Souza- george.dsouza@stjohns.in and Dr. Chirag Dhar- cdhar@ucsd.edu/ chirag@worldviewofmedicine.org

Abstract

Lung cancer is the cause of a fourth of all cancer-related deaths. About a third of all lung adenocarcinoma tumours harbour mutations on exons 18, 19, 20 and 21 of the epidermal growth factor receptor (*EGFR*) gene. Detection of these mutations allows for targeted therapies in the form of EGFR Tyrosine kinase inhibitors. In our study, we utilized malignant pleural effusions (MPEs) as "liquid biopsies" to detect *EGFR* mutations when tissue biopsies were unavailable. We showed that a direct sequencing approach was likely to miss SNVs in MPEs. We then optimized an *EGFR* mutant-specific quantitative polymerase chain reaction-based assay and piloted it on n=10 pleural effusion samples (1 non-malignant pleural effusion as a negative control). 5/9 (55.55%) samples harboured EGFR mutations with 2/9 (22.22%) being exon 19 deletions and 3/9 (33.33%) had the S768I exon 20 mutation. The frequency of the S768I SNV in our study was significantly higher than that observed in other studies (~0.3%). Utilizing publicly available cBioPortal data, we report that patients with the S768I SNV had a shorter median survival time, progression-free survival time and lower tumor mutation count compared to patients with other *EGFR* mutations. These data suggest that this point mutation predicts poor prognosis as a result of aggressive disease, though studies in larger cohorts are necessary to confirm these findings. The high frequency of S768I mutations seen in our study also suggests that cancer cells harbouring these mutations may be superior in their ability to migrate, home or reside in pleural fluid.

Introduction

Lung cancer is a major contributor to death due to cancer. Amongst the various types of lung cancer, non-small cell lung cancer (NSCLC) accounts for approximately 80% of all

cases(1). A significant rise has been seen in the time trends of lung cancer in the Indian cities of Delhi, Chennai and Bangalore. Lung cancer is the cause of 6.9 percent of all new cancer cases and 9.3 percent of all cancer related deaths in both men and women in India (2). Mutations seen on Exons 18, 19, 20 and 21 of the Epidermal growth factor receptor (*EGFR*) gene are frequently observed in these cases and have been shown to be present in nearly a third of all lung adenocarcinoma cases (3)(4)(5). *EGFR* is a family member of receptor tyrosine kinases that play a central role in cellular signalling promoting cell growth and proliferation. Some mutations in this protein strongly predict the efficacy of *EGFR* tyrosine kinase inhibitors used in the treatment of these cases. There are reportedly more than 32 different mutations that have been detected in this gene distributed across exons 18, 19, 20, and 21 (6)(7). More than 20 different in-frame deletions on exon 19 have been reported accounting for nearly half of the cases of *EGFR* mutant lung adenocarcinoma. Importantly, tumours harbouring these deletions are found to be sensitive to EGFR TKIs such as erlotinib and gefitinib (8). This targeted therapy has been found to increase clinical outcomes in these patients (9). Response rates of more than 70 percent have been observed in patients on EGFR targeted therapy (10). Patients with EGFR mutant positive advanced NSCLC treated with Erlotinib, an EGFR Tyrosine kinase inhibitor (TKI) have shown better response rates and Progression-free survival (PFS) as compared to those on first line chemotherapeutic agents (11)(12)(13)(14).

Diagnostic tests for the detection of these EGFR mutations are now part of the routine management of lung adenocarcinoma(15). In spite of the invasiveness and morbidity associated with lung biopsies(16), the primary tumour is preferred for detection of mutations(17). Most often, an amplification-refractory mutation system (ARMS) is used for the detection of these mutations from the biopsy. The utility of malignant pleural effusion (MPE) in the detection of

EGFR mutations has been well established over the years in the field of lung cancer diagnosis(18). In this study, we optimized a method to detect EGFR mutations in MPEs at our tertiary care centre. Notably, the single nucleotide variant (SNV) S768I was found in 3 out of the 9 pleural effusions (33.33%). We then analysed the prevalence and clinical features of patients harboring this SNV by utilizing cBioPortal (www.cbioportal.org).

Results

Dilutions of mutant in wildtype EGFR DNA predict poor sensitivity of direct sequencing-based detection of SNVs in MPEs.

First, a direct sequencing-based approach was optimized to detect EGFR mutations from MPEs. A common exon 19 deletion was detected using this method (figure 1 a and b). We then asked if this method would be suitable to pick up mutant alleles in the backdrop of wildtype inflammatory cells present in MPEs. To answer this question, we serially diluted DNA with the exon 19 deletion with wildtype DNA and subjected the mixture to direct sequencing. While signatures of the large 18bp deletion were visible even in 1:100 dilutions, the results suggested poor base-call confidence (figure 1 c). This result suggested that the direct sequencing method was unlikely to pick up SNVs.

Figure 1: Poor base-call confidence in diluted DNA samples. a) Common exon 19 deletion detected. Alignment of MPE sequence with reference

sequence shows an 18bp deletion (highlighted in the yellow box). **b) Chromatogram view of the same mutation** (highlighted in the yellow box) **c) Alignment of diluted DNA sample.** Poor base-call confidence in sequence of 1:100 dilution of mutant DNA in wildtype DNA (healthy laboratory volunteer).

High frequency of S768I exon 20 mutation detected by mutant-specific quantitative PCR of MPE.

Subsequently, an established *EGFR* mutation real-time PCR kit was used to identify EGFR mutations in malignant pleural effusions. While this kit is optimized to detect mutations in genomic DNA obtained from primary tumors, we attempted to utilize MPEs instead. We demonstrated for the first time that this kit could detect EGFR mutations from MPEs. Our results showed that 5/9 (55.55%) of pleural effusions probed had EGFR mutations. Of these, 2/9 (22.22%) were the exon 19 deletion E746_S750del18 and 3/9 (33.33%) were the exon 20 S768I mutations (Figure 2 a and b). The amplification plots and ΔCt values are provided in supplementary figures 1-7 and supplementary tables 1 and 2. The occurrence of this SNV and clinical features of these patients was then studied using publicly available data on cBioPortal. This SNV has been detected only 17 times in the nearly 5836 (0.3%) non-small cell lung cancer samples surveyed for *EGFR* mutations (Figure 2 c). These detections include duplicates and the frequency at the patient level is 0.18% (10/5490 patients). This SNV is mapped onto the tyrosine-kinase domain of *EGFR* and is a known cancer hotspot (statistically significantly recurrent mutations identified from large scale cancer genomic studies) (Figure 2 d). Patients with this mutation are suitable candidates for the FDA-approved targeted therapy Afatanib.

a

Sample. No.	Sex	Mutation detected	Exon
S1	M	WT	-
S2	M	E746_S750DEL	Exon 19
S3	M	WT	-
S4	F	E746_S750DEL	Exon 19
S5	F	WT	-
S6	M	WT	-
S7	M	S768I	Exon 20
S8	M	WT	-
S9	M	S768I	Exon 20
S10	M	S768I	Exon 20

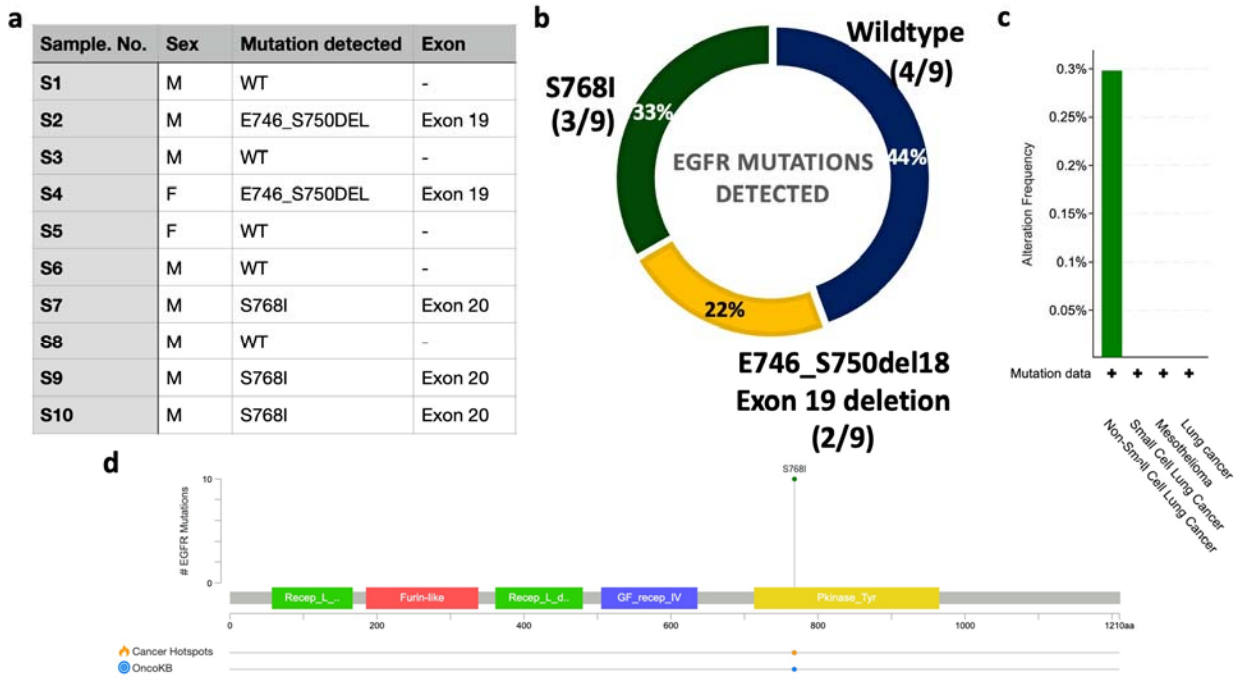
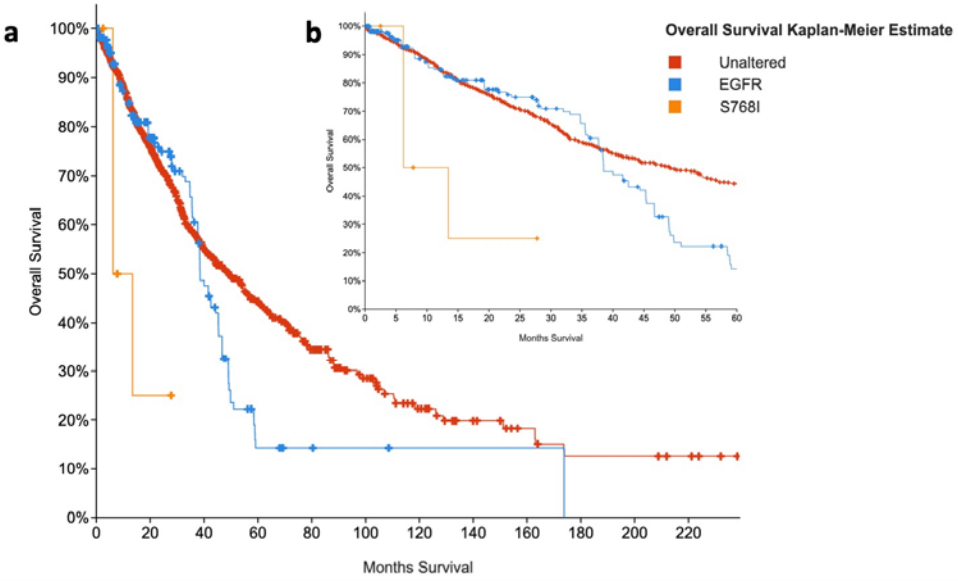


Figure 2: High frequency of S768I detected from MPEs. a) Table depicting gender, mutation detected and exon of occurrence. M= Male, F= Female, WT= wildtype *EGFR*, S5 was a non-malignant pleural fluid sample that served as an internal assay negative control **b) Doughnut chart representing the frequency of various mutations detected in this study.** Percentages are embedded in the slices and number of cases observed are indicated in parenthesis beneath the genotype. **c) Frequency of S768I in other studies.** cBioPortal data shows that this SNV has been detected only 17 times at a frequency of 0.4%. **d) Mapping of S768I.** S768I maps to the tyrosine kinase domain of *EGFR* and is a known cancer hotspot.

