Germline *ERBB3* mutation in familial non-small cell lung carcinoma: expanding ErbB's role in oncogenesis.

Authors: Aideen M. McInerney-Leo, PhD,[1] Hui Yi Chew, BSc,[1] Po-Ling Inglis, MBBS FRACP,[2] Paul J. Leo, BSc,[3] Shannon R. Joseph, PhD,[1] Caroline L. Cooper, MBBS FRCPA,[4,5] Satomi Okano, MBiostatistics,[1] Tim Hassall, MBBS FRACP,[6] Lisa Anderson, BSc,[2] Rayleen V. Bowman, MBBS FRACP PhD, [5,7] Michael Gattas, MBBS FRACP,[8] Jessica E. Harris, MSc,[3] Mhairi S. Marshall, MSc,[3] Janet G. Shaw, BSc, [5,7] Lawrie Wheeler, BSc,[3] Ian A. Yang, MBBS FRACP PhD,[5,7] Matthew A. Brown, MBBS MD FRACP,[3,9,10] Kwun M. Fong, MBBS FRACP PhD,[5,7*] Fiona Simpson, PhD,[1*] Emma L. Duncan, MBBS MRCP FRACP PhD,[3,5,11*].

[*]Joint senior authors

[1]The Dermatology Research Centre, The University of Queensland Diamantina Institute, The University of Queensland, Woolloongabba, QLD, 4102.

[2]Medical Oncology, Royal Brisbane and Women's Hospital, Herston, QLD, 4029.

[3]Australian Translational Genomics Centre, Institute of Health and Biomedical Innovation, School of Biomedical Sciences, Queensland University of Technology (QUT), Translational Research Institute, 37 Kent St, Woolloongabba, QLD, 4102.

[4]Department of Anatomical Pathology, Pathology Queensland, Princess Alexandra Hospital, Brisbane

[5] UQTRC, Faculty of Medicine, The University of Queensland, 288 Herston Road, Herston, QLD, 4006.

[6] Queensland Children's Hospital, South Brisbane, QLD, 4101.

[7]Department of Thoracic Medicine, The Prince Charles Hospital, Rode Road, Chermside, QLD, 4032.

[8]Genetic Health Queensland, Royal Brisbane and Women's Hospital, Herston, QLD, 4029.

[9] Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom

[10] King's College London NIHR Biomedical Research Centre, King's College London, United Kingdom

[11]Department of Endocrinology, Royal Brisbane and Women's Hospital, Butterfield St, Herston, QLD, 4029.

Running title: Germline *ERBB3* mutation in autosomal dominant NSCLC

Keywords: *ERBB3*, germline mutation, autosomal dominant, lung cancer, signalling pathway activation.

Corresponding author:

Professor Emma L. Duncan

Department of Twin Research & Genetic Epidemiology, School of Life Course Sciences

Faculty of Life Sciences and Medicine

King's College London

Mailing Address:

Department of Twin Research & Genetic Epidemiology

St Thomas' Campus

Lambeth Palace Road

London SE1 7EH

United Kingdom

Email: emma.duncan@kcl.ac.uk

Disclaimers: The authors declare no competing financial interests.

Word Count: 2426 (excluding abstract, reference list, tables, and figure legends)

Number of Figures and Tables: Three figures and one table.

Abbreviations: CNV, copy number variation; EGFR, Epidermal Growth Factor Receptor; GWAS, genome-wide association studies; MAF, minor allele frequency; NSCLC, non-small-cell lung cancer; TKIs, tyrosine kinase inhibitors.

ABSTRACT

Background

Lung cancer is the commonest cause of cancer deaths worldwide. Although strongly associated with smoking, predisposition to lung cancer is also heritable with multiple common risk variants identified. Rarely, dominantly inherited non-small-cell lung cancer (NSCLC) has been reported due to somatic mutations in *EGFR/ErbB1* and *ERBB2*.

Methods

Germline exome sequencing was performed in a multi-generation family with autosomal dominant NSCLC, including an affected child. Tumour samples were also sequenced. Full-length wild-type (wtErbB3) and mutant ERBB3 (mutErbB3) constructs were transfected into HeLa cells. Protein expression, stability, and sub-cellular localisation were assessed; and cellular proliferation, pAkt/Akt, and pERK levels were determined.

Results

A novel germline variant in *ERBB3* (c.1946T>G: p.Iso649Arg), coding for receptor tyrosine-protein kinase erbB-3 (ErbB3), was identified, with appropriate segregation. There was no loss-of-heterozygosity in tumour samples. Both wtErbB3 and mutErbB3 were stably expressed. MutErbB3-transfected cells demonstrated an increased ratio of the 80kD form (which enhances proliferation) compared to the full-length (180kD) form. MutErbB3 and wtErbB3 had similar punctate cytoplasmic localisation pre- and post-EGF stimulation; however, EGFR levels decreased faster post-stimulation in mutErbB3-transfected cells, suggesting more rapid processing of the mutErbB3/EGFR heterodimer. Cellular proliferation was increased in mutErbB3-transfected cells compared to wtErbB3 transfection. MutErbB3-transfected cells also showed decreased pAkt/tAkt ratios and increased pERK/tERK 30 minutes post-stimulation compared to wtErbB3 transfection, demonstrating altered signalling

pathway activation by mutErbB3. Cumulatively, these results support this mutation as tumorogenic.

Conclusions

This is the first reported family with a germline *ERBB3* mutation causing heritable NSCLC, furthering understanding of the ErbB family pathway in oncogenesis.

INTRODUCTION

Lung cancer is the leading cause of cancer deaths worldwide (World Health Organisation) [1], with over 80% of cases attributable to smoking. However, lung cancer is also heritable, with heritability of ~18% [2]. Genome-wide association studies (GWAS) have identified multiple susceptibility loci for lung cancer overall (reviewed [3, 4]), for non-small cell lung cancer (NSCLC) [5] and for histology-specific sub-types of NSCLC [6] (with specific GWAS in squamous cell carcinoma [7] and adenocarcinoma [8], but not large cell to date). There have also been many reports of familial aggregation of lung cancer, (summarised [9]), with increased familial risk particularly observed in cases with younger age of onset [10, 11], of female gender, and with adenocarcinoma [12], even after adjusting for smoking status [12, 13]. Linkage and association studies in familial lung cancer have identified unique susceptibility loci, as well as confirming loci associated with NSCLC overall and with specific NSCLC subtypes [14-18]. Additionally, GWAS have identified unique susceptibility loci for NSCLC cases carrying somatic EGFR mutations [19, 20].

