

Genomic Alterations and Possible Druggable Mutations in Carcinoma of Unknown Primary (CUP)

Hamidreza Abulkheyr Es

Royan Institute

Hamid Mahdizadeh

Royan Institute

Amir Abbas Hedayati Asl

Royan Institute

Mehdi Totonchi (✉ totonchimehdi@gmail.com)

Royan Institute

Research Article

Keywords: Carcinoma of unknown primary, Mutation, Copy number alterations, Targeted therapy, Personalized cancer medicine

Posted Date: May 27th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-554581/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Scientific Reports on July 23rd, 2021. See the published version at <https://doi.org/10.1038/s41598-021-94678-4>.

Abstract

Carcinoma of Unknown Primary (CUP) is a heterogeneous and metastatic disease where the primary site of origin is undetectable. Currently, chemotherapy is the only state-of-art treatment option for CUP patients. Using molecular profiling of the tumour, particularly mutation detection, offers a new treatment for CUP in a personalized fashion using targeted agents. We analyzed the mutation and copy number alterations profile of 1,709 CUP samples deposited in the AACR Project Genomics Evidence Neoplasia Information Exchange (GENIE) cohort and explored potentially druggable mutations. We identified 52 significant mutated genes (SMG) among CUP samples, in which 13 (25%) of SMG were potentially targetable with either drug is approved for the known primary tumour or undergoing clinical trials. The most variants detected were *TP53* (43%), *KRAS* (19.90%), *KMT2D* (12.60%), and *CDKN2A* (10.30%). Additionally, using pan-cancer analysis, we found similar variants of *TERT* promoter in CUP and NSCLC samples, suggesting that these mutations may serve as a diagnostic marker for identifying the primary tumour in CUP. Taken together, analyzing mutation profiling of the CUP tumours may open a new way of identifying druggable targets and consequently administrating appropriate treatment in a personalized manner.

Introduction

From all patient diagnosed with cancer, 3–5% present as metastatic carcinoma of unknown primary site (CUP)¹. It is classified as any metastatic epithelial tumour where, following extensive clinical history, physical examination, radiological studies and histopathological investigations failed to identify the primary site of tumours². Clinically, there has been no consensus and standard treatment guidelines that result in various chemotherapy regimens. Because of CUP tumour heterogeneity, the current clinical trials are difficult to perform, resulting in a poor prognosis with a median survival of less than 12 months and 5-year survival of 14%³. Thus, there is an urgent need to improve treatment modalities and prolong patients' survival with CUP⁴.

Personalized cancer medicine using genomics technologies opened new ways to treat various types of cancers using the identification of targetable mutations^{5–9}. Recent studies have highlighted the crucial role of precision medicine in patient stratification and the selection of effective treatment in malignant types of cancer^{10,11}. Moreover, several studies have demonstrated improved overall survival in patients with advanced and metastatic cancers who have received genetically-matched targeted therapies^{12,13}. In CUP tumours, the implementation of this approach may improve treatment by targeting tumour-specific and druggable somatic variants in a personalized manner³. The AACR Project Genomics Evidence Neoplasia Information Exchange (GENIE) has recently collected the genomic information including mutations and copy number variation of the wide range of solid tumours including CUP from both primary and metastatic tumours^{14–16}. Using these public data, we analyzed the genomic mutations and copy number alterations of 1,709 CUP samples to provide insight into the genetic makeup of these tumours and determined potentially druggable targets.

Results

Clinical characteristic of samples

In total, 45,048 samples across 17 cancer types, including CUP were included in this study. The sample type distribution was 24,567 primary and 15,484 metastasis tumours in GENIE cohorts. The hotspot regional mutations and copy number variations of these samples were available from GENIE and cBioportal. According to the information provided by GENIE, we divided samples into 17 broader cancer types, including CUP samples (Fig. 1A). The cancer categories containing the most samples were non-small cell lung cancer (9,085 (15.3%)), breast invasive ductal carcinoma (8,712 (14.7%)), colorectal cancer (5,961 (10.0%)), Glioma (3,214 (5.4%)), Melanoma (2,492 (4.2%)), prostate cancer (2,214 (3.7%)). The number of CUP samples registered in this cohort was 1709 (2.9%), dividing to 1222 metastatic (71.5%), 288 primaries (16.9%), 182 (10.6%) not applicable or heme and 17 (1.0%) unspecified (Fig. 1B). For gender information among CUP patients, 50.5 % of patients were female and 49.5% were male (Fig. 1C).