Patients having the S768I mutation have a shorter median survival time and a shorter progression-free survival time compared to patients having other *EGFR* mutations.

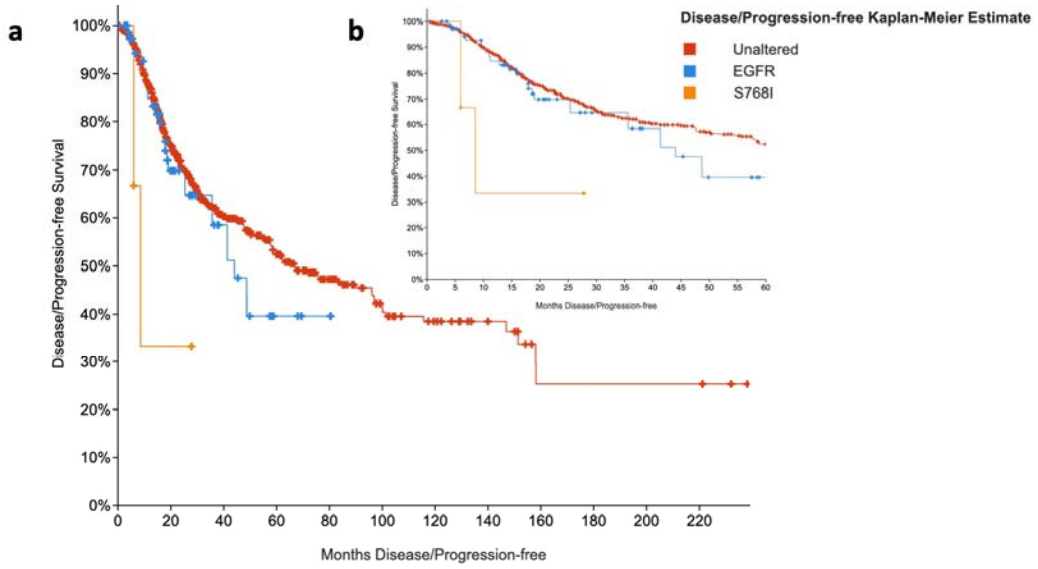
From publicly available cBioPortal data, patients with the S768I mutation have a median survival of 6.20 months compared to 38.40 months patients with other *EGFR* mutations (figure 3 a, b and c). Additionally, patients with the SNV had a shorter progression-free survival times (8.55 months for patients with S768I vs 44.02 months for patients with other *EGFR* mutations, figures 4 a, b and c). These survival statistics (limited by number of patients) suggest that patients with this SNV present with an advanced form of NSCLC.



c

	Number of Cases, Total	Number of Cases, Deceased	Median Months Survival
Unaltered	**2984**	**1081**	**49.25**
EGFR	**238**	**93**	**38.50**
S768I	**9**	**5**	**6.21**

Figure 3: Patients having the S768I mutation have a shorter median survival time compared to patients having other *EGFR* mutations. a) Kaplan-Meir (KM) survival curve b) KM curve for first 60 months inset c) Median survival in months suggest shorter median survival in patients with S768I. Unaltered refers to NSCLC patients without *EGFR* mutations, EGFR to patients with *EGFR* mutations except S768I and S768I refers to patients with this SNV.



c

	Number of Cases, Total	Number of Cases, Relapsed/Progressed	Median Months Disease-free
Unaltered	**1331**	**442**	**66.84**
EGFR	**76**	**27**	**44.02**
S768I	**3**	**2**	**8.55**

Figure 4: Patients having the S768I mutation have a shorter progression-free survival time compared to patients having other *EGFR* mutations. a) Kaplan-Meir (KM) survival curve b) KM curve for first 60 months inset c) Median survival in months suggest shorter median time to relapse in patients with S768I. Unaltered refers to NSCLC patients without *EGFR* mutations, EGFR to patients with *EGFR* mutations except S768I and S768I refers to patients with this SNV.

Presence of the S768I SNV was not associated with a significant difference in stage at diagnosis, lymph node involvement, size of tumor and metastasis.

Given the poorer survival in the S768I group, we compared the stage at diagnosis, lymph nodal involvement, size of tumor and metastasis between the groups. There was no difference in the stage at diagnosis between patients with or without the S768I mutation (figure 5 a). There also was no discernible difference in lymph nodal involvement (figure 5 b) or tumor size (figure 5 c)

between the two groups possibly due to the small sample size in the S768I group. We then queried the number metastasis samples obtained from these patients as a proxy for frequency of metastasis. Again, there was no difference in the frequency between the groups (figure 5 d). It is likely that a larger meta-analysis of this SNV may reveal discernible differences in these factors but in a limited number of cases with available data this was not the case.

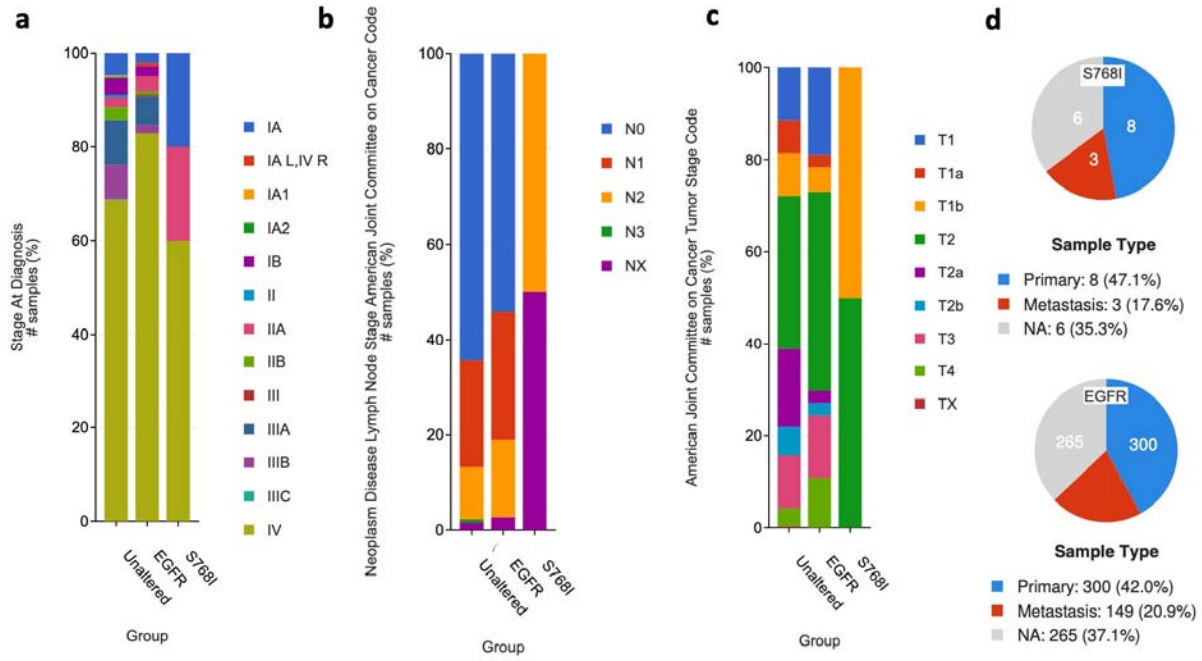


Figure 5: No observable differences in stage at diagnosis, lymph nodal involvement, tumor stage code or metastasis in tumors with S768I. Differences between groups by a) Stage at diagnosis b) Lymph nodal involvement c) Tumor stage code and d) metastatic lesions sampled as a proxy for metastatic frequency. Unaltered refers to NSCLC patients without *EGFR* mutations, EGFR to patients with *EGFR* mutations except S768I and S768I refers to patients with this SNV.

Tumors with the S768I SNV have a lower mutation burden but co-occur with mutations on other genes.

We then asked if patients with the S768I patients had a lower mutation count compared to patients with other *EGFR* mutations. Indeed, the number of mutations was lower in the S768I group (figure 6 a) suggesting a more deleterious effect of this SNV. These data when taken together with the survival data point towards a likely pathogenic role of S768I in advanced NSCLC. Further analysis revealed association of the S768I SNV with other mutations. This SNV was as likely to co-occur with TP53 mutations as tumors with no EGFR mutations or ones with other EGFR mutations. Yet, S768I co-occurred with mutations on other genes such as MUC17, BMPR2, DOCK3, TPTE, ATRX and FSHR (figure 6 b, c and d). The clinical significance or mechanistic importance of these co-occurring mutations is unknown.