Somatic gain-of-function mutations affecting the tyrosine kinase (TK) domain of Epidermal Growth Factor Receptor (EGFR) are common in non-small cell lung cancer (NSCLC), particularly adenocarcinoma, and predict responsiveness to EGFR-targeting tyrosine kinase inhibitors (TKIs) [21]. Extremely rarely, germline carriage of *EGFR* mutations has been described in families with autosomal dominant NSCLC, occasionally with additional somatic *EGFR* mutations [22, 23]. EGFR (ErbB1, Human EGF Receptor [HER] 1) belongs to the ErbB family of receptor tyrosine kinases which includes ErbB2 (neu, HER2), ErbB3 (HER3) and ErbB4 (HER4). A germline *ERBB2* mutation was identified in another family with autosomal dominant NSCLC, without additional *HER2* somatic variant(s) [24]. No paediatric NSCLC were reported in these families; indeed, primary lung

116 cancers in children are very rare [25, 26]. Notably, none of the loci associated with lung
117 cancer in the many GWAS to date have included *EGFR* or other ERBB family members [4].
118
119 Here, we report a new causative gene in a family with autosomal dominant NSCLC.
120
121 **MATERIALS AND METHODS**
122 This study was approved by The Prince Charles Hospital Metro North Human
123 Research Ethics Committee (approval HREC/13/QPCH/216). Participants gave informed
124 written consent.
125
126 Detailed methods are presented in Supplementary Data. Briefly, exome sequencing
127 was performed on germline DNA in a multi-generational family with autosomal dominant
128 NSCLC (Fig. 1). Given the rarity of autosomal dominant NSCLC, and paediatric lung
129 malignancies overall [25, 26], analysis focussed on rare variants (previously unreported; and
130 with minor allele frequency [MAF] <0.001), assessed against internal and external databases
131 (e.g. gnomAD [27], 1000 Genomes [28], and dbSNP137 [29]).
132
133 Formalin-fixed paraffin-embedded [FFPE] samples were obtained from individuals
134 LGCA-1.2, LGCA-1.3 and LGCA-1.6, with DNA extracted and sequenced. Expression and
135 localisation of ErbB3 in normal and tumour tissue was assessed by immunohistochemistry.
136
137 Full-length wild-type (wtErbB3) and mutant (mutErbB3, c.1946T>G: p.Iso649Arg)
138 *ERBB3* expression constructs were produced and transfected into HeLa cells (which do not
139 express endogenous ErbB3 or ErbB2, but do express EGFR (ErbB1), the preferred
140 dimerisation partner of ErbB3 [30]). To evaluate protein size and conformation, Western

blotting was performed on lysates from transfected HeLa cells (vector-only, wtErbB3, or mutErbB3), probed with commercial anti-bodies against ErbB3 with β-tubulin used as protein-loading control. To assess localisation pre- and post-stimulation, transfected cells were either fixed (0') or stimulated with 10ng EGF-Alexa Fluor 488 (30') prior to fixation, and immunostained for ErbB3 and endogenous EGFR, with nuclei stained using DAPI. Transfected cells (vector-only, wtErbB3 or mutErbB3, co-transfected with green fluorescence protein) were separated by fluorescence-activated cell sorting and proliferation rate assessed. Signalling pathway activation of ErbB3 and mutErbB3 transfected cells were analysed by immunoblotting for ErbB3, EGFR, Akt (phospho- and total) and ERK (phospho- and total) in cells grown in full serum (control), 3 hours post-serum starvation (0) and post-EGF stimulation (10ng/ml) at 10 minutes and 30 minutes. Relative protein expression was quantified and the ratio normalised to β-Tubulin (used as a loading control) to enable the quantification of phospho- to total-Akt (pAkt/tAkt) and phospho- to total-ERK (pERK/tERK). Results are presented without formal statistical assessment, as is conventional for these analyses [31].

RESULTS

Clinical details

The proband (LGCA-1.2) presented with lung adenocarcinoma aged 51 years. Her father and paternal grandfather, died of NSCLC aged 39 and 34 years, respectively. Two of her five children have lung adenocarcinoma, presenting aged 12 and 30 years (Fig. 1). The proband, her father and grandfather had all smoked at some stage; however, neither of the children had ever smoked.

Exome sequencing

Four novel good-quality variants affecting highly conserved bases and with appropriate familial segregation were identified, three of which were predicted damaging by at least two protein prediction algorithms (Table 1; Filtering steps presented in Supplementary Data: Table S1). Of these, the *ERBB3* variant (NM_001982; c.1946T>G; p.Ile649Arg) was of particular interest given the known oncogenic role of ErbB3 itself [32], and of other ErbB family members in heritable NSCLC In considering the other two variants: *SORBS1* (*Sorbin and SH3 Domain Containing 1*) is involved in cell adhesion, growth factor signalling and cytoskeleton formation; but appears mainly to regulate insulin-mediated glucose uptake [33]. *ATG2B* (*Autophagy-Related Protein 2 Homolog B*) is involved in autophagy, a key pathway mediating stress-induced adaptation and cellular damage control. However, although exploited by cancer cells to survive stressors (e.g., starvation, hypoxia, and chemotherapy), autophagy is not considered an oncogenic driver process *per se* [34].

Filtering the data with a less stringent MAF threshold (MAF<0.001) identified variants in eight additional candidate genes (Supplementary Data: Table S2), of which one (*PAXIP1)* is a genome stability gene previously associated with cancer [35]. Somatic copy number variation (CNV) of *PAXIP1* has also been associated with breast cancer prognosis [36]. However, the identified variant (rs199937188) is predicted benign and tolerated by Polyphen [37] and SIFT [38].

