

DNA polymerase alpha subunit B is a binding protein for Erlotinib resistance in non-small lung cancer.

Tae Young Kim ¹, Eun Sun Ji ², Ju Yeon Lee ², Jin Young Kim ², Jong Shin Yoo ², Marcell Szasz ^{3,4}, Balazs Dome ^{4,5}, Gyorgy Marko-Varga ³ and Ho Jeong Kwon ^{1,*}

¹ Chemical Genomics Global Research Lab, Department of Biotechnology, College of Life Science & Biotechnology, Yonsei University, Seoul 120-749, Korea; kty1273@yonsei.ac.kr (T.Y.K.); kwonhj@yonsei.ac.kr (H.J.K.)

² Korea Basic Science Institute, Ochang, 28119, Korea; eunsunji@kbsi.re.kr (E.S.J.); jylee@kbsi.re.kr (J.Y.L.); jinyoung@kbsi.re.kr (J.Y.K.); jongshin@kbsi.re.kr (J.S.Y.)

³ Div. Clinical Protein Science & Imaging, Dept. of Clinical Sciences (Lund) and Dept. of Biomedical Engineering, Lund University, SE-221 84, Lund, Sweden; szaszam@gmail.com (M.S.); gyorgy.marko-varga@bme.lth.se (G.M.V.)

⁴ Hungarian Academy of Sciences, 1117 Budapest, Hungary; Department of Thoracic Surgery, National Institute of Oncology and Semmelweis University, Budapest, Hungary; szaszam@gmail.com (M.S.); balazs.dome@meduniwien.ac.at (B.D.)

⁵ Division of Thoracic Surgery, Department of Surgery, Medical University of Vienna, Vienna, Austria; balazs.dome@meduniwien.ac.at (B.D.)

* Correspondence: kwonhj@yonsei.ac.kr (H.J.K); Tel.: +82-2-2123-5883; Fax: 82-2-362-7265

Ho Jeong Kwon, Ph.D.

Professor, Department of Biotechnology, College of Life Science and Biotechnology, Yonsei University, Engineering Hall B, B410, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea

Phone: 82-2-2123-5883

Fax: 82-2-362-7265

E-mail: kwonhj@yonsei.ac.kr. Tel: +82-2-2123-5883

Supplementary information

Materials and Methods

LC-MS/MS Sample Preparation

For each sample, DTT was added to a final concentration of 10 mM, and the samples were then incubated for 30 min at room temperature, with agitation. Next, IAA was added to each sample to a final concentration of 10 mM, and each sample was then placed in the dark for 30 min at room temperature. Subsequently, protein samples were digested into peptides using trypsin; these peptides were incubated with end-over-end rotation at 37 °C for 4 h. The final solutions were dried using a SpeedVac, and the dried peptides were re-dissolved in an aqueous solution of 0.1% formic acid prior to quantification

Proteome Analysis

We referenced previous publication [8]. LC-MS/MS Analysis: IDA (information-dependent acquisition) and SWATH (sequential windowed acquisition of all theoretical fragment-ion spectra) mass spectrum techniques were used to acquire tandem MS data on a Triple TOF 6600 tandem mass spectrometer (AB Sciex, Concord, Ontario, Canada) coupled to an Ekspert nanoLC 400 system (AB Sciex). An autosampler was used to load individual 5 µL aliquots of the peptide solutions into a C18 trap column of inner diameter (i.d.) 180 µm, length 20 mm, and particle size 5 µm (Waters, Milford, MA). Peptides were desalted and concentrated on the trap column for 10 min at a flow rate of 5 µL/ min. Trapped peptides were then back-flushed and separated on a homemade microcapillary C18 column of i.d. 100 µm and length 200 mm (Reposil Pur 120, C18-AQ, particle size 3 µm) at a flow rate of 300 nL/min. The mobile phases were composed of 100%