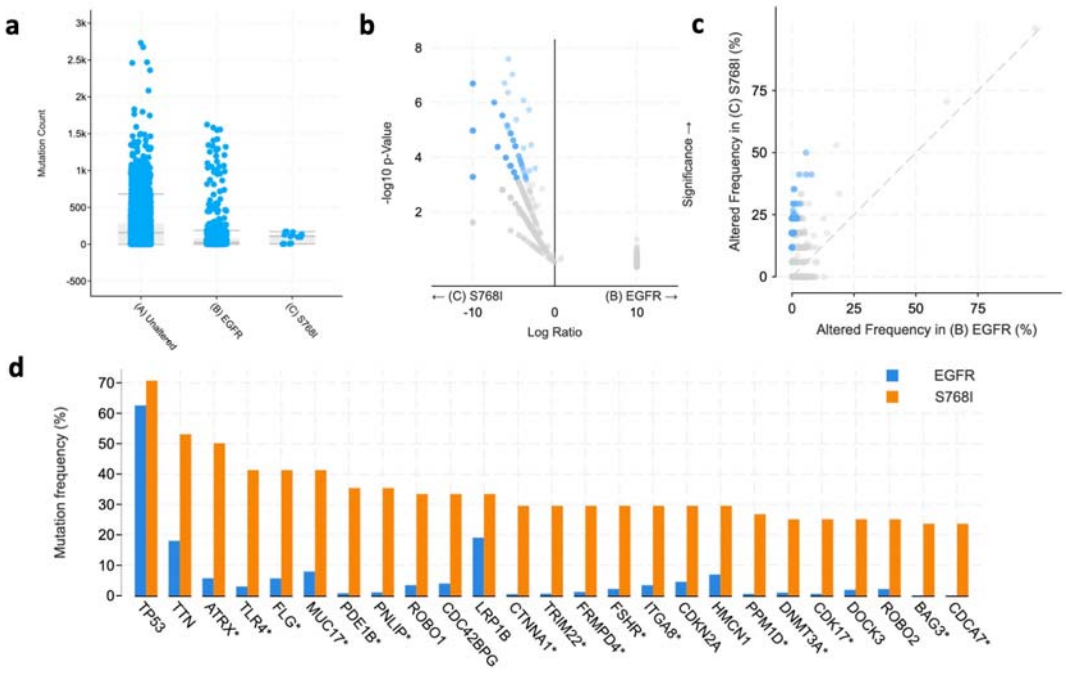


Figure 6: Lower tumor mutation in patients with S768I and co-occurrence with other mutations. a) Mutation load in tumors with S768I. The overall mutation load in tumors with the S768I SNV is lower than in tumors with other EGFR mutations or NSCLC samples with

other mutations. **b) Graphical representation of genes mutated more frequently in tumors positive for S768I but not in tumors with other EGFR mutations. c) Dot plot depicting frequency of co-occuring mutations in S768I tumors.** Blue dots represent genes with a statistically significant p-value. **d) Bar graphs of 25 most frequently detected mutations (apart from EGFR).** TP53 mutation frequency is nearly the same but mutations in genes such as FLG, MUC17, BMPR2, DOCK3, TPTE, ATRX and FSHR mutations appear more frequently in conjunction with the S768I mutation. Asterisked (*) genes represent those with statically significant differences. A list of all 175 genes with statistically significant co-occurrence with S768I are listed in the supplementary information.

Discussion

In this study, we showed that a direct sequencing approach may miss *EGFR* SNVs in MPEs and therefore utilized a Taqman-based mutant-specific quantitative PCR assay. We optimized this assay and detected both large deletions and SNVs in *EGFR*. Notably, a high frequency of the S768I mutation occurring in the absence of other *EGFR* mutations was seen. We then compared this frequency to publicly available data on cBioPortal ® and observed a significant difference in frequency (33.33% in our study vs 0.4% in reported literature). To further understand the role of this SNV, a cBioPortal ® search and comparison strategy was devised to compare the clinical features of patients with the SNV, patients with other *EGFR* mutations and patients with no mutaions in *EGFR*. Patients with the S768I had an 84% reduction in estimated median survival time (and progression-free survival time) compared to those with other *EGFR* mutations. Tumors bearing the S768I mutation also had a lower mutational than the other *EGFR* mutations group. These data when taken together suggests that patients with the

S768I mutation have poorer clinical outcomes and this mutation is likely to promote an aggressive tumor phenotype.

Very little is known about the implications of having the S768I mutation in the absence of other *EGFR* mutations. OncoKB suggests the use of afatinib as the FDA-approved therapy for patients with this SNV. One other study shown that S768I conferred reduced sensitivity to gefitinib *in vitro* as compared to other mutations (19)(20). A study by Leventakos and group reported that the S7681 is rare and is usually found in combination with sensitizing mutations and they have concluded that the predictive and prognostic role of this mutation is yet to be fully explored (21). In our study we have found that 33.3% of the samples exhibited this mutation in isolation, with an absence of any other mutation. It did not escape our attention that all reports of this SNV in studies discussed here occur in the presence of other EGFR mutations (such as at the G719 or T790 position). Some of the reasons for this mutation being detected in the absence of other *EGFR* mutations in our study are discussed here. Most studies utilize the primary tissue and a sub-population with this isolated mutation may appear later in the pathogenesis of the disease. One may also simply come to the conclusion that our cohort has a larger number of patients with this mutation, but we think otherwise. We hypothesize that S768I is a secondary mutation that develops later in the course of the disease and cells harbouring this mutation are able to migrate to, home in or reside within the pleural cavity more easily that cells without this mutation. One could further this speculation that cells having this mutation have a greater metastatic potential (figure 7) though further studies will be needed to prove these hypotheses. Additionally, *in vitro* and *in vivo* studies will determine if this mutation is likely to respond to EGFR TKIs such as afatinib.

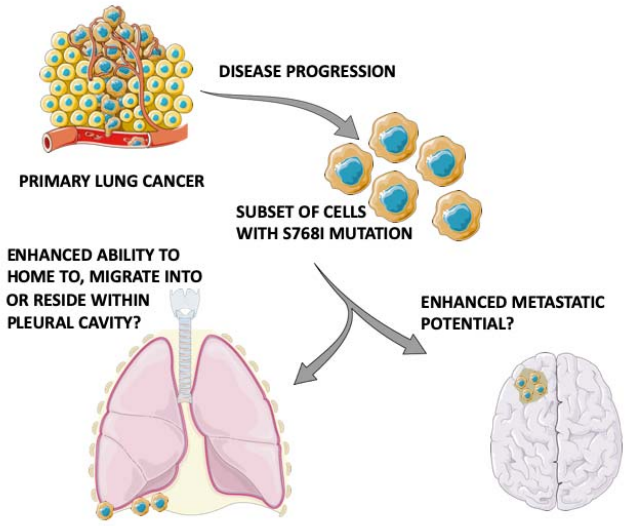


Figure 7: Hypothesized role of the S768I mutation in the pathogenesis of metastasis in lung cancer (Elements of images sourced from http://smart.servier.com/).

The prevalence of somatic mutations in the *EGFR* gene is high (~ 33%) in lung adenocarcinomas(22)(23). The detection of any *EGFR* mutation and the resultant treatment with targeted therapy with EGFR TKIs (Tyrosine Kinase Inhibitors) aid in sensitizing these tumours. *EGFR* mutation testing on tissue biopsies is the gold standard for diagnosis. However, biopsy procedures may result in uncontrolled bleeding and morbidity(24). Utility of pleural effusion fluid for mutation testing along with other secondary specimens (pericardial effusion, fine needle aspirates, CSF, cfDNA, etc) have been used as alternate strategies for reliable testing (25). In our study we first used PCR coupled to direct sequencing to detect mutation in pleural effusion derived malignant cells. We were able to detect a common 18bp deletion mutation using this method. One of the biggest drawbacks of direct sequencing versus allele specific detection lies in its limit of detection. Pleural fluid consists of a heterogeneous pool of stromal and inflammatory cells apart from the metastatic cells. Since the non-tumor cells harbour wildtype genomic sequence, the signal to noise ratio in electropherogram is lower for mutated sequence. The

mutation sequence identification may be compromised in the event of suboptimal malignant cells numbers. We showed that at dilution of 1:100 of a mutated DNA with wildtype genomic DNA signatures of the mutation are visible but the wildtype sequence bases provide conflicting results in certain positions of the gene. These inadequate results are more likely to be seen in single base pair mutations. We then utilized an *EGFR* mutation detection kit. This diagnostic kit is designed to test for *EGFR* mutation on FFPE sections but we were able to optimize its use on MPEs. To the best of our knowledge, this is the first report of MPE *EGFR* mutation testing in India utilizing the Taqman technology platform. Jong Sik Lee and colleagues have used PNA clamping technology to detect EGFR mutations in the supernatant derived from the pleural effusion where in the objective was to detect mutations using cell free DNA (26). Jie Lin et al have shown that EGFR mutations can be detected in the supernatant, cell pellets of the pleural effusion using High Resolution Melting (HRM) analysis and Sanger sequencing (27). When implemented on a small pilot scale study (n=9), we were successful in detecting EGFR mutations in 5/9 (55%) subjects. The mutations that were detected in these samples were either S768I or E746_S750del.

The need to develop minimally invasive diagnostic tests in the management of cancers is gaining momentum. However, our study is limited in its impact as it evaluates only 11 patients and further study needs to be undertaken to evaluate and validate this method in a larger cohort. Additionally, the detection of circulating tumour DNA should also be explored as an alternative to repeated biopsies is one of the techniques being studied (28). This detection is being done using highly sensitive techniques such as droplet digital polymerase chain reaction. Resources

must be directed towards the development of such minimally/non-invasive techniques, especially if these are inexpensive as well.

Ethics

Ethical clearance was obtained from the Institutional Ethics Committee of St. John's National Academy of Health as per IERB guidelines (Study No. 1/2016). Appropriate consents for carrying out experimental genetic tests were obtained.

Methods

Pleural fluid processing and DNA isolation

Pleural fluid was tapped from patients with malignant pleural effusions and centrifuged at 3500 rpm for 5 mins at 4°C to collect the cell pellet. The cell pellets were stored in -80°C until further processing. The DNA was extracted from the samples using the QIAamp DNA Mini Kit (Qiagen) as per the manufacturer's instructions. Samples were analyzed for their quality using BioSpec Nano UV-VIS Spectrophotometer (Shimadzu Corp.) and electrophoresed on an agarose gel to ascertain integrity.

Direct sequencing

Appropriate primers for exons 18, 19, 20 and 21 of the EGFR gene were designed and obtained based on the reference sequence available on NCBI (Supplementary table 2). Genomic DNA obtained was subjected to a gradient PCR to optimize the annealing temperature and was

visualized on a 1.2% agarose gel (Supplementary figure 8). The PCR products were subjected to direct sequencing.

Real-time PCR based detection of EGFR mutations

DNA samples were subjected to an assay to a detect 10 *EGFR* mutations in a 96-well Real Time PCR format using the TRUPCR® EGFR Kit (3B BlackBio Biotech India Ltd.) as per the manufacturer's instructions. Real-time PCR data analysis and identification of EGFR mutant samples was done. To the best of our knowledge, this is the first time this kit is being using to detect mutations from MPEs.

cBioPortal search strategy

A search strategy was used to access publicly available cancer genomic/clinical data on cBioPortal(29)(30) utilizing multiple studies(31)(32)(33)(34)(35)(36)(37)(38)(39)(40)(41)(4)(42) with the gene alias “EGFR: MUT = S768I”, “EGFR” and to access all NSCLC studies. Venn-diagram based sorting was utilized to isolate cases with the S768I mutation and those with other *EGFR* mutations. “Unaltered” refers to patients with mutations other than EGFR, “EGFR” refers to patients with *EGFR* mutations other than S768I and “S768I” refers to patients with the S768I *EGFR* mutation. OncoKB(43) was used to identify any FDA-approved drug for this mutation.