The data were also interrogated for coding variants in genes previously implicated in familial lung cancer (specifically, *EGFR, ERBB2, TP53* or *PARK2*); none were detected.

Tumour sequencing

Sanger sequencing of tumour DNA excluded homozygosity of the *ERBB3* variant (data not shown). Unfortunately, tumour DNA from FFPE samples was too degraded for massively parallel sequencing, precluding comprehensive assessment of *ERBB3* somatic variants.

Immunohistochemistry for ERBB3

ErbB3 is typically upregulated in NSCLC, staining both membrane and cytoplasm [39]. Tumour tissue from the proband (LGCA-1.2) showed weak cytoplasmic ErbB3, with absent staining of normal surrounding lung tissue (Supplementary Data: Figure S1) Results from other tumour samples were inconsistent; notably, less tissue was made available from these other individuals for this study, as their tumour samples were required to inform their ongoing clinical care.

ErbB3 expression, folding, and cytoplasmic localisation

Both wtErbB3 and mutErbB3 were folded and expressed stably (Fig. 2A) with similar sub-cellular localisation (Fig. 2B). Cytoplasmic organelle distribution mirrored that of endogenous ErbB3 (in other, non-HeLa cells; data not shown). Compared with wtErbB3, cells expressing mutErbB3 showed a higher ratio of the 80kDa to full-length (~180kDa) forms (Fig. 2A). Without EGF stimulation, EGFR co-localised with mutErbB3 in concentrated puncta in the endosomal system, which was less evident with wtErbB3 (Fig. 2B). After EGF stimulation, both wtErbB3 and mutErbB3 increased in the perinuclear region (Fig. 2B), also co-localising with EGFR at this time-point (Fig. 2B, 30’ time-point).

Cell proliferation

HeLa cells expressing mutErbB3 demonstrated increased cellular proliferation, when compared to either HeLa cells expressing wtErbB3 or vector-only (Fig. 2C).

Signalling Pathway Activation

After EGF stimulation, expression levels of mutErbB3 and wtErbB3 were comparable at all time-points; though, EGFR levels decreased over time in cells expressing mutErbB3 compared with wtErbB3 (Fig. 3A).

Cells expressing either mutErbB3 or wtErbB3 had increased pAkt levels, compared with vector-only transfected cells (Fig. 3B). After starvation followed by 10 minutes' EGF stimulation, similar pAkt/tAkt ratios were observed in mutErbB3, wtErbB3, and vector-only transfected cells (Fig. 3B). However, by 30 minutes, mutErbB3-transfected cells had decreased pAkt/tAkt ratios compared with both wtErBB3 and vector-only transfected cells (Fig. 3B). Of note, mutErbB3-expressing cells had increased pERK at the 30-minute time-point (Fig. 3B). These findings show that mutErbB3 is changing the signalling activation profile in response to ligand stimulation.

Together, these results suggest that the EGFR/mutErbB3 heterodimer is more efficiently activated, internalised and degraded, compared with EGFR/wtErbB3.

DISCUSSION

We have identified a novel germline mutation in *ERBB3* segregating with autosomal dominant NSCLC. We demonstrate that mutErbB3 is stably expressed, functional with EGFR heterodimerisation and signalling, with an increased ratio of 80kDa vs. full-length 180kDa ErbB, a faster time-course of signalling activation and degradation, and increased cellular proliferation, compared to wtErbB3. These results support this mutation as the oncogenic driver of NSCLC in this family.

Multiple studies demonstrate the importance of ErbB3 in oncogenesis generally and NSCLC specifically [40]. *ERBB3* is part of a five-gene expression "signature" predictive of relapse-free and overall survival in NSCLC, independent of age, gender, stage and histological characteristics [41]. In a gene expression of ten "signature" genes in early lung adenocarcinoma, a two-gene signature comprising only *ERBB3* and *BRCA1* expression was an independent risk factor in predicting survival, improving the discriminatory power of conventional classification systems [42]. Other studies also identified increased ErbB3 expression correlating with shorter survival in NSCLC [43]. Within NSCLC, *ERBB3* expression is higher in adenocarcinoma compared with squamous [44] and other forms of lung cancer [45]; and circulating *ERBB3* mRNA levels correlate with higher TNM stage and poorer survival in adenocarcinoma [46].

The mutation reported here (c.1946T>G; p.Ile649Arg) lies in a conserved transmembrane motif, key to dimerization [47]. Notably the ERBB2 variant (p.Gly660Asp) previously associated with autosomal dominant NSCLC is located in the corresponding transmembrane motif of HER2 [24]. Although germline *ERBB3* variants have been reported previously [48], pathogenic variants have been reported extremely rarely – viz., a germline *ERBB3* mutation (c.4009G>A;p.Ala1337Thr), affecting the C-terminus of the protein, was reported in association with familial erythroleukemia [49]; and a homozygous loss-of-function mutation in *ERBB3* was associated with Lethal Congenital Contractural Syndrome Type 2 (OMIM 607598) in two Israeli-Bedouin families [50].

Although it has been hypothesised that germline polymorphisms in ErbB genes would contribute to lung cancer risk [51], no such associations have been identified in GWAS of lung cancer to date (neither lung cancer overall nor individual histopathological subtype) [4-

8, 14]. Indeed, *ERBB3* has been 'relatively under-investigated' in lung cancer [51]. A very small single-candidate gene study suggested association of a variant in the *ERBB3* promoter region with lung cancer – but only with analysis restricted to a recessive model and limited to a non-smoking subset of 119 cases and 191 controls (P=0.037) [52]. Reduced ERBB3 expression was reported with the protective allele, consistent with an oncogenic role of ERBB3; however, these results have not been replicated in an independent cohort. Somatic *ERBB3* mutations, whilst common in colonic and gastric carcinomas, appear to be rare in NSCLC (Supplementary Data: Table S3). However, a study assessing CNVs in ErbB genes found that half of all lung adenocarcinomas have CNVs of *EGFR*, *ERBB2*, *ERBB3* and *ERBB4*, with higher CNV number corresponding to poorer prognosis [53].