1 water (A) and 100% acetonitrile (ACN) (B), each containing 0.1% formic acid. The liquid
2 chromatography (LC) gradient began with 5% mobile phase B and ramped to 10% mobile phase
3 B, for 1 min. 10% mobile phase B was maintained for 4 min. 10% mobile phase B was linearly
4 ramped to 40% for 90 min and then to 95% for 1 min. Then, 95% mobile phase B was maintained
5 for 13 min before descending to 5% B for another 1 min. The column was re-equilibrated with 5%
6 B for 10 min before the next run. The voltage of the nanoelectrospray was 2.0 kV. During the
7 chromatographic separation, triple time-of-flight mass spectrometry (TOF-MS) was performed
8 using either an IDA MS or a SWATH MS. In IDA mode, a TOF-MS survey scan was acquired at
9 m/z 400–2000 with 250 ms accumulation time. The 10 most intense precursor ions (2^{+} – 5^{+}) in the
10 survey scan were consecutively isolated for product ion scans. Dynamic exclusion was used for
11 45 s. Product ion spectra were accumulated for 100 ms in the mass range m/z 100–2000, with
12 rolling collision energy.

13 IDA DATA-Based SWATH Library Generation: SWATH data were acquired three times for each
14 sample. In SWATH mode, TOF-MS survey scans were acquired (m/z 400–2000, 250 ms); this was
15 followed by 34 product ion scans, with consecutive constant Q1 windows (25 Da) in the mass
16 range m/z 400–1250. Product ion spectra were accumulated for 49 ms in the mass range m/z
17 100–2000, with rolling collision energy. IDA Data-Based SWATH Library Generation. IDA MS/
18 MS data were subjected to database searches by ProteinPilot (V5.0, AB Sciex) using the Paragon
19 algorithm. The UniProt human database (August 2014 version) was used for database searches.
20 Search parameters were as follows. Sample type: identification. Cys alkylation: iodoacetamide.
21 Digestion: trypsin. Special factors: none. ID focus: allow biological modifications. The group file
22 from the database search was loaded to PeakView (V2.2 with SWATH Quantitation plug-in) and

1 exported as libraries in CSV format.

2 SWATH Peak Extraction: PeakView V2.2 with SWATH quantitation plug-in (AB Sciex) was
3 used to extract SWATH MS peak areas using the library generated in our study. PeakView uses a
4 set of processing settings to filter the ion library and determine which peptides or transitions should
5 be used for quantification. These settings include: the maximum number of peptides (50) per
6 protein to be included from the imported ion library; the number of transitions (three) per peptide;
7 a peptide confidence threshold (95%) in percentage that is used to remove the peptides with a
8 confidence below that value; a false discovery rate ($FDR < 1\%$) in percentage, which is used to
9 filter the SWATH extraction results and only export the peptide peak groups displaying a false
10 positive rate below that value; and XIC (Extracted Ion Chromatogram) retention time window (15
11 min) and mass tolerance window (50 ppm) for RT and m/z tolerance, to pick the transitions. After
12 SWATH peak extraction, the transition ion peak areas, peptide peak areas, and protein peak areas
13 were exported in Excel format for further statistical analysis. The ion peak areas exported by
14 PeakView for each sample were normalized, by total area normalization, using MarkerView 1.2.1.
15 Specifically, the sum total of all ion intensities for each sample was calculated, and the maximum
16 of those totals was determined. The ion intensities of each sample were divided by the ratio of
17 $[\text{total ion intensities}] / [\text{maximum total ion intensity}]$, ensuring that the normalized ion intensities
18 sum to the same value for all samples. Normalized protein peak areas were automatically
19 performed via PCA (principal component analysis) and then grouped by PC variance in
20 MarkerView; 25 groups were generated. These results were then exported into Excel format
21 according to peptide and protein, which we refer to henceforth as peptide peak area and protein
22 peak area.

1 Finally, the results of proteins showing significant changes in expression level were checked by
2 manual validation, and their peak areas were recalculated based solely on reliable peptide peaks
3 acquired during SWATH extraction.