Acknowledgements

This work is supported by a grant from the Advanced Research Wing of the Rajiv Gandhi University of Health Sciences, Bangalore. C.D. was awarded a C.R.E.S.T MAS scholarship by CTRI at the School of Medicine, University of California, San Diego. We thank Prof. Sudhir

Krishna, Ms. Pranatharthi Annapurna and the sequencing division of NCBS, Bangalore for their support. Some of these data were presented at the National Lung Cancer Conference 2016 in Bhubaneshwar and at the American Society for Cellular Biology-European Molecular Biology Organization joint conference 2018 at San Diego(44)(45)(46) . A special mention to Dr. Paul Kalanithi for his battle against EGFR mutant lung cancer. His memoir When Breath becomes Air is humbling and yet, inspiring.

Conflict of Interest

The authors declare no conflict of interest.

Author contributions

G.D.S and C.D contributed equally this work which was conceived, planned, supervised and interpreted by G.D.S, C.D and S.S. Funding acquisition was by G.D.S. Experimental assays were performed and interpreted by C.D., V.K., L.Y. and M.S. Clinical sample acquisition was by C.D., M.N.S. and G.D.S. C.D. performed cBioPortal/bioinformatic analysis. Manuscript was written and edited by C.D. and V.K. S.S. edited the manuscript. All authors approved the final manuscript.

References

1. Detection and comparison of epidermal growth factor receptor mutations in cells and fluid of malignant pleural effusion in non-small cell lung cancer.

2. Malik PS, Raina V. Lung cancer: prevalent trends & emerging concepts. Indian J Med Res. 2015 Jan;141(1):5–7.

3. Doval D, Prabhash K, Patil S, Chaturvedi H, Goswami C, Vaid A, et al. Clinical and epidemiological study of EGFR mutations and EML4-ALK fusion genes among Indian patients with adenocarcinoma of the lung. OncoTargets Ther. 2015;8:117–23.

4. Ding L, Getz G, Wheeler DA, Mardis ER, McLellan MD, Cibulskis K, et al. Somatic mutations affect key pathways in lung adenocarcinoma. Nature. 2008 Oct 23;455(7216):1069–75.

5. Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, et al. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A. 2004 Sep 7;101(36):13306–11.

6. Sharma SV, Bell DW, Settleman J, Haber DA. Epidermal growth factor receptor mutations in lung cancer. Nat Rev Cancer. 2007 Mar;7(3):169–81.

7. Cosmic. COSMIC - Catalogue of Somatic Mutations in Cancer [Internet]. [cited 2020 Mar 6]. Available from: https://cancer.sanger.ac.uk/cosmic

8. Cheng L, Alexander RE, Maclennan GT, Cummings OW, Montironi R, Lopez-Beltran A, et al. Molecular pathology of lung cancer: key to personalized medicine. Mod Pathol Off J U S Can Acad Pathol Inc. 2012 Mar;25(3):347–69.

9. Rocha-Lima CM, Raez LE. Erlotinib (tarceva) for the treatment of non-small-cell lung cancer and pancreatic cancer. P T Peer-Rev J Formul Manag. 2009 Oct;34(10):554–64.

10. Mok TS, Wu Y-L, Thongprasert S, Yang C-H, Chu D-T, Saijo N, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009 Sep 3;361(10):947–57.

11. Zhou C, Wu Y-L, Chen G, Feng J, Liu X-Q, Wang C, et al. Erlotinib versus chemotherapy as first-line treatment for patients with advanced EGFR mutation-positive non-small-cell lung cancer (OPTIMAL, CTONG-0802): a multicentre, open-label, randomised, phase 3 study. Lancet Oncol. 2011 Aug;12(8):735–42.

12. Rosell R, Carcereny E, Gervais R, Vergnenegre A, Massuti B, Felip E, et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. Lancet Oncol. 2012 Mar;13(3):239–46.

13. Sequist LV, Martins RG, Spigel D, Grunberg SM, Spira A, Jänne PA, et al. First-line gefitinib in patients with advanced non-small-cell lung cancer harboring somatic EGFR mutations. J Clin Oncol Off J Am Soc Clin Oncol. 2008 May 20;26(15):2442–9.

14. Shigematsu H, Lin L, Takahashi T, Nomura M, Suzuki M, Wistuba II, et al. Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst. 2005 Mar 2;97(5):339–46.

15. Peters S, Adjei AA, Gridelli C, Reck M, Kerr K, Felip E, et al. Metastatic non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol Off J Eur Soc Med Oncol. 2012 Oct;23 Suppl 7:vii56-64.

16. Wiener RS, Schwartz LM, Woloshin S, Welch HG. Population-based risk for complications after transthoracic needle lung biopsy of a pulmonary nodule: an analysis of discharge records. Ann Intern Med. 2011 Aug 2;155(3):137–44.

17. Pirker R, Herth FJF, Kerr KM, Filipits M, Taron M, Gandara D, et al. Consensus for EGFR mutation testing in non-small cell lung cancer: results from a European workshop. J Thorac Oncol Off Publ Int Assoc Study Lung Cancer. 2010 Oct;5(10):1706–13.

18. Liu D, Lu Y, Hu Z, Wu N, Nie X, Xia Y, et al. Malignant Pleural Effusion Supernatants Are Substitutes for Metastatic Pleural Tumor Tissues in EGFR Mutation Test in Patients with Advanced Lung Adenocarcinoma. PLoS ONE [Internet]. 2014 Feb 28 [cited 2020 Apr 3];9(2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3938554/

19. Kancha RK, von Bubnoff N, Peschel C, Duyster J. Functional analysis of epidermal growth factor receptor (EGFR) mutations and potential implications for EGFR targeted therapy. Clin Cancer Res Off J Am Assoc Cancer Res. 2009 Jan 15;15(2):460–7.

20. Chen Y-R, Fu Y-N, Lin C-H, Yang S-T, Hu S-F, Chen Y-T, et al. Distinctive activation patterns in constitutively active and gefitinib-sensitive EGFR mutants. Oncogene. 2006 Feb 23;25(8):1205–15.

21. Leventakos K, Kipp BR, Rumilla KM, Winters JL, Yi ES, Mansfield AS. S768I Mutation in EGFR in Patients with Lung Cancer. J Thorac Oncol Off Publ Int Assoc Study Lung Cancer. 2016;11(10):1798–801.

22. Zhang Y-L, Yuan J-Q, Wang K-F, Fu X-H, Han X-R, Threapleton D, et al. The prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis. Oncotarget. 2016 Oct 12;7(48):78985–93.

23. Han B, Tjulandin S, Hagiwara K, Normanno N, Wulandari L, Laktionov K, et al. EGFR mutation prevalence in Asia-Pacific and Russian patients with advanced NSCLC of adenocarcinoma and non-adenocarcinoma histology: The IGNITE study. Lung Cancer Amst Neth. 2017;113:37–44.

24. Lim M, Kim C-J, Sunkara V, Kim M-H, Cho Y-K. Liquid Biopsy in Lung Cancer: Clinical Applications of Circulating Biomarkers (CTCs and ctDNA). Micromachines. 2018 Feb 28;9(3).

25. Ellison G, Zhu G, Moulis A, Dearden S, Speake G, McCormack R. EGFR mutation testing in lung cancer: a review of available methods and their use for analysis of tumour tissue and cytology samples. J Clin Pathol. 2013 Feb;66(2):79–89.

26. Lee JS, Hur JY, Kim IA, Kim HJ, Choi CM, Lee JC, et al. Liquid biopsy using the supernatant of a pleural effusion for EGFR genotyping in pulmonary adenocarcinoma patients: a comparison between cell-free DNA and extracellular vesicle-derived DNA. BMC Cancer. 2018 Dec 10;18(1):1236.

27. Lin J, Gu Y, Du R, Deng M, Lu Y, Ding Y. Detection of EGFR mutation in supernatant, cell pellets of pleural effusion and tumor tissues from non-small cell lung cancer patients by high resolution melting analysis and sequencing. Int J Clin Exp Pathol. 2014;7(12):8813–22.

28. Suraj S, Dhar C, Srivastava S. Circulating nucleic acids: An analysis of their occurrence in malignancies. Biomed Rep. 2017 Jan;6(1):8–14.

29. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012 May;2(5):401–4.

30. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013 Apr 2;6(269):pl1.

31. Thoracic PDX (MSK, provisional) [Internet]. Available from: cBioPortal.org

32. Yaeger R, Chatila WK, Lipsyc MD, Hechtman JF, Cercek A, Sanchez-Vega F, et al. Clinical Sequencing Defines the Genomic Landscape of Metastatic Colorectal Cancer. Cancer Cell. 2018 08;33(1):125-136.e3.

33. Rizvi H, Sanchez-Vega F, La K, Chatila W, Jonsson P, Halpenny D, et al. Molecular Determinants of Response to Anti-Programmed Cell Death (PD)-1 and Anti-Programmed Death-Ligand 1 (PD-L1) Blockade in Patients With Non-Small-Cell Lung Cancer Profiled With Targeted Next-Generation Sequencing. J Clin Oncol Off J Am Soc Clin Oncol. 2018 01;36(7):633–41.

34. Jamal-Hanjani M, Wilson GA, McGranahan N, Birkbak NJ, Watkins TBK, Veeriah S, et al. Tracking the Evolution of Non-Small-Cell Lung Cancer. N Engl J Med. 2017 01;376(22):2109–21.

35. Gobbini E, Galetta D, Tiseo M, Graziano P, Rossi A, Bria E, et al. Molecular profiling in Italian patients with advanced non-small-cell lung cancer: An observational prospective study. Lung Cancer Amst Neth. 2017;111:30–7.

36. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science. 2015 Apr 3;348(6230):124–8.

37. Campbell JD, Alexandrov A, Kim J, Wala J, Berger AH, Pedamallu CS, et al. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. Nat Genet. 2016;48(6):607–16.

38. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell. 2012 Sep 14;150(6):1107–20.

39. Lung Adenocarcinoma, Lung Squamous Cell Carcinoma (TCGA, Firehose Legacy) [Internet]. Available from: cBioPortal.org

40. Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. Nature. 2014 Jul 31;511(7511):543–50.

41. TCGA PanCancer Atlas [Internet]. TCGA PanCancer Atlas. Available from: https://www.cell.com/pbassets/consortium/pancanceratlas/pancani3/index.html

42. Jordan EJ, Kim HR, Arcila ME, Barron D, Chakravarty D, Gao J, et al. Prospective Comprehensive Molecular Characterization of Lung Adenocarcinomas for Efficient Patient Matching to Approved and Emerging Therapies. Cancer Discov. 2017;7(6):596–609.