Our germline *ERBB3* mutation is novel for NSCLC; and has not been reported (either as a somatic or germline mutation) in any other tumour type. Attributing causality to a variant segregating within a relatively small family just because of its rarity can lead to misattribution [54, 55]; hence our comprehensive functional assessment supporting this mutation as causative. Unsurprisingly, given the rarity of autosomal dominant NSCLC, no additional families were available for replication. However, our data concord with previous reports of germline mutations in ErbB family members *EGFR* (*ERBB1)* [56] and *ERBB2* [24] in other pedigrees with autosomal dominant NSCLC. Poor quality tumour DNA precluded assessment of *ERBB3* mutation(s) in our family, noting that somatic mutations were not identified in the single family with the germline *ERBB2* mutation and NSCLC [24], and inconsistently in individuals and families with *EGFR/ERBB1* mutations [56]. Although ErbB3 is expressed widely, this family has not manifested other malignancies (the proband has had non-cancerous colonic polyps). The apparent tissue specificity for malignancy is

unclear, although again is consistent with NSCLC families with *EGFR* (ERBB1) [56] and *ERBB2* [24] mutations.

Our functional data support the identified variant as tumorogenic. Normally, ErbB3 (180kDa) is expressed as a transmembrane protein dimerised with another ErbB family member; upon activation, the heterodimer is internalised via ligand-induced receptor mediated endocytosis to endosomes and subsequently trafficked to lysosomes for proteolytical degradation. Additionally, some transmembrane ErB3 is directly cleaved, forming a cytoplasmic stable and active 80kDa form, which effects are normally offset by the tumour suppressor p14ARF sequestering the 80kDa form for degradation [57]. Our results suggest that mutErbB3 is more prone to cleavage, resulting in increased amounts of the cytoplasmic 80kDa form; moreover, this increase in the 80kDa form may exceed the sequestration capacity of p14ARF. Further, the 80kDa form may independently drive proliferation, as it can increase transcription of proliferative genes without requiring activation of cytoplasmic pathways [57]. Notably, our results demonstrated increased cellular proliferation in mutErbB3-transfected cells compared to wtErbB3.

We also demonstrated that mutErbB3 co-localised with EGFR; with EGFR levels decreasing over time in mutErbB3-transfected cells, compared to wtErbB3-transfected cells. Together, these results suggest that EGFR/mutErbB3 dimers internalise and reach the late endosomal/lysosomal system faster than wtErbB3-transfected cells, consistent with more rapid signal transduction with mutErbB3. Further, activation profiles of downstream signalling pathways differed in mutErbB3-transfected cells compared to wtErbB3-transfected cells; these pathways affect transcriptional regulation of cell proliferation and migration, both critical for cancer initiation and metastasis [32].

Our results may have clinical implications beyond genetic counselling. Both germline and somatic EGFR mutations affect NSCLC responsiveness to TKIs; and HER family-targeted therapy can induce prolonged progression-free survival specifically in individuals with TK domain mutations (including NSCLC) [58]. The identified *ERBB3* mutation does not lie within this domain; thus, HER family inhibitors may not benefit this family. However, ErbB3 downregulation (e.g. by siRNA) can restore tumour responsiveness to various therapeutic approaches, including TKIs, potentially of clinical relevance for this family [58].

In conclusion, we report the first family with heritable NSCLC segregating with a germline mutation in *ERBB3*, with functional data strongly supporting this mutation as oncogenic.

Funding

This work was supported by the Cancer Council Queensland (#1041390), the Queensland Head and Neck Cancer Centre, the Princess Alexandra Research Foundation (#2016030) [FS] and a Queensland University of Technology Cancer Programme Publication Award. AML is funded by a National Health and Medical Research Council (NHMRC) Early Career Fellowship (ID 1158111). SRJ is supported by Princess Alexandra Research Foundation. AML is supported by an NHMRC ECF (#1158111). MAB was supported by a NHMRC Senior Principal Research Fellowship (ID 1024879). The Translational Research Institute was supported by a grant from the Australian Government.

Acknowledgements

We thank the family for their gracious participation. Additionally, we thank pathologist David Godbolt; research nurses Deborah Courtney, Linda Passmore, Elisabeth

McCaul; Sharon Song for technical support; David Pennisi and Karolina Slater for writing and administrative support; and Malcolm Lim. The authors acknowledge the Translational Research Institute for providing the excellent research environment and microscopy and histology core facilities.

Author Contributions:

KF, RB, IY, JS, AML, ELD and MAB established the study. KF, PI, MG and TH identified and recruited family members to the study. ELD, AML, MB, JH, LA, PL, FS, SO, SRJ, CC and HYC designed and optimized experimental approach, performed the experiments and analyzed the data. ELD, AML and FS wrote the first draft of the manuscript, with additional input from, LW, HYC, SRJ and SO. All authors critically reviewed the final manuscript.

Web resources

COSMIC: https://cancer.sanger.ac.uk/cosmic

World Health Organisation statistics: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx

Supplementary Data

References

1. World Health Organisation. *Cancer Fact Sheet*. https://www.who.int/news-room/fact-sheets/detail/cancer.
2. Mucci LA, Hjelmborg JB, Harris JR, *et al.* Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. JAMA 2016;315(1):68-76.
3. Bosse Y, Amos CI. A Decade of GWAS Results in Lung Cancer. Cancer Epidemiol Biomarkers Prev 2018;27(4):363-379.
4. EMBL-EBI. Lung Cancer GWAS loci. In; 2020.
5. EMBL-EBI. NSCLC GWAS Database. In; 2020.
6. McKay JD, Hung RJ, Han Y, *et al.* Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. Nat Genet 2017;49(7):1126-1132.
7. EMBL-EBI. Squamous Cell Carcinoma of the lung GWAS database. In; 2020.
8. EMBL-EBI. Adenocarcinoma of the lung GWAS database. In; 2020.
9. Sellers TA, Bailey-Wilson JE, Elston RC, *et al.* Evidence for mendelian inheritance in the pathogenesis of lung cancer. J Natl Cancer Inst 1990;82(15):1272-9.
10. Schwartz AG, Yang P, Swanson GM. Familial risk of lung cancer among nonsmokers and their relatives. Am J Epidemiol 1996;144(6):554-62.
11. Wu PF, Lee CH, Wang MJ, *et al.* Cancer aggregation and complex segregation analysis of families with female non-smoking lung cancer probands in Taiwan. Eur J Cancer 2004;40(2):260-6.
12. Wu AH, Fontham ET, Reynolds P, *et al.* Family history of cancer and risk of lung cancer among lifetime nonsmoking women in the United States. Am J Epidemiol 1996;143(6):535-42.