4

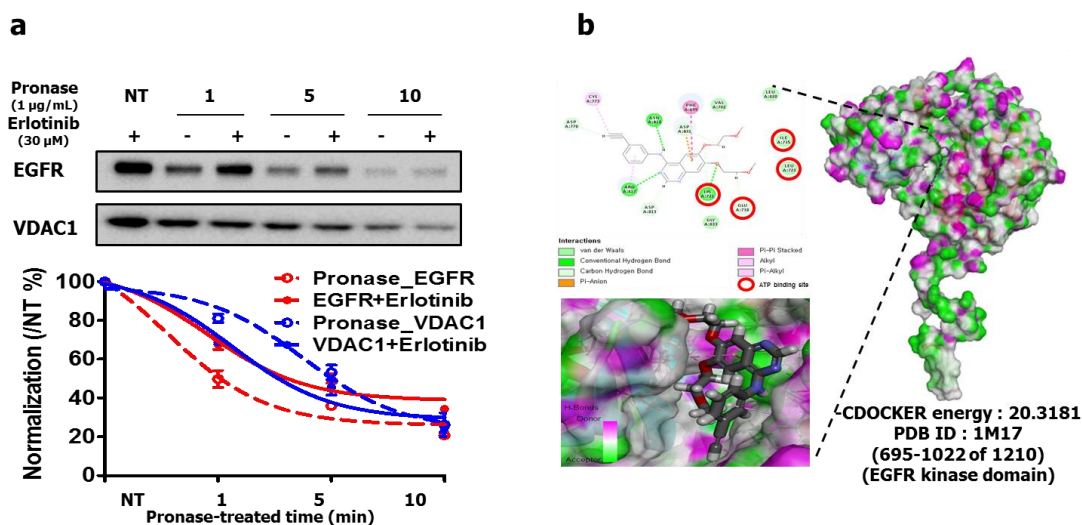


Figure 1. Characterization of EGFR as an Erlotinib target. (a) Optimal condition of DARTS assay for target validation. EGFR protein was stabilized upon Pronase treatment compared with destabilization of VDAC1 in H1299 cell lysates. The sigmoidal curve shows the quantitative data. (b) *In silico* docking model of Erlotinib interacting with EGFR (Human EGFR kinase domain, RCSB PDB ID: 1M17). Erlotinib binds to amino acid LYS 721, GLU 738, ILE 735, and LEU 723 of ATP binding site (Red circles) with van der Waals force and hydrogen bond (Up panel). Ligands are shown as bold sticks in displayed charge receptor surfaces. Bonds are shown as dashed lines color-coded as shown in that picture (Down panel)