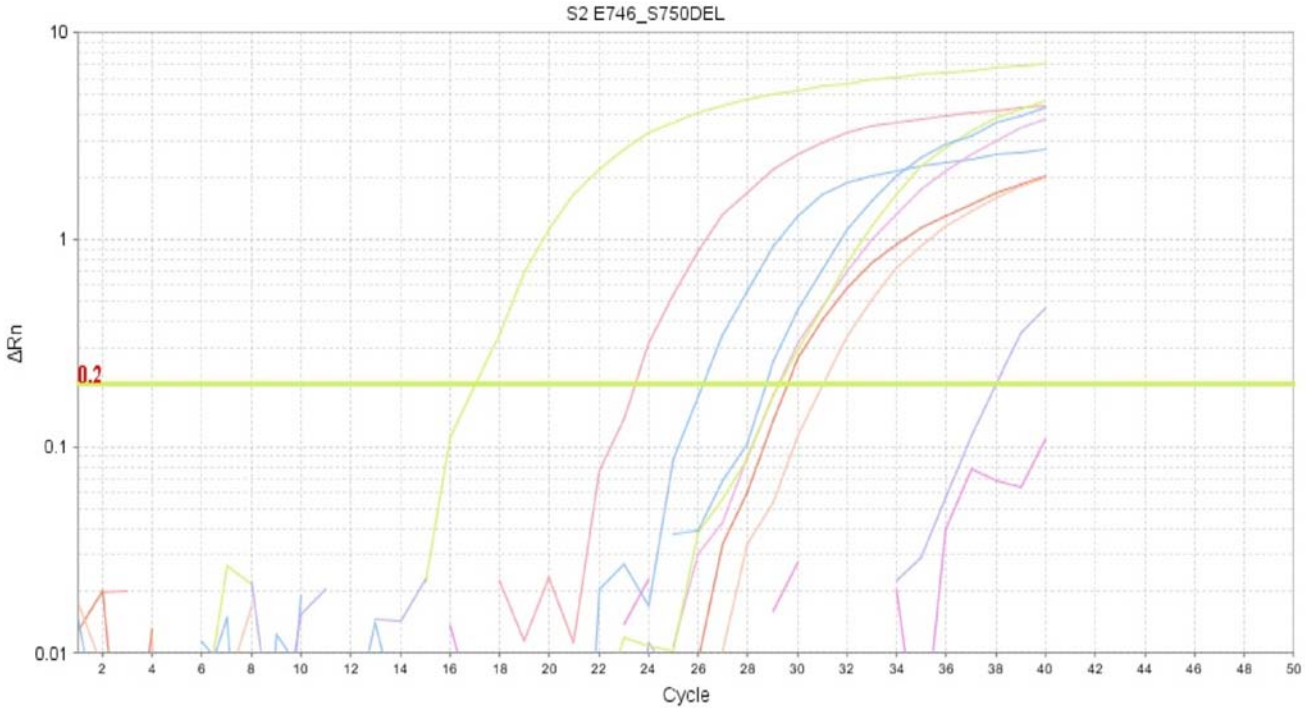
43. Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, et al. OncoKB: A Precision Oncology Knowledge Base. JCO Precis Oncol [Internet]. 2017 Jul [cited 2020 Apr 3];2017. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5586540/

44. Dhar C, Sharma M, Nawaz M, D'Souza G, Srivastava S. Sanger sequencing of metastatic pleural effusion: An inexpensive, minimally invasive tool for the detection of EGFR mutations in lung adenocarcinoma- Presented at the National Lung Cancer Conference, Bhubaneshwar 2016 (https://www.researchgate.net/publication/313108485_Sanger_sequencing_of_metastatic_pleural_effusion_An_inexpensive_minimally_invasive_tool_for_the_detection_of_EGFR_mutations_in_lung_adenocarcinoma-_Presented_at_the_National_Lung_Cancer_Conference_201). In.

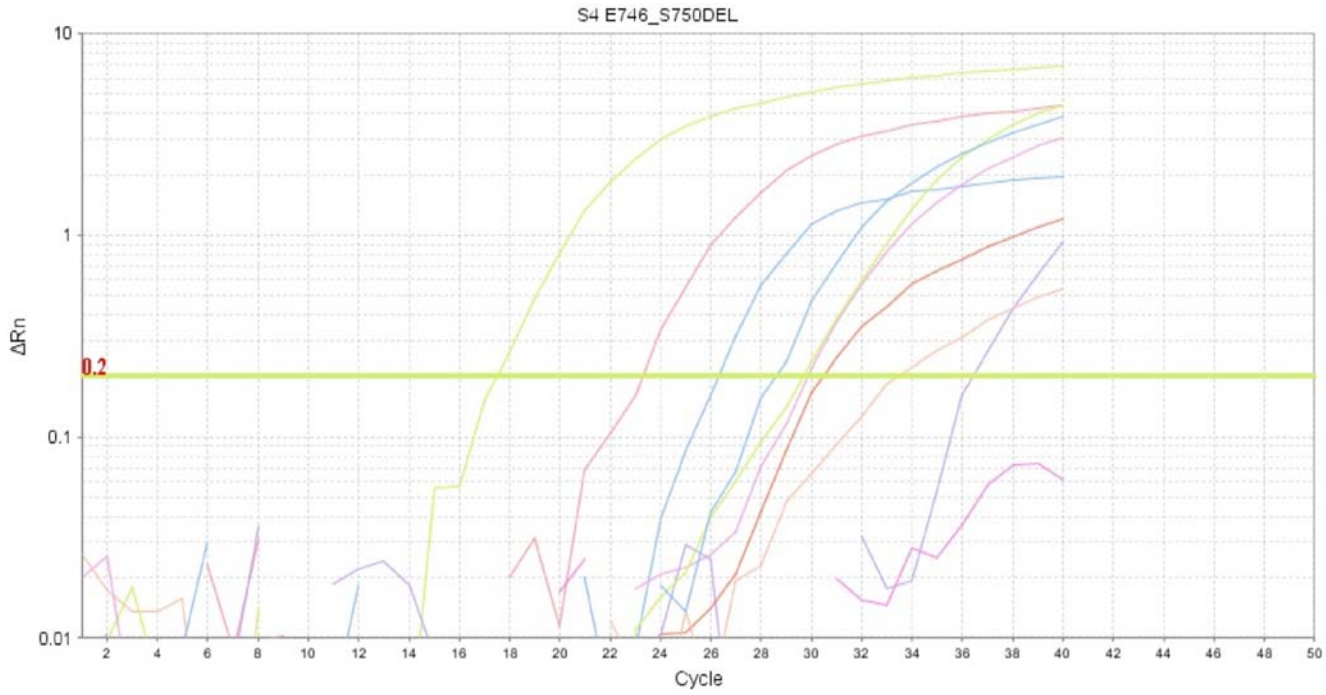
45. Dhar C, Sharma M, Nawaz S M et al. Detection of EGFR TKI sensitizing mutations from metastatic pleural fluid secondary to lung adenocarcinoma: a perspective from Southern India. [version 1; not peer reviewed]. F1000Research 2018, 7:1943 (poster) (https://doi.org/10.7490/f1000research.1116357.1). In.

46. 2018 ASCB Annual Meeting abstracts. Mol Biol Cell. 2018 Dec 15;29(26):3063.

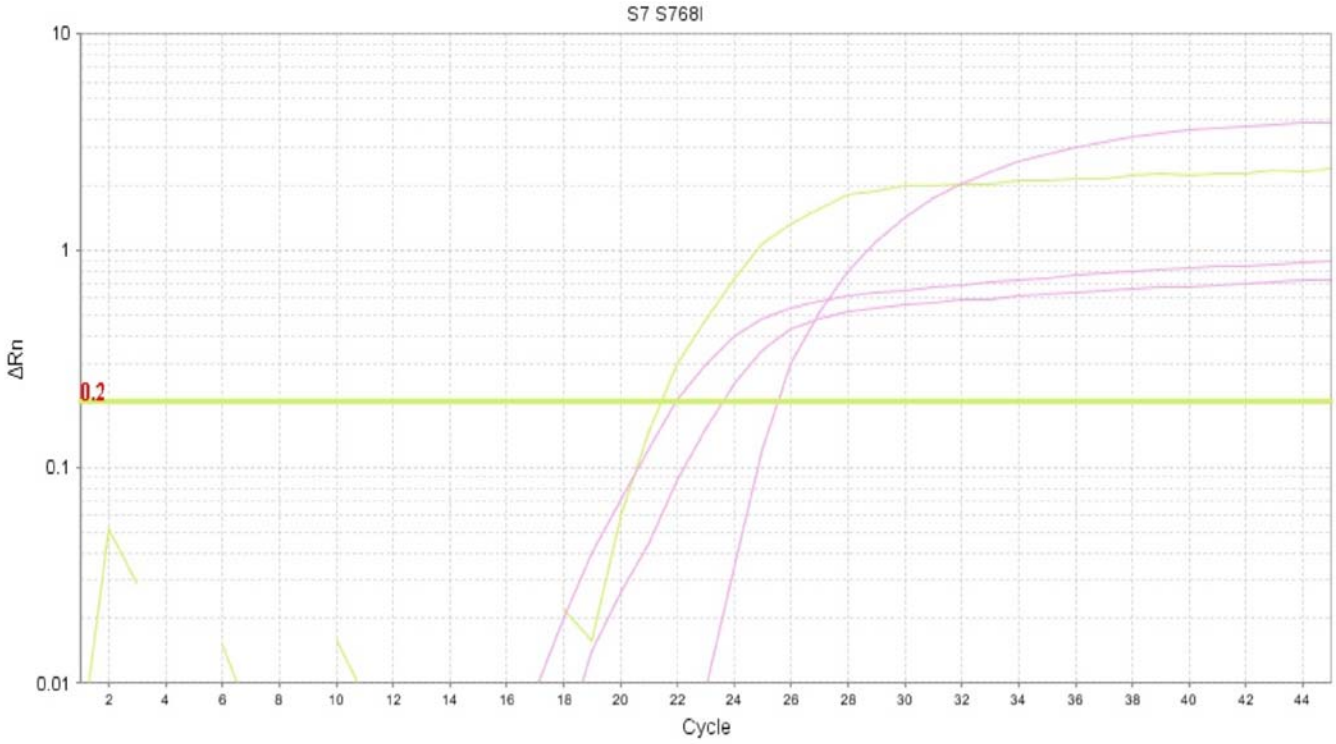
Supplementary Information



Supplemental figure 1 shows the amplification curves for different mutations tested in Sample S2



Supplemental figure 2 shows the amplification curves for different mutations tested in Sample S4



Supplemental figure 3 shows the amplification curves for different mutations tested in Sample S7