13. Lorenzo Bermejo J, Hemminki K. Familial lung cancer and aggregation of smoking habits: a simulation of the effect of shared environmental factors on the familial risk of cancer. Cancer Epidemiol Biomarkers Prev 2005;14(7):1738-40.

14. Byun J, Schwartz AG, Lusk C, *et al.* Genome-wide association study of familial lung cancer. Carcinogenesis 2018;39(9):1135-1140.

15. Fang S, Pinney SM, Bailey-Wilson JE, *et al.* Ordered subset analysis identifies loci influencing lung cancer risk on chromosomes 6q and 12q. Cancer Epidemiol Biomarkers Prev 2010;19(12):3157-66.

16. Musolf AM, Simpson CL, de Andrade M, *et al.* Parametric Linkage Analysis Identifies Five Novel Genome-Wide Significant Loci for Familial Lung Cancer. Hum Hered 2016;82(1-2):64-74.

17. Poirier JG, Brennan P, McKay JD, *et al.* Informed genome-wide association analysis with family history as a secondary phenotype identifies novel loci of lung cancer. Genet Epidemiol 2015;39(3):197-206.

18. Xiong D, Wang Y, Kupert E, *et al.* A recurrent mutation in PARK2 is associated with familial lung cancer. Am J Hum Genet 2015;96(2):301-8.

19. Seow WJ, Matsuo K, Hsiung CA, *et al.* Association between GWAS-identified lung adenocarcinoma susceptibility loci and EGFR mutations in never-smoking Asian women, and comparison with findings from Western populations. Hum Mol Genet 2017;26(2):454-465.

20. Cheng YI, Gan YC, Liu D, *et al.* Potential genetic modifiers for somatic EGFR mutation in lung cancer: a meta-analysis and literature review. BMC Cancer 2019;19(1):1068.

21. Zhou C, Yao LD. Strategies to Improve Outcomes of Patients with EGRF-Mutant Non-Small Cell Lung Cancer: Review of the Literature. J Thorac Oncol 2016;11(2):174-86.

404 22. Kobayashi S, Boggon TJ, Dayaram T, *et al.* EGFR mutation and resistance of non-
405 small-cell lung cancer to gefitinib. N Engl J Med 2005;352(8):786-92.
406 23. Gazdar A, Robinson L, Oliver D, *et al.* Hereditary lung cancer syndrome targets never
407 smokers with germline EGFR gene T790M mutations. J Thorac Oncol 2014;9(4):456-63.
408 24. Yamamoto H, Higasa K, Sakaguchi M, *et al.* Novel germline mutation in the
409 transmembrane domain of HER2 in familial lung adenocarcinomas. J Natl Cancer Inst
410 2014;106(1):djt338.
411 25. Cohen MC, Kaschula RO. Primary pulmonary tumors in childhood: a review of 31
412 years' experience and the literature. Pediatr Pulmonol 1992;14(4):222-32.
413 26. Giuseppucci C, Reusmann A, Giubergia V, *et al.* Primary lung tumors in children: 24
414 years of experience at a referral center. Pediatr Surg Int 2016;32(5):451-7.
415 27. Karczewski KJ, Francioli L, tiao G, *et al.* Variation across 141,456 human exomes
416 and genomes reveals the spectrum of loss-of-function intolerance across human protein-
417 coding genes. *bioRxiv 531210*, 2019; https://doi.org/10.1101/531210
418 https://www.biorxiv.org/content/10.1101/531210v3.
419 28. Genomes Project C, Auton A, Brooks LD, *et al.* A global reference for human genetic
420 variation. Nature 2015;526(7571):68-74.
421 29. Sherry ST, Ward MH, Kholodov M, *et al.* dbSNP: the NCBI database of genetic
422 variation. Nucleic Acids Res 2001;29(1):308-11.
423 30. Guy PM, Platko JV, Cantley LC, *et al.* Insect cell-expressed p180erbB3 possesses an
424 impaired tyrosine kinase activity. Proc Natl Acad Sci U S A 1994;91(17):8132-6.
425 31. Pillai-Kastoori L, Schutz-Geschwender AR, Harford JA. A systematic approach to
426 quantitative Western blot analysis. Anal Biochem 2020;593:113608.
427 32. Jaiswal BS, Kljavin NM, Stawiski EW, *et al.* Oncogenic ERBB3 mutations in human
428 cancers. Cancer Cell 2013;23(5):603-17.

429 33. Lesniewski LA, Hosch SE, Neels JG, *et al.* Bone marrow-specific Cap gene deletion
430 protects against high-fat diet-induced insulin resistance. Nat Med 2007;13(4):455-62.
431 34. Amaravadi R, Kimmelman AC, White E. Recent insights into the function of
432 autophagy in cancer. Genes Dev 2016;30(17):1913-30.
433 35. Mailand N, Gibbs-Seymour I, Bekker-Jensen S. Regulation of PCNA-protein
434 interactions for genome stability. Nat Rev Mol Cell Biol 2013;14(5):269-82.
435 36. De Gregoriis G, Ramos JA, Fernandes PV, *et al.* DNA repair genes PAXIP1 and
436 TP53BP1 expression is associated with breast cancer prognosis. Cancer Biol Ther
437 2017;18(6):439-449.
438 37. Adzhubei IA, Schmidt S, Peshkin L, *et al.* A method and server for predicting
439 damaging missense mutations. Nat Methods 2010;7(4):248-9.
440 38. Vaser R, Adusumalli S, Leng SN, *et al.* SIFT missense predictions for genomes. Nat
441 Protoc 2016;11(1):1-9.
442 39. Scharpenseel H, Hanssen A, Loges S, *et al.* EGFR and HER3 expression in
443 circulating tumor cells and tumor tissue from non-small cell lung cancer patients. Sci Rep
444 2019;9(1):7406.
445 40. Ma S, Jia S, Ren Y, *et al.* ErbB3 Ligand Heregulin1 Is a Major Mitogenic Factor for
446 Uncontrolled Lung Cancer Cell Proliferation. Neoplasia 2019;21(4):343-352.
447 41. Chen HY, Yu SL, Chen CH, *et al.* A five-gene signature and clinical outcome in non-
448 small-cell lung cancer. N Engl J Med 2007;356(1):11-20.
449 42. Sun Y, Hou L, Yang Y, *et al.* Two-gene signature improves the discriminatory power
450 of IASLC/ATS/ERS classification to predict the survival of patients with early-stage lung
451 adenocarcinoma. Onco Targets Ther 2016;9:4583-91.