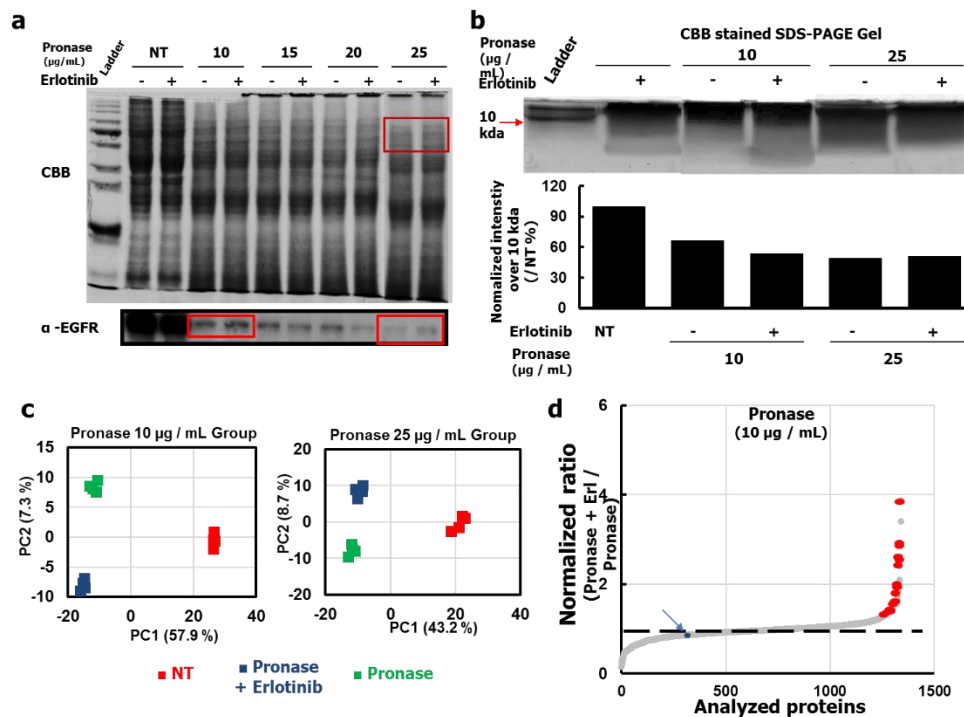


Figure 2. SDS-PAGE analysis of the H1299 proteome pool visualized using CBB staining. H1299 cells were lysed and diluted to 5 mg/mL. After dilution, these cells were treated in 30 µM Erlotinib and incubated at 4 °C for 4 h. After the drug treatment, Pronase treatment was administered at 25 °C for 15 min. After digestion with Pronase, aliquots of samples were mixed with SDS and boiled. (a) This samples were loaded in CBB stained 8% SDS-PAGE gel. 10 differential protein bands were analyzed by image analysis of Pronase dose dependent (10 ~ 25 µg/mL) with Erlotinib 30 µM treatment and confirmed EGFR expression in same sample. (b) Sample preparation for DARTS LC-MS/MS analysis. CBB stained 15% SDS-PAGE gel. It was loaded with remaining sample after demonstrating the proteome change in Figure S2a. The graph shows intensity data over 10 kda normalized by NT levels. c) The proteome status of DARTS LC-MS/MS samples measured in 4 replicate segregated into NT (Red), Pronase only (Blue), and Pronase with Erlotinib (Green) subgroup by PCA analysis. d) Identified proteins were arranged by quantitative SWATH analysis data from the DARTS sample (Pronase 10 µg/mL). Blue dot with an arrow means EGFR protein and red dots are coincided with protein in Figure 2b. Dash line is 1 of normalized ratio, which means increased quantitatively proteins by Erlotinib.