Significantly mutated genes (SMG) in CUP samples

We analyzed the most genomic mutations of hotspot regions at the gene level in CUP samples in GENIE according to the previously developed method ^{8,9}. In total, 52 SMG was identified (Fig. 2A, Supplementary Table 1). Among SMGs, the mutation rate of *TP53*, *KRAS*, *ARID1A*, *SMARCA4* and *KMT2D* were recorded significantly higher than other identified SMGs (Fig. 2B, Supplementary Table 1). The pathway enrichment analysis of identified SMGs resulted to the involvement of SMGs in a wide range of cellular processes, (Fig. 2C, Supplementary Table 2), including transcription factors/regulators, receptor tyrosine kinase signalling, cell cycle, IGF pathway-protein kinase B signalling, phosphatidylinositol-3-OH kinase (PI(3)K) signalling, Wnt/β-catenin signalling, *PDGF*, *FGF*, *EGF*, *TGF-beta*, and Notch signalling pathways and integrin signalling pathway. The identification of *MAPK*, *PI(3)K* and Wnt/β-catenin signalling pathways is consistent with classical cancer studies. Notably, almost all samples had at least one non-synonymous mutation in at least one SMG. The average number of point mutations in these genes varies across samples, with the highest (512 mutations for *TP53* across 727 cases) and the lowest (15 mutations for *GLI3* across 15 cases) (Fig. 2B. Supplementary table 1). This suggests that the numbers of both cancer-related genes (52 identified in this study) and cooperating driver mutations required during oncogenesis are few, although large-scale structural rearrangements were not included in this analysis. Interestingly, in line with the previous study performed by Zehir et al. 2017 ⁸ highlighting *TERT* promoter mutations across few primary tumours, we observed similar mutation of *TERT* promoter among CUP samples (n=91) (Fig. 2D). Although the clinical relevance of mutations in the *TERT* promoter remains incompletely understood, our results reaffirm the high prevalence of these alterations in patients with advanced solid tumours and suggest an association with disease progression and poor outcome. Additionally, the presence of similar mutation of *TERT* promoter in CUP and NSCLC samples suggests these mutations may serve as a diagnostic marker for identification of primary tumour in CUP patients.

Mutual exclusivity and co-occurrence among SMGs

The 1,035 pair-wise exclusivity and co-occurrence analysis of the 52 SMGs found 198 mutually exclusive (P-value < 0.001) and 837 co-occurring (P-value < 0.001) pairs (Fig. 3 and supplementary Table 3) among cup samples. Pairs with significant exclusivity were include *KRAS* and *FAT1*, *KRAS* and *NOTCH3*, *KRAS* and *NF1*, *KRAS* and *DMD* and *CDKN2A* and *RB1* in CUP samples. Additionally, the cohort analysis identified pairs with significant co-occurrence, including *KRAS* and *APC*, *TP53* and *APC*, *KRAS* and *CDKN2A*, *KRAS* and *STK11*, *KRAS KEAP1*, and *SMARCA4* and *KEAP1* highlighting the importance of this oncogene in cup tumours.

Copy number alteration among cup samples

The copy number variation differences within cup samples resulted into identification of 624 frequently amplified/deleted regions. Significant amplification of *MYC*, *FGF4* and *FGF19* observed in a small fraction of patients (Fig. 4A) while deletion of cell cycle-related genes *CDKN2B* and *CDKN2A* were detected in only 10 and 20 percent of patients respectively (Fig. 4A). Further, we analyzed copy number alteration of the CUP-SMGs within CUP samples (Fig. 4B) and across primary tumours of 14 cancer types registered in GENIE (Fig. 4C, Supplementary Table 4). Among CUP samples, a deep deletion of *TP53*, *RB1*, *CDKN2A*, and *STK11* and amplification of *KRAS* and *PIK3CA* were observed. In a pan-cancer analysis, amplification of *KRAS* and *PIK3CA* in the breast (66 and 114 of cases) and non-small cell lung cancer (46 and 48 of cases), *TERT* in non-small cell lung cancer (114 of cases) and *ATR* in breast cancer (36 of cases), were the most amplified genes, while deletion in *CDKN2A* in glioma (676 of cases), *RB1* and *TP53* in small cell lung cancer (15 of cases) were observed in these 14 different cancer types (Fig. 4C). Among these genes with significantly altered copy numbers between CUP and primary tumours, a significant amplification of *TERT* promoter was observed in both CUP and non-small cell lung cancer samples compared to glioma and breast primary tumours suggesting that copy number variation of *TERT* may play diagnostic role for identification of origin of CUP tumours (Fig. 4D).

Mutation frequency of CUP-SMGs across 17 known primary tumours

To identify similar and targetable mutation pattern in CUP, we analyzed and compared genomic alteration frequency of identified CUP-SMGs in primary tumour types across 17 cancer types in GENIE (Fig. 5A). The majority of CUP-SMGs mutations were enriched in non-small cell lung cancer (4,221 cases) colon cancer (4,011 cases) and breast cancer (3,376 cases) (Fig. 5A).

The most frequently mutated gene in this cohort was *TP53* (44% of total samples) (Fig. 5B). Its mutations predominate in non-small cell lung cancer (46.36%, 2,517 cases), colon cancer (65.55%, 2,365 cases) and breast cancer (36.26%, 2,060 cases) (Fig. 5B). *KRAS* is the second most commonly mutated genes, occurring frequently (>10%) in most cancer types (pancreatic:74.6%, colon cancer:44.24%, non-small cell lung cancer:30.93%) except hepatobiliary carcinoma, cervical cancer, bladder cancer, thyroid cancer, melanoma, small-cell lung cancer, head and neck carcinoma, prostate and breast cancer (Fig. 5B). *PIK3CA* mutations was frequented in breast cancer (36.7%) and cervical cancer (25.14%), being

specifically enriched in luminal subtype tumours. Many cancer types carried mutations in chromatin remodelling genes. In particular, histone-lysine N-methyltransferase genes *