452 43. Yi ES, Harclerode D, Gondo M, *et al*. High c-erbB-3 protein expression is associated
453 with shorter survival in advanced non-small cell lung carcinomas. Mod Pathol
454 1997;10(2):142-8.
455 44. Skrzypski M, Dziadziuszko R, Jassem E, *et al.* Main histologic types of non-small-
456 cell lung cancer differ in expression of prognosis-related genes. Clin Lung Cancer
457 2013;14(6):666-673 e2.
458 45. Kawano O, Sasaki H, Endo K, *et al.* ErbB3 mRNA expression correlated with
459 specific clinicopathologic features of Japanese lung cancers. J Surg Res 2008;146(1):43-8.
460 46. Masroor M, Javid J, Mir R, *et al.* Prognostic significance of serum ERBB3 and
461 ERBB4 mRNA in lung adenocarcinoma patients. Tumour Biol 2016;37(1):857-63.
462 47. Mendrola JM, Berger MB, King MC, *et al.* The single transmembrane domains of
463 ErbB receptors self-associate in cell membranes. J Biol Chem 2002;277(7):4704-12.
464 48. Karczewski KJ, Francioli LC, Tiao G, *et al.* The mutational constraint spectrum
465 quantified from variation in 141,456 humans. Nature 2020;581(7809):434-443.
466 49. Braunstein EM, Li R, Sobreira N, *et al.* A germline ERBB3 variant is a candidate for
467 predisposition to erythroid MDS/erythroleukemia. Leukemia 2016;30(11):2242-2245.
468 50. Narkis G, Ofir R, Manor E, *et al.* Lethal congenital contractural syndrome type 2
469 (LCCS2) is caused by a mutation in ERBB3 (Her3), a modulator of the phosphatidylinositol-
470 3-kinase/Akt pathway. Am J Hum Genet 2007;81(3):589-95.
471 51. Alaoui-Jamali MA, Morand GB, da Silva SD. ErbB polymorphisms: insights and
472 implications for response to targeted cancer therapeutics. Front Genet 2015;6:17.
473 52. Sung JS, Jin L, Jo U, *et al.* Association between -276 C/T polymorphism of the
474 ERBB3 gene and lung cancer risk in a Korean population. Anticancer Res 2012;32(10):4433-
475 7.

476 53. Chen HY, Liu CH, Chang YH, *et al.* EGFR-activating mutations, DNA copy number
477 abundance of ErbB family, and prognosis in lung adenocarcinoma. Oncotarget
478 2016;7(8):9017-25.
479 54. MacArthur DG, Manolio TA, Dimmock DP, *et al.* Guidelines for investigating
480 causality of sequence variants in human disease. Nature 2014;508(7497):469-76.
481 55. Minikel EV, Vallabh SM, Lek M, *et al.* Quantifying prion disease penetrance using
482 large population control cohorts. Sci Transl Med 2016;8(322):322ra9.
483 56. Yamamoto H, Yatabe Y, Toyooka S. Inherited lung cancer syndromes targeting never
484 smokers. Transl Lung Cancer Res 2018;7(4):498-504.
485 57. Andrique L, Fauvin D, El Maassarani M, *et al.* ErbB3(80 kDa), a nuclear variant of
486 the ErbB3 receptor, binds to the Cyclin D1 promoter to activate cell proliferation but is
487 negatively controlled by p14ARF. Cell Signal 2012;24(5):1074-85.
488 58. Verlingue L, Hollebecque A, Lacroix L, *et al.* Human epidermal receptor family
489 inhibitors in patients with ERBB3 mutated cancers: Entering the back door. Eur J Cancer
490 2018;92:1-10.
491

492 **Table 1. Characteristics of Variants Fulfilling Filtering Criteria**

Gene	Variant	GERP score	SIFT	MutationTaster	PolyPhen2
ERBB3 (*Erb-B2 Receptor Tyrosine Kinase 3*)	NM_001982: c.1946T>G: p.Ile649Arg	3.87	0.00 (deleterious)	0.995 (disease-causing)	0.121 (benign)
ATG2B (Autophagy related 2B)	NM_018036: c.4057T>A: p.Cys1353Ser	5.48	0.05 (deleterious)	0.999 (disease-causing)	0.978 (probably damaging)
SORBS1 (Sorbin and SH3 domain-containing protein 1)	NM_001034954: c.2464A>G: p.Ile822Val	6.07	0.05 (deleterious*)	0.999 (deleterious)	0.903 (possibly damaging)

493

FIGURE LEGENDS

Figure 1. Germline *ERBB3* mutation segregating with NSCLC in an affected family. **A**. Family Pedigree. **B**. Sanger sequencing chromatograms of germline DNA, demonstrating heterozygosity for *ERBB3* c.1946T>G variant (arrow) in three affected individuals and wildtype in the proband's unaffected mother.

Figure 2. ErbB3 expression, folding, response to stimulation with EGF, and effect on cellular proliferation, in HeLa cells transfected with vector-only, wtErbB3, or mutErbB3.

A. Western blot of lysates from transfected cells using two different commercial anti-ErbB3 antibodies. MutErbB3 is stably expressed and normally folded. A higher ratio of 80kDa to full length 180kDa form (arrows) is observed with mutErbB3 compared to wtErbB3.