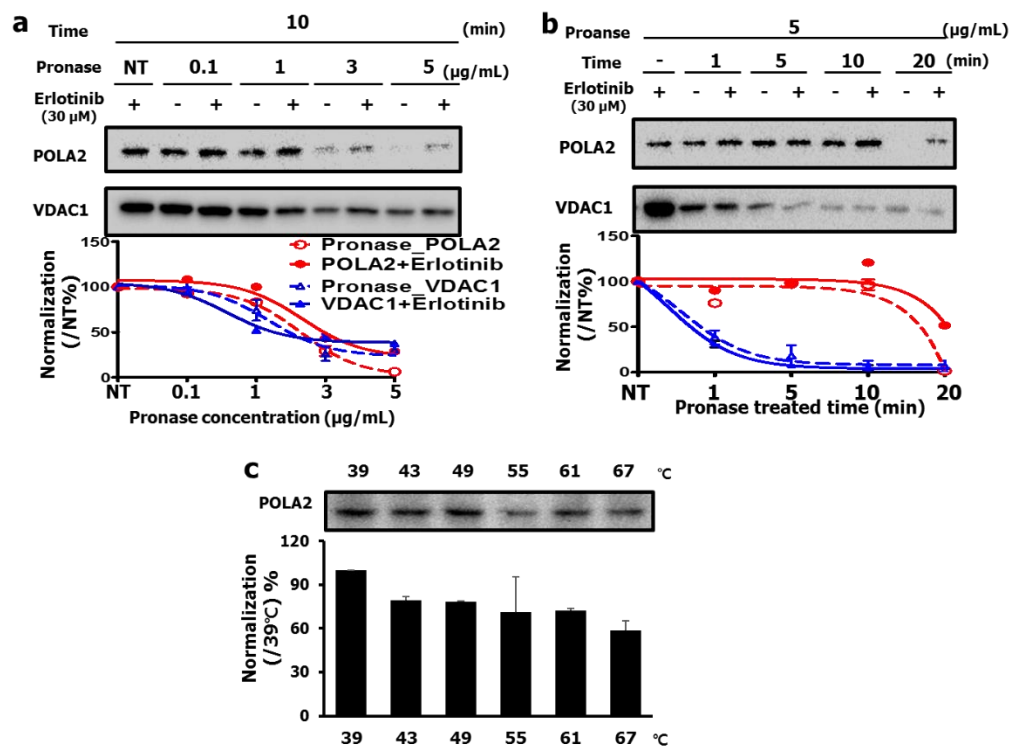


Figure 3. Characterization of POLA2 as a novel Erlotinib target and the optimal concentration. DARTS assays were performed by Pronase-dose (a) and time dependent (b). (c) The CETSA assay demonstrated the intensity of POLA2 on temperature-dependent (39 ~ 67 °C) for 5 min. More details of western blot, please view at the material & method. Values in this graph represent mean \pm SEM from three independent experiments.

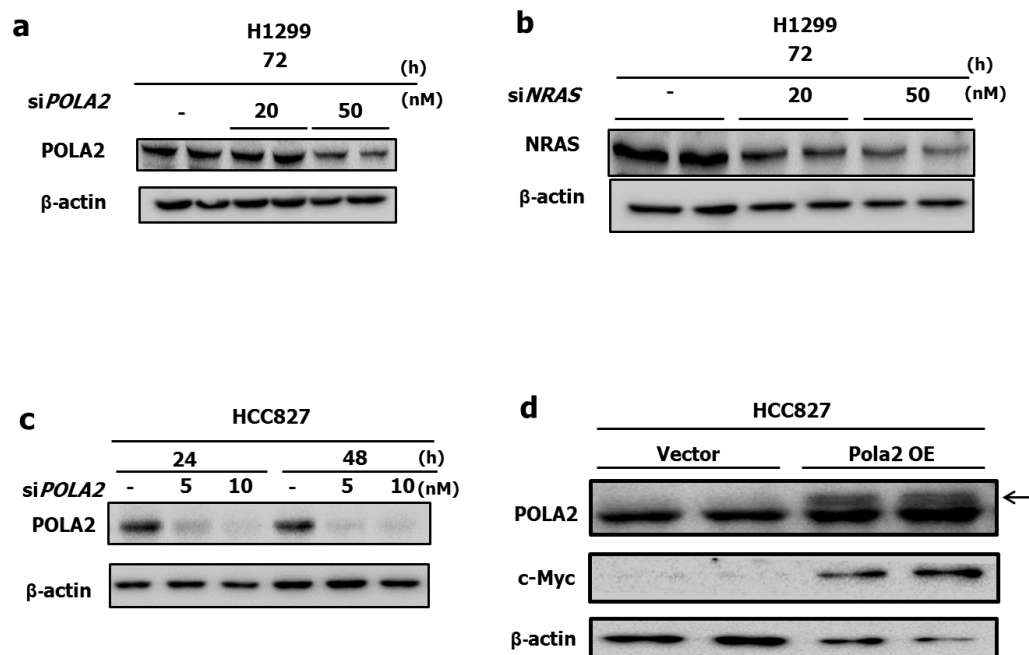


Figure 4. Genetic knockdown of POLA2 and NRAS in H1299 and HCC827 cells. (a) siPOLA2 20 and 50 nM were transfected in H1299 for 72 h. (b) siNRAS 20 and 50 nM were transfected in H1299 for 72 h. (c) Genetic knockdown of POLA2 by siRNA in HCC827 cell line at time (24, 48 h) and dose (5, 10 nM) response. (d) Overexpression of POLA2 in HCC827. Vector and POLA2 were transfected in HCC827. Indicated arrow means shifted bands by tagged POLA2 with c-myc and DDK. Expression of Beta-actin was detected to use to inner control.