501

502

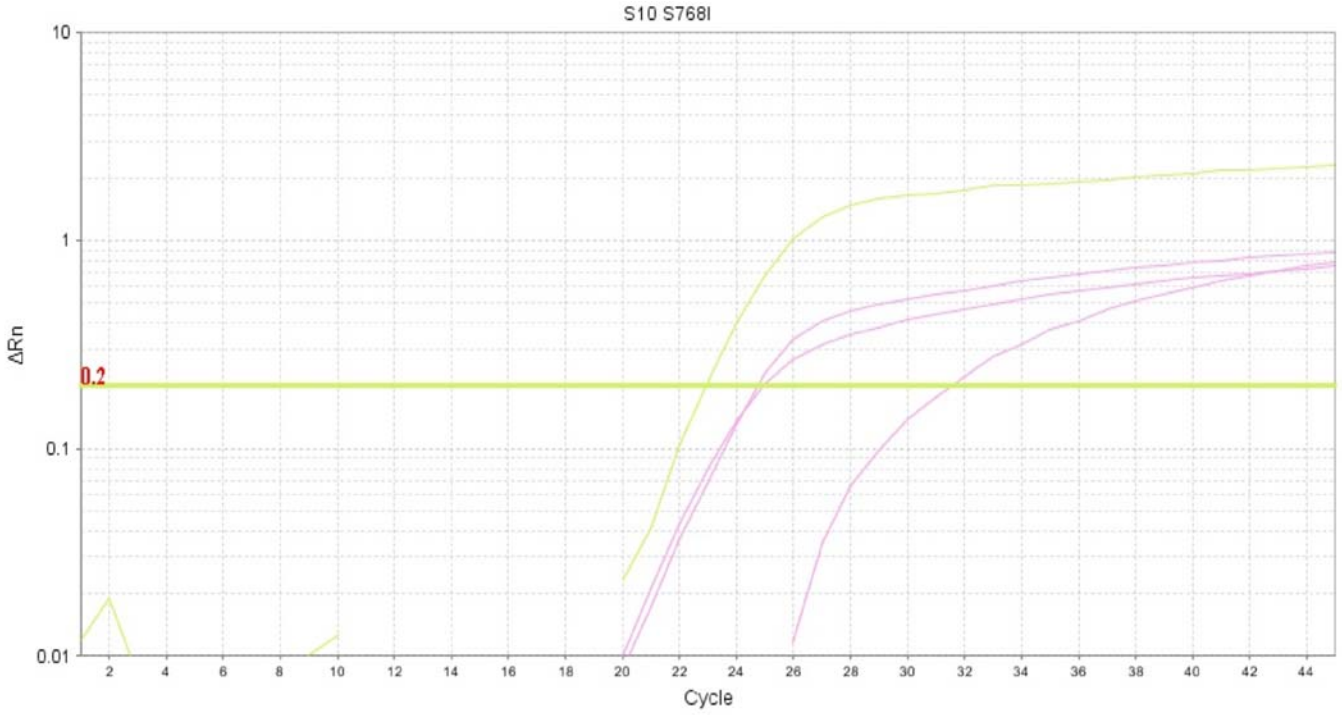
503

504

505

506

507



Supplemental figure 5 shows the amplification curves for different mutations tested in Sample S9

508

509

510

511

512

513

514

515

516

517

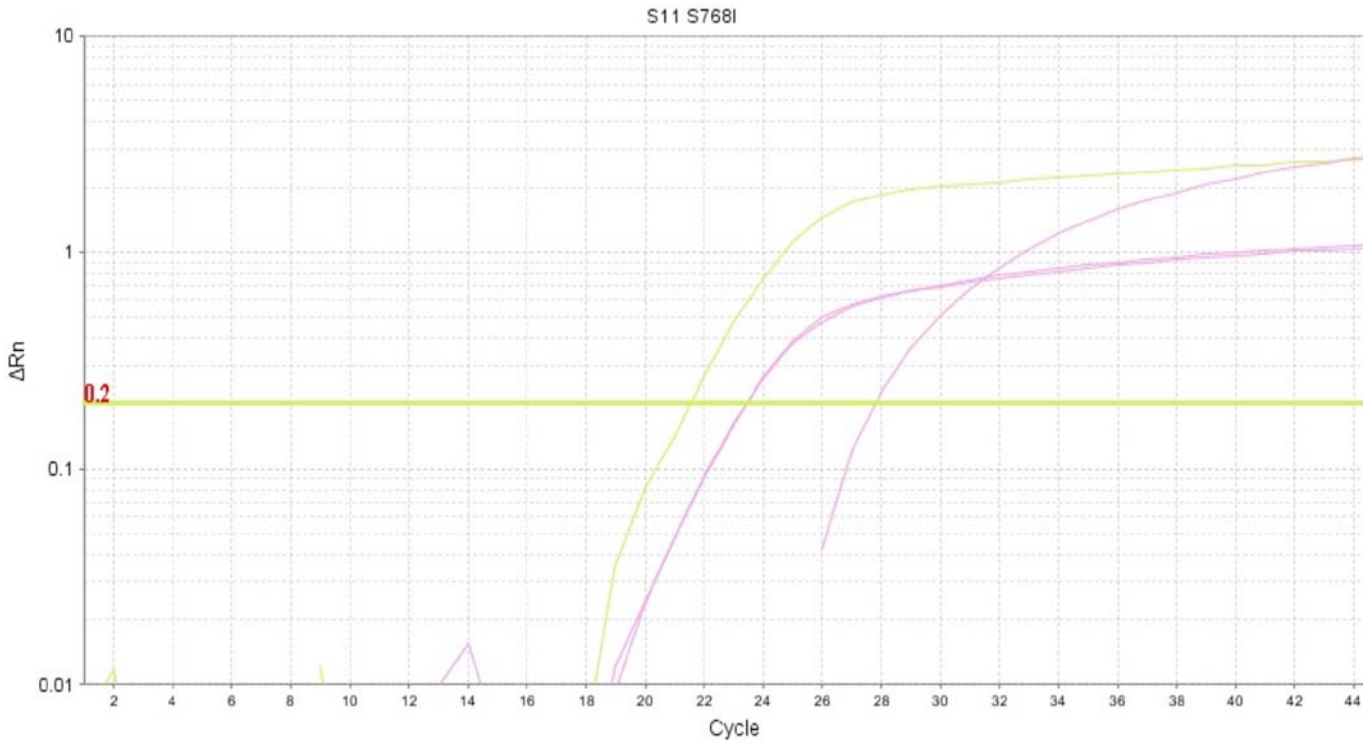
518

519

520

521

522



Supplemental figure 6 shows the amplification curves for different mutations tested in Sample S10

523

524

525

526

527

528

529

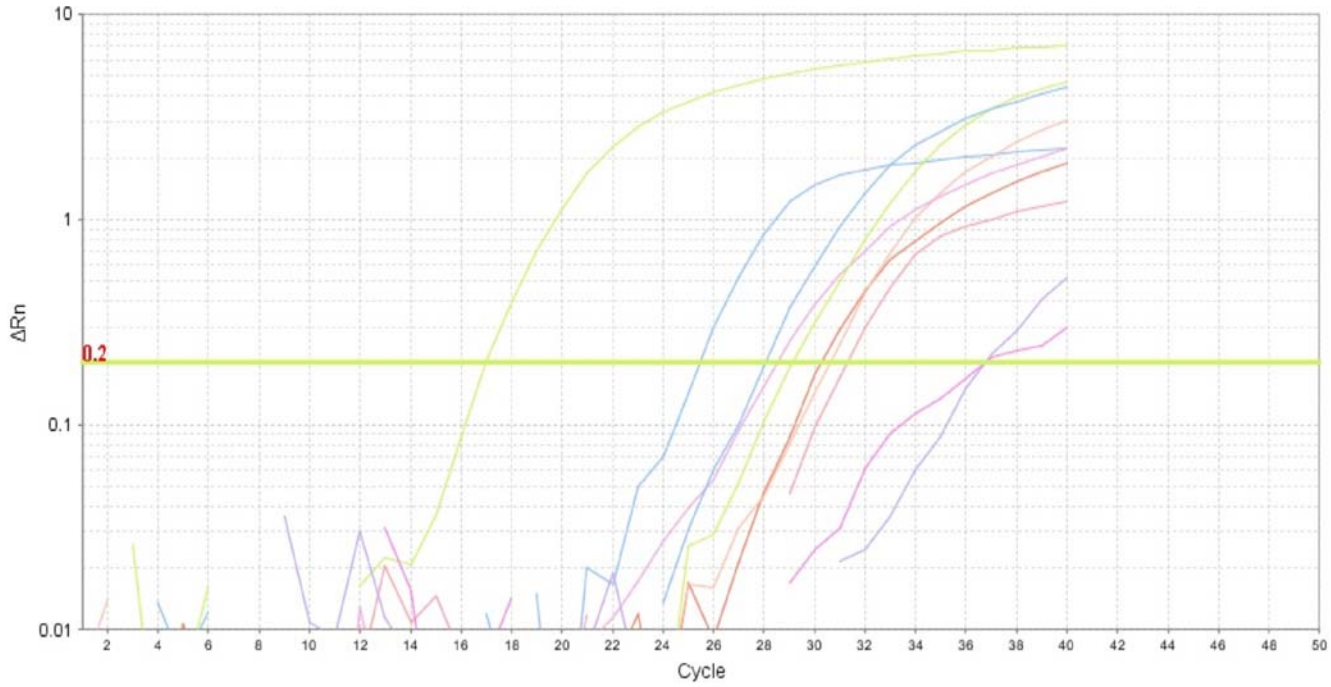
530

531

532

533

534



Supplemental figure 7 shows the representative amplification curves for different mutations tested and classified as a Wild Type sample or below LOD

535

536

537

538

Sample No.	ΔCt								
	G719X	**T790M**	**S768I**	**EX20INS**	**C797S**	**L858R**	**L861Q**	**E746_S750DEL**	**EX19DEL**
S1	11.261	12.919	12.393	15.518	13.704	21.527	19.990	14.589	8.697
S2	11.595	12.284	12.148	13.915	12.470	20.835	ND	**6.327**	9.159
S3	10.997	12.154	11.452	13.572	13.223	19.834	19.788	14.197	8.416
S4	10.981	12.106	12.276	15.961	12.879	18.905	ND	**5.597**	8.640
S5	12.438	11.107	11.385	14.670	19.424	ND	ND	ND	8.957
S6	11.314	12.079	11.776	17.363	12.901	ND	ND	14.173	8.642