B. Transfected HeLa cells fixed pre- (0') and post (30')-EGF stimulation and immunostained for ErbB3 (red), endogenous EGFR (purple) and nuclei stained using DAPI (blue). Right column shows merged image. Scale bars, 20µm.

Without stimulation EGFR co-localised with mutErbB3 in concentrated puncta, less evident with wtErbB3. After stimulation, both wtErbB3 and mutErbB3 increased in the perinuclear region, co-localising with EGFR.

C. Proliferation assay of transfected cells quantified and described as fold change relative to vector only (data shown as mean ± S.E.M).

MutErbB3-transfected cells showed increased rates of proliferation.

Figure 3. Analysis of protein expression and down-stream signalling pathway activation in HeLa cells transfected with vector-only, wtErbB3, or mutErbB3 constructs, before

518 **starvation (control), after 3 hours starvation (0 min) followed by EGF stimulation**
519 **(assessed at 10 min and 30 min).**

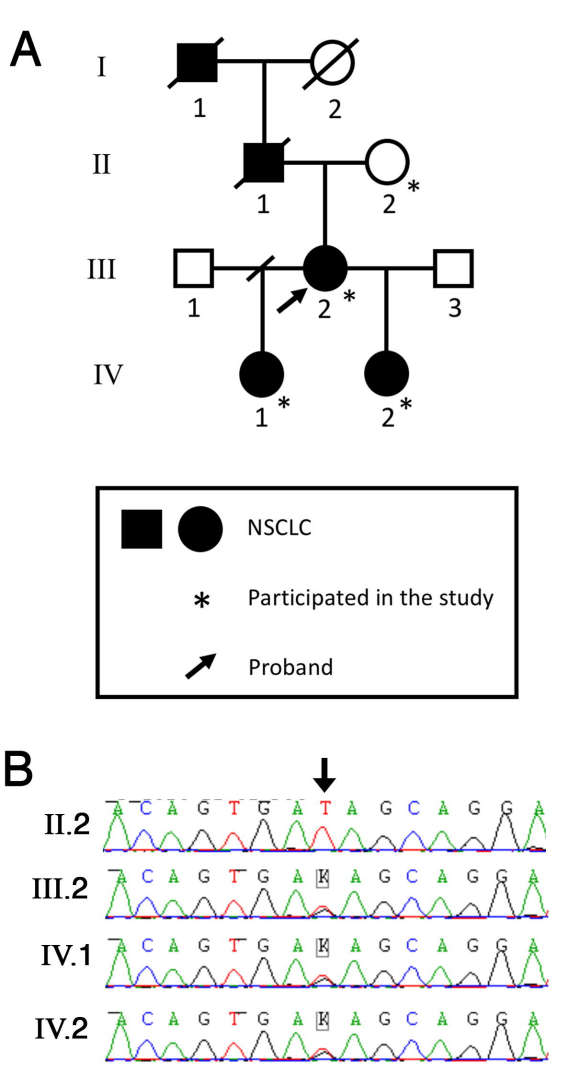
520 **A.** Western blots of ErbB3, EGFR, phospho-Akt (pAkt), total Akt (tAkt), phospho-ERK
521 (pERK) and total ERK (tERK), performed on transfected cell lysates.

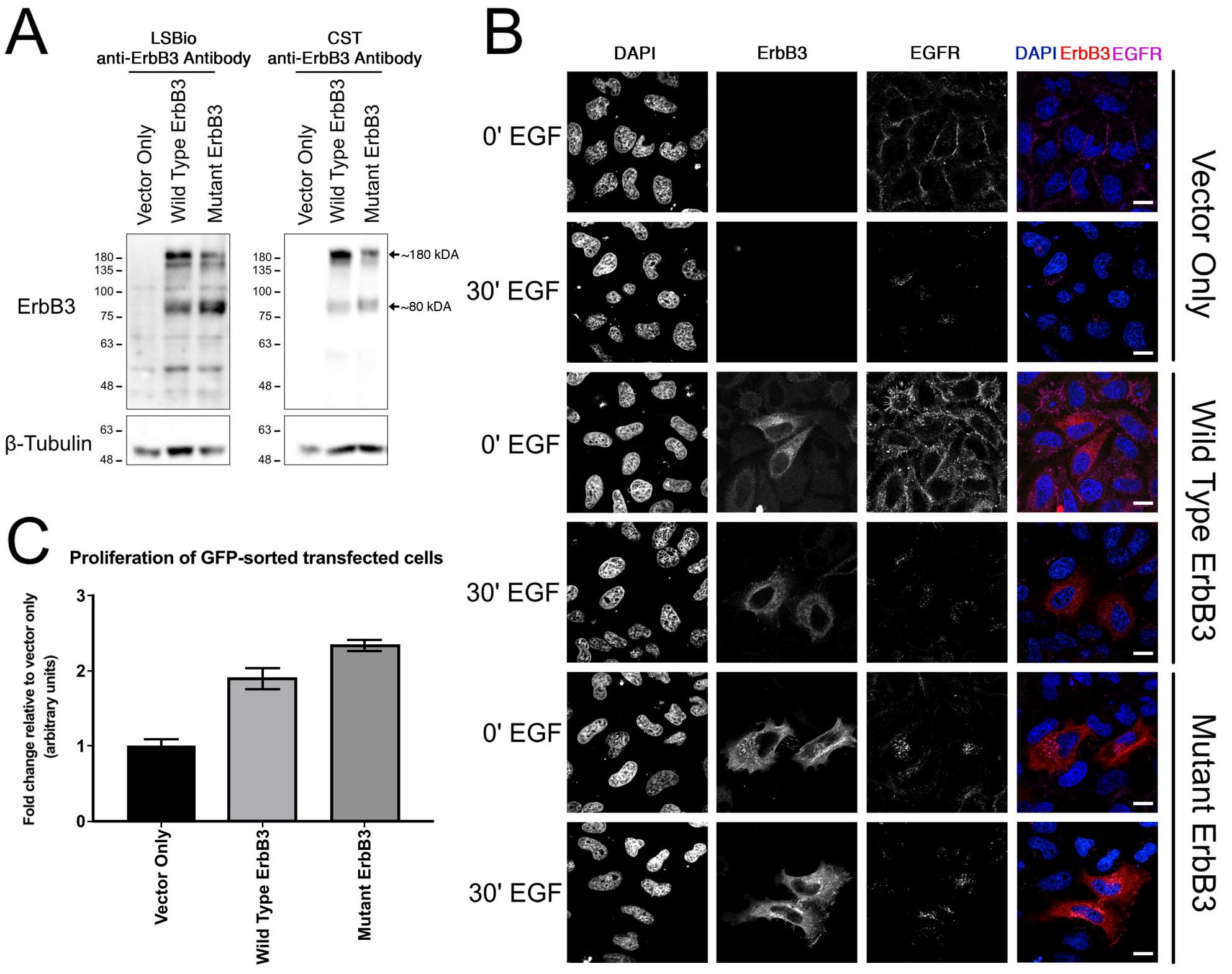
522 Following starvation and EGF stimulation, mutErbB3-transfected cells demonstrated
523 decreased EGFR levels compared with wtErbB3.