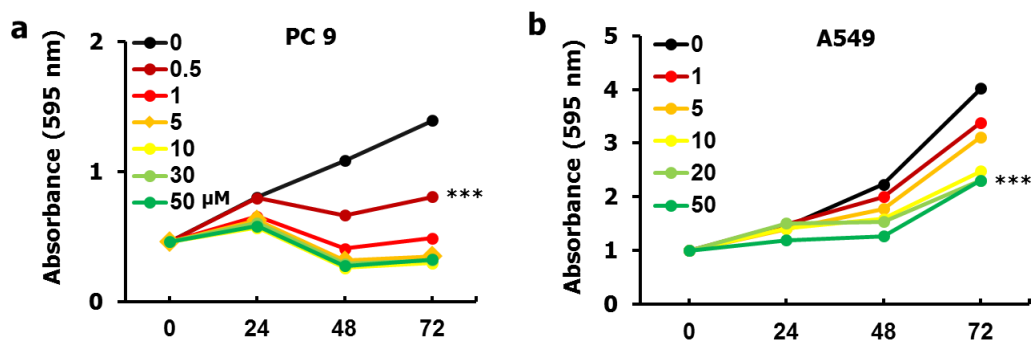


Figure 5. Proliferative effect of Erlotinib on 72-h MTT assays. NSCLC cell lines ((a) PC9 and b) A549 were measured absorbance (595 nm) by plate reader. Values in this graph represent means from three independent experiments.

1

Table 1. List of the target candidates (red box of Figure 2c)

Protein description	Relationship with Resistance
DNA polymerase alpha subunit B OS=Homo sapiens GN=POLA2 PE=1 SV=2	+++
Ephrin type-A receptor 2 OS=Homo sapiens GN=EPHA2 PE=1 SV=2	+++
RNA-binding protein 14 OS=Homo sapiens GN=RBM14 PE=1 SV=2	+++
Nuclear protein localization protein 4 homolog OS=Homo sapiens GN=NPLOC4 PE=1 SV=3	-
Heterogeneous nuclear ribonucleoprotein U-like protein 2 OS=Homo sapiens GN=HNRNPUL2 PE=1 SV=1	-
Ubiquitin-associated protein 2-like OS=Homo sapiens GN=UBAP2L PE=1 SV=2	-
Cyclin-H OS=Homo sapiens GN=CCNH PE=1 SV=1	-
Nuclear pore complex protein Nup153 OS=Homo sapiens GN=NUP153 PE=1 SV=2	-
Cytosolic Fe-S cluster assembly factor NUBP2 OS=Homo sapiens GN=NUBP2 PE=1 SV=1	-
AP-3 complex subunit beta-1 OS=Homo sapiens GN=AP3B1 PE=1 SV=3	-
Serine/threonine-protein phosphatase 2A 56 kDa regulatory subunit delta isoform OS=Homo sapiens GN=PPP2R5D PE=1 SV=2	-
Centromere/kinetochore protein zw10 homolog OS=Homo sapiens GN=ZW10 PE=1 SV=3	-
RNA-binding protein 4 OS=Homo sapiens GN=RBM4 PE=1 SV=1	-

2

3

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

Table 2. IC₅₀ values for Erlotinib on NSCLC cell lines.