Supplemental table 1 shows the ΔCt values of Samples S1- S6

539

540

541

542

Sample No.	ΔCt								
	G719X	**T790M**	**S768I**	**EX20INS**	**C797S**	**L858R**	**L861Q**	**E746_S750DEL**	**EX19DEL**
S7	12.212	20.890	**4.077**	ND	ND	ND	ND	ND	ND
S8	15.164	ND	13.393	ND	ND	ND	ND	ND	ND
S9	13.942	ND	**8.558**	ND	ND	ND	ND	ND	ND
S10	12.041	ND	**6.280**	ND	ND	ND	ND	ND	ND

Supplemental table 2 shows the ΔCt values of Samples S7-S11

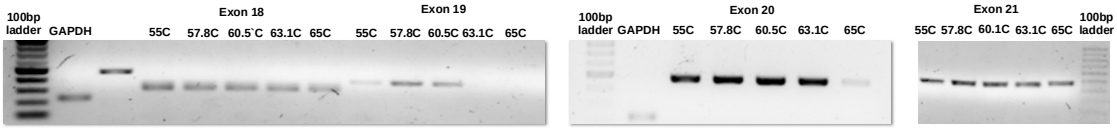
Supplementary table 3

List of newly designed primers

Region of Interest	Forward Primer 5’ to 3’	Reverse Primer 5’ to 3’
Exon 18	GTCCTTCCAAATGAGCTGGCAAG	ACAAAGAGTAAAGTAGATGATGG
Exon 19	TGTCCCTCACCTTCGGGGTGCAT	ACATTTAGGATGTGGAGATGAGCA
Exon 20	CTCAAGATCGCATTCATGCGTC	GACAGGCACTGATTTGTGCAC
Exon 21	AGTAGTCACTAACGTTCGCCAG	TCCCAGCAAGTACTGTTCCC

Optimization of a polymerase chain reaction protocol to amplify EGFR exon 18, 19, 20 and 21 from malignant pleural effusion (MPE).

Initially, we attempted a PCR amplification coupled with a direct sequencing approach to detect EGFR mutations from malignant pleural effusions. Appropriate primers (sequences listed in table 1) were designed and optimized on sample S0. Successful amplification of exons 18-21 by PCR on S0 are shown in figure 1. The primers for exon 18 and 21 had 55 ℃ as the annealing temperature while exon 19 and 20 were best amplified at an annealing temperature of 57.8 ℃.



Supplementary figure 8: 1.2% Agarose gel imaged under ultra-violet light: 100bp ladder, A: GAPDH, B: Exon 20, lanes C-G: Exon 19 at annealing temperatures of 55 C, 57.8 C, 60.5 C, 63.1 C and 65 C lanes H-L: Exon 18 at annealing temperatures of 55 C, 57.8 C, 60.5 C, 63.1 C and 65 C, lanes M-Q: Exon 21 at annealing temperatures of 55 C, 57.8 C, 60.5 C, 63.1 C and 65 C Interestingly, two distinct bands are visible for Exon 19 (lanes C-G)

List of genes with mutations co-occurring with S768I

Gene	Cytoband	(B) EGFR	(C) S768I	Alteration Overlap	Log Ratio	p-Value	q-Value	Enriched in
PDE1B	12q13.2	5 (0.70%)	6 (35.29%)		-5.66	2.58E-08	1.823E-04	(C) S768I
PNLIP	10q25.3	7 (0.98%)	6 (35.29%)		-5.17	9.33E-08	1.823E-04	(C) S768I
CTNNA1	5q31.2	3 (0.42%)	5 (29.41%)		-6.13	1.94E-07	1.823E-04	(C) S768I
BAG3	10q26.11	0 (0.00%)	4 (23.53%)		<-10	2.02E-07	1.823E-04	(C) S768I
CDCA7	2q31.1	0 (0.00%)	4 (23.53%)		<-10	2.02E-07	1.823E-04	(C) S768I
DACH2	Xq21.2	0 (0.00%)	4 (23.53%)		<-10	2.02E-07	1.823E-04	(C) S768I
KCNA7	19q13.33	0 (0.00%)	4 (23.53%)		<-10	2.02E-07	1.823E-04	(C) S768I
NKAP	Xq24	0 (0.00%)	4 (23.53%)		<-10	2.02E-07	1.823E-04	(C) S768I
ODF2	9q34.11	0 (0.00%)	4 (23.53%)		<-10	2.02E-07	1.823E-04	(C) S768I
PLSCR2	3q24	0 (0.00%)	4 (23.53%)		<-10	2.02E-07	1.823E-04	(C) S768I
RP2	Xp11.3	0 (0.00%)	4 (23.53%)		<-10	2.02E-07	1.823E-04	(C) S768I
SLC7A2	8p22	0 (0.00%)	4 (23.53%)		<-10	2.02E-07	1.823E-04	(C) S768I
TEKT3	17p12	0 (0.00%)	4 (23.53%)		<-10	2.02E-07	1.823E-04	(C) S768I
TOMM70	3q12.2	0 (0.00%)	4 (23.53%)		<-10	2.02E-07	1.823E-04	(C) S768I
TRIM22	11p15.4	4 (0.56%)	5 (29.41%)		-5.71	4.3E-07	3.627E-04	(C) S768I
TLR4	9q33.1	21 (2.94%)	7 (41.18%)		-3.81	8.27E-07	5.99E-04	(C) S768I
C2CD2L	11q23.3	1 (0.14%)	4 (23.53%)		-7.39	9.94E-07	5.99E-04	(C) S768I
GTF2IRD1	7q11.23	1 (0.14%)	4 (23.53%)		-7.39	9.94E-07	5.99E-04	(C) S768I
HIST2H2AB	1q21.2	1 (0.14%)	4 (23.53%)		-7.39	9.94E-07	5.99E-04	(C) S768I
PHC3	3q26.2	1 (0.14%)	4 (23.53%)		-7.39	9.94E-07	5.99E-04	(C) S768I
ZNF304	19q13.43	1 (0.14%)	4 (23.53%)		-7.39	9.94E-07	5.99E-04	(C) S768I
ATRX	Xq21.1	38 (5.65%)	8 (50.00%)		-3.15	1.847E-06	1.062E-03	(C) S768I
CCR6	6q27	2 (0.28%)	4 (23.53%)		-6.39	2.939E-06	1.329E-03	(C) S768I
FCMR	1q32.1	2 (0.28%)	4 (23.53%)		-6.39	2.939E-06	1.329E-03	(C) S768I
SMCR8	17p11.2	2 (0.28%)	4 (23.53%)		-6.39	2.939E-06	1.329E-03	(C) S768I
TNPO2	19p13.13	2 (0.28%)	4 (23.53%)		-6.39	2.939E-06	1.329E-03	(C) S768I
ZBTB10	8q21.13	2 (0.28%)	4 (23.53%)		-6.39	2.939E-06	1.329E-03	(C) S768I
ZNF428	19q13.31	2 (0.28%)	4 (23.53%)		-6.39	2.939E-06	1.329E-03	(C) S768I
FRMPD4	Xp22.2	8 (1.12%)	5 (29.41%)		-4.71	4.153E-06	1.812E-03	(C) S768I
CA1	8q21.2	3 (0.42%)	4 (23.53%)		-5.81	6.76E-06	2.413E-03	(C) S768I
NRIP1	21q11.2-q21.1	3 (0.42%)	4 (23.53%)		-5.81	6.76E-06	2.413E-03	(C) S768I
TMPRSS11D	4q13.2	3 (0.42%)	4 (23.53%)		-5.81	6.76E-06	2.413E-03	(C) S768I
PPM1D	17q23.2	3 (0.51%)	4 (26.67%)		-5.70	8.603E-06	2.413E-03	(C) S768I
ASB10	7q36.1	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
BEST3	12q15	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
CEP44	4q34.1	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
CPT1A	11q13.3	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
ENTPD7	10q24.2	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I

575

EPDR1	7p14.1	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
FOXL1	16q24.1	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
GGNBP2	17q12	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
IFNLR1	1p36.11	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
KCNJ6	21q22.13	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
KRT3	12q13.13	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
MIOS	7p21.3	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
MOGAT2	11q13.5	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
OR6C2	12q13.2	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
PARVA	11p15.3	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
PPP2R5A	1q32.3	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
PSMD7	16q23.1	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
SIRPD	20p13	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
SKIV2L	6p21.33	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
WDR3	1p12	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
ZNF474	5q23.2	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
ZNF720	16p11.2	0 (0.00%)	3 (17.65%)		<-10	1.049E-05	2.413E-03	(C) S768I
ETNPPL	4q25	4 (0.56%)	4 (23.53%)		-5.39	1.333E-05	2.635E-03	(C) S768I
GPR34	Xp11.4	4 (0.56%)	4 (23.53%)		-5.39	1.333E-05	2.635E-03	(C) S768I
GTPBP1	22q13.1	4 (0.56%)	4 (23.53%)		-5.39	1.333E-05	2.635E-03	(C) S768I
PARP8	5q11.1	4 (0.56%)	4 (23.53%)		-5.39	1.333E-05	2.635E-03	(C) S768I
SLTM	15q22.1	4 (0.56%)	4 (23.53%)		-5.39	1.333E-05	2.635E-03	(C) S768I
SPOCD1	1p35.2	4 (0.56%)	4 (23.53%)		-5.39	1.333E-05	2.635E-03	(C) S768I
TEX11	Xq13.1	4 (0.56%)	4 (23.53%)		-5.39	1.333E-05	2.635E-03	(C) S768I
TMEM246	9q31.1	4 (0.56%)	4 (23.53%)		-5.39	1.333E-05	2.635E-03	(C) S768I
ZGPAT	20q13.33	4 (0.56%)	4 (23.53%)		-5.39	1.333E-05	2.635E-03	(C) S768I
ARHGAP10	4q31.23	5 (0.70%)	4 (23.53%)		-5.07	2.364E-05	4.4E-03	(C) S768I
CD5	11q12.2	5 (0.70%)	4 (23.53%)		-5.07	2.364E-05	4.4E-03	(C) S768I
CEP85L	6q22.31	5 (0.70%)	4 (23.53%)		-5.07	2.364E-05	4.4E-03	(C) S768I
MTHFD1	14q23.3	5 (0.70%)	4 (23.53%)		-5.07	2.364E-05	4.4E-03	(C) S768I
FLG	1q21.3	40 (5.60%)	7 (41.18%)		-2.88	3.463E-05	5.946E-03	(C) S768I
DNMT3A	2p23.3	6 (0.89%)	4 (25.00%)		-4.81	3.772E-05	5.946E-03	(C) S768I
ACOX3	4p16.1	6 (0.84%)	4 (23.53%)		-4.81	3.884E-05	5.946E-03	(C) S768I
DNAH6	2p11.2	6 (0.84%)	4 (23.53%)		-4.81	3.884E-05	5.946E-03	(C) S768I
GSTA1	6p12.2	6 (0.84%)	4 (23.53%)		-4.81	3.884E-05	5.946E-03	(C) S768I
KAT6B	10q22.2	6 (0.84%)	4 (23.53%)		-4.81	3.884E-05	5.946E-03	(C) S768I
NCOA7	6q22.31-q22.32	6 (0.84%)	4 (23.53%)		-4.81	3.884E-05	5.946E-03	(C) S768I
PLCB3	11q13.1	6 (0.84%)	4 (23.53%)		-4.81	3.884E-05	5.946E-03	(C) S768I
RNF6	13q12.13	6 (0.84%)	4 (23.53%)		-4.81	3.884E-05	5.946E-03	(C) S768I
SERPINA7	Xq22.3	6 (0.84%)	4 (23.53%)		-4.81	3.884E-05	5.946E-03	(C) S768I