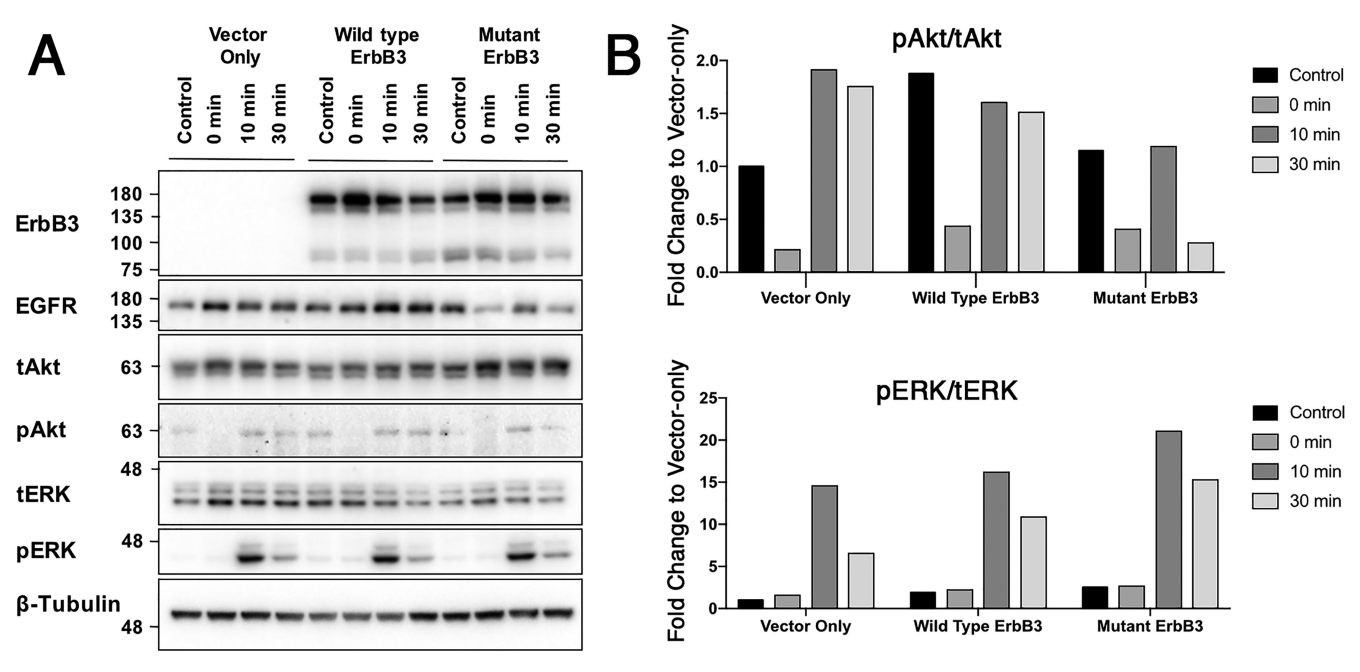
524 **B.** Ratios of phospho-Akt to total-Akt (pAkt/tAkt) (upper graphs), and phospho-ERK to total-
525 ERK (pERK/tERK) (lower graphs) quantified and normalised to β-Tubulin.

526 By 30 minutes, mutErbB3-transfected cells show decreased pAkt/tAkt ratio and increased
527 pERK/tERK ratio compared with wtErbB3-transfected cells

528 Both blot images and ratio quantification are representative of at least three separate
529 biological replicates.


A
I
1
2
II
1
2*
III
1
2*
3
IV
1*
2*
NSCLC
Participated in the study
Proband
B
II.2
III.2
IV.1
IV.2


A
LSBio
anti-ErbB3 Antibody
CST
anti-ErbB3 Antibody
Vector Only
Wild Type ErbB3
Mutant ErbB3
Vector Only
Wild Type ErbB3
Mutant ErbB3
ErbB3
β-Tubulin
180
135
100
75
63
48
63
48
63
48
~180 kDA
~80 kDA
B
DAPI
ErbB3
EGFR
DAPI ErbB3 EGFR
0' EGF
30' EGF
0' EGF
30' EGF
0' EGF
30' EGF
Vector Only
Wild Type ErbB3
Mutant ErbB3
C
Proliferation of GFP-sorted transfected cells
Fold change relative to vector only (arbitrary units)
3
2
1
0
Vector Only
Wild Type ErbB3
Mutant ErbB3


A
Vector Only
Wild type ErbB3
Mutant ErbB3
Control
0 min
10 min
30 min
ErbB3
EGFR
tAkt
pAkt
tERK
pERK
β-Tubulin
180
135
100
75
63
48
B
pAkt/tAkt
Fold Change to Vector-only
2.0
1.5
1.0
0.5
0.0
Vector Only
Wild Type ErbB3
Mutant ErbB3
Control
0 min
10 min
30 min
pERK/tERK
25
20
15
10
5
0