Cell line	PC9	HCC827	H1299	A549
IC ₅₀ (μM)	0.3	0.2	65	50

Figure 1

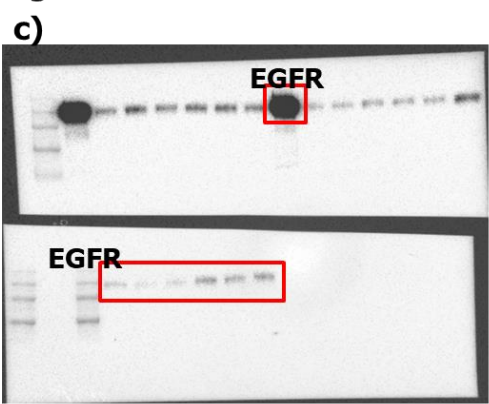


Figure 3

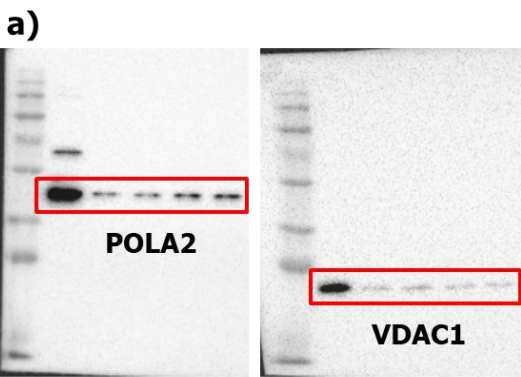


Figure 3

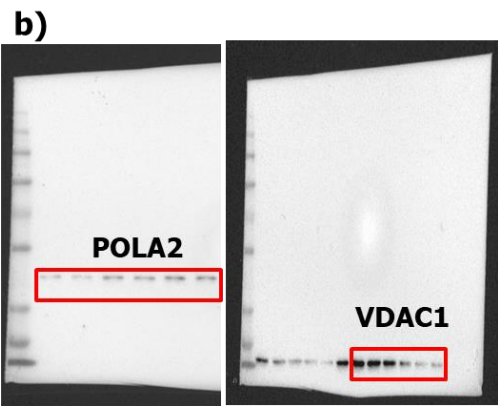
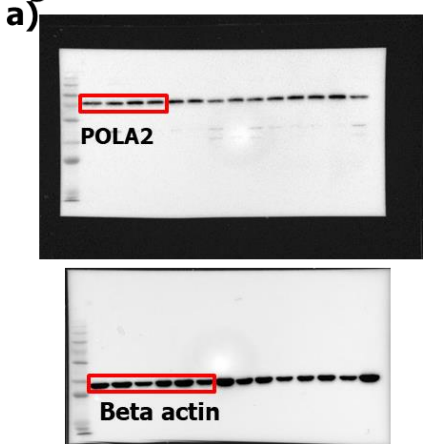


Figure 5



1
2
3
4
5
6
7

Figure S1

a)

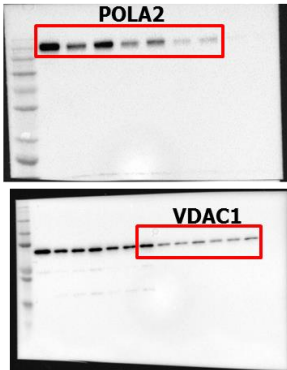


Figure S2

a)

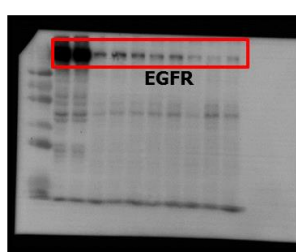


Figure S3

a)

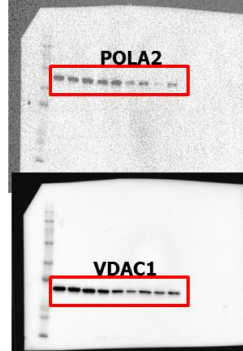


Figure S3

b)



Figure S4

a)

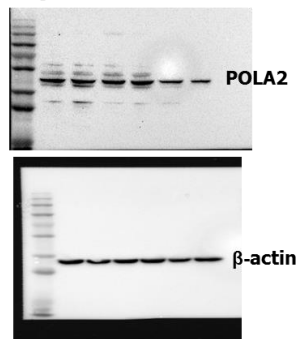


Figure S4

b)

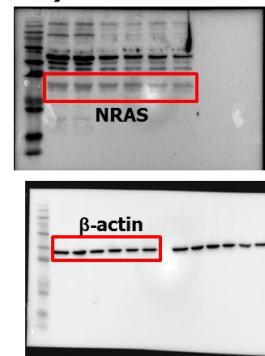


Figure S3

c)

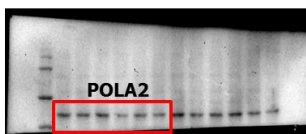


Figure S4

c)

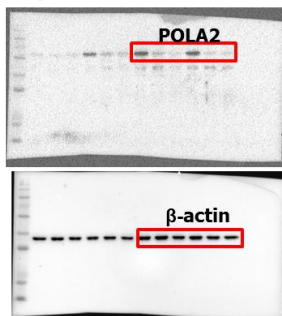


Figure S4

c)