576

TUBGCP6	22q13.33	6 (0.84%)	4 (23.53%)		-4.81	3.884E-05	5.946E-03	(C) S768I
C6ORF136	6p21.33	1 (0.14%)	3 (17.65%)		-6.98	4.135E-05	5.946E-03	(C) S768I
CIZ1	9q34.11	1 (0.14%)	3 (17.65%)		-6.98	4.135E-05	5.946E-03	(C) S768I
DNA2	10q21.3	1 (0.14%)	3 (17.65%)		-6.98	4.135E-05	5.946E-03	(C) S768I
H6PD	1p36.22	1 (0.14%)	3 (17.65%)		-6.98	4.135E-05	5.946E-03	(C) S768I
LRIG1	3p14.1	1 (0.14%)	3 (17.65%)		-6.98	4.135E-05	5.946E-03	(C) S768I
LUZP4	Xq23	1 (0.14%)	3 (17.65%)		-6.98	4.135E-05	5.946E-03	(C) S768I
MTMR14	3p25.3	1 (0.14%)	3 (17.65%)		-6.98	4.135E-05	5.946E-03	(C) S768I
RAB3C	5q11.2	1 (0.14%)	3 (17.65%)		-6.98	4.135E-05	5.946E-03	(C) S768I
SYT3	19q13.33	1 (0.14%)	3 (17.65%)		-6.98	4.135E-05	5.946E-03	(C) S768I
FSHR	2p16.3	15 (2.10%)	5 (29.41%)		-3.81	4.536E-05	6.45E-03	(C) S768I
ADAM8	10q26.3	8 (1.12%)	4 (23.53%)		-4.39	8.895E-05	0.0122	(C) S768I
GPR52	1q25.1	8 (1.12%)	4 (23.53%)		-4.39	8.895E-05	0.0122	(C) S768I
MYO16	13q33.3	8 (1.12%)	4 (23.53%)		-4.39	8.895E-05	0.0122	(C) S768I
ANAPC7	12q24.11	2 (0.28%)	3 (17.65%)		-5.98	1.019E-04	0.0126	(C) S768I
CARD14	17q25.3	2 (0.28%)	3 (17.65%)		-5.98	1.019E-04	0.0126	(C) S768I
CH25H	10q23.31	2 (0.28%)	3 (17.65%)		-5.98	1.019E-04	0.0126	(C) S768I
FFAR1	19q13.12	2 (0.28%)	3 (17.65%)		-5.98	1.019E-04	0.0126	(C) S768I
GADD45A	1p31.3	2 (0.28%)	3 (17.65%)		-5.98	1.019E-04	0.0126	(C) S768I
MEIS3	19q13.32	2 (0.28%)	3 (17.65%)		-5.98	1.019E-04	0.0126	(C) S768I
OR2M7	1q44	2 (0.28%)	3 (17.65%)		-5.98	1.019E-04	0.0126	(C) S768I
TBC1D9B	5q35.3	2 (0.28%)	3 (17.65%)		-5.98	1.019E-04	0.0126	(C) S768I
TSNARE1	8q24.3	2 (0.28%)	3 (17.65%)		-5.98	1.019E-04	0.0126	(C) S768I
USP9X	Xp11.4	2 (0.28%)	3 (17.65%)		-5.98	1.019E-04	0.0126	(C) S768I
A2ML1	12p13.31	9 (1.26%)	4 (23.53%)		-4.22	1.266E-04	0.0151	(C) S768I
MCOLN3	1p22.3	9 (1.26%)	4 (23.53%)		-4.22	1.266E-04	0.0151	(C) S768I
PCDHA6	5q31.3	9 (1.26%)	4 (23.53%)		-4.22	1.266E-04	0.0151	(C) S768I
RAG1	11p12	9 (1.26%)	4 (23.53%)		-4.22	1.266E-04	0.0151	(C) S768I
NAALAD2	11q14.3	10 (1.40%)	4 (23.53%)		-4.07	1.747E-04	0.0201	(C) S768I
NME8	7p14.1	10 (1.40%)	4 (23.53%)		-4.07	1.747E-04	0.0201	(C) S768I
PKD1	16p13.3	10 (1.40%)	4 (23.53%)		-4.07	1.747E-04	0.0201	(C) S768I
TEX10	9q31.1	10 (1.40%)	4 (23.53%)		-4.07	1.747E-04	0.0201	(C) S768I
C11ORF88	11q23.1	3 (0.42%)	3 (17.65%)		-5.39	2.008E-04	0.0210	(C) S768I
CASZ1	1p36.22	3 (0.42%)	3 (17.65%)		-5.39	2.008E-04	0.0210	(C) S768I
DMWD	19q13.32	3 (0.42%)	3 (17.65%)		-5.39	2.008E-04	0.0210	(C) S768I
DPYSL4	10q26.3	3 (0.42%)	3 (17.65%)		-5.39	2.008E-04	0.0210	(C) S768I
GABRG3	15q12	3 (0.42%)	3 (17.65%)		-5.39	2.008E-04	0.0210	(C) S768I
GALNT15	3p25.1	3 (0.42%)	3 (17.65%)		-5.39	2.008E-04	0.0210	(C) S768I
GZMA	5q11.2	3 (0.42%)	3 (17.65%)		-5.39	2.008E-04	0.0210	(C) S768I
LDHA	11p15.1	3 (0.42%)	3 (17.65%)		-5.39	2.008E-04	0.0210	(C) S768I

577

OPRL1	20q13.33	3 (0.42%)	3 (17.65%)		-5.39	2.008E-04	0.0210	(C) S768I
PCSK6	15q26.3	3 (0.42%)	3 (17.65%)		-5.39	2.008E-04	0.0210	(C) S768I
SMC1B	22q13.31	3 (0.42%)	3 (17.65%)		-5.39	2.008E-04	0.0210	(C) S768I
CDK17	12q23.1	2 (0.52%)	3 (25.00%)		-5.59	2.053E-04	0.0213	(C) S768I
FGF4	11q13.3	3 (0.43%)	3 (17.65%)		-5.37	2.109E-04	0.0217	(C) S768I
ADGRL1	19p13.12	11 (1.54%)	4 (23.53%)		-3.93	2.349E-04	0.0238	(C) S768I
MED13	17q23.2	11 (1.54%)	4 (23.53%)		-3.93	2.349E-04	0.0238	(C) S768I
MUC17	7q22.1	56 (7.84%)	7 (41.18%)		-2.39	2.483E-04	0.0249	(C) S768I
ITGA8	10p13	24 (3.36%)	5 (29.41%)		-3.13	3.061E-04	0.0305	(C) S768I
UHRF1BP1L	12q23.1	12 (1.68%)	4 (23.53%)		-3.81	3.086E-04	0.0305	(C) S768I
ADGRF2	6p12.3	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
AKR7A3	1p36.13	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
CEP89	19q13.11	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
COL23A1	5q35.3	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
CSF2RB	22q12.3	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
ELP4	11p13	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
MMEL1	1p36.32	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
NKTR	3p22.1	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
OR2M4	1q44	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
PARM1	4q13.3	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
TNNT3	11p15.5	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
VGLL1	Xq26.3	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
ZNF226	19q13.31	4 (0.56%)	3 (17.65%)		-4.98	3.464E-04	0.0311	(C) S768I
ABCB4	7q21.12	13 (1.82%)	4 (23.53%)		-3.69	3.978E-04	0.0355	(C) S768I
ACE	17q23.3	14 (1.96%)	4 (23.53%)		-3.58	5.041E-04	0.0398	(C) S768I
DSG2	18q12.1	14 (1.96%)	4 (23.53%)		-3.58	5.041E-04	0.0398	(C) S768I
TTC14	3q26.33	14 (1.96%)	4 (23.53%)		-3.58	5.041E-04	0.0398	(C) S768I
C7ORF31	7p15.3	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
CC2D2A	4p15.32	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
EHD2	19q13.33	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
EYA1	8q13.3	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
FAM193B	5q35.3	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
FIBP	11q13.1	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
GYG2	Xp22.33	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
HYKK	15q25.1	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
KLK13	19q13.41	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
LMAN2L	2q11.2	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
OR5AN1	11q12.1	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
SPANXN3	Xq27.3	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
TMEM161A	19p13.11	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I

578

TRMT1	19p13.13	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
TSPAN31	12q14.1	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
TUBA3D	2q21.1	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
ZNF829	19q13.12	0 (0.00%)	2 (11.76%)		<-10	5.097E-04	0.0398	(C) S768I
ADAM2	8p11.22	5 (0.70%)	3 (17.65%)		-4.66	5.462E-04	0.0400	(C) S768I
CPA4	7q32.2	5 (0.70%)	3 (17.65%)		-4.66	5.462E-04	0.0400	(C) S768I
CTTNBP2	7q31.31	5 (0.70%)	3 (17.65%)		-4.66	5.462E-04	0.0400	(C) S768I
DUSP19	2q32.1	5 (0.70%)	3 (17.65%)		-4.66	5.462E-04	0.0400	(C) S768I
ELL2	5q15	5 (0.70%)	3 (17.65%)		-4.66	5.462E-04	0.0400	(C) S768I
GPR141	7p14.1	5 (0.70%)	3 (17.65%)		-4.66	5.462E-04	0.0400	(C) S768I
ITGB4	17q25.1	5 (0.70%)	3 (17.65%)		-4.66	5.462E-04	0.0400	(C) S768I
KCTD19	16q22.1	5 (0.70%)	3 (17.65%)		-4.66	5.462E-04	0.0400	(C) S768I
SNX13	7p21.1	5 (0.70%)	3 (17.65%)		-4.66	5.462E-04	0.0400	(C) S768I
TSGA10	2q11.2	5 (0.70%)	3 (17.65%)		-4.66	5.462E-04	0.0400	(C) S768I
TULP1	6p21.31	5 (0.70%)	3 (17.65%)		-4.66	5.462E-04	0.0400	(C) S768I
GLI2	2q14.2	15 (2.10%)	4 (23.53%)		-3.49	6.293E-04	0.0455	(C) S768I
GRM1	6q24.3	15 (2.10%)	4 (23.53%)		-3.49	6.293E-04	0.0455	(C) S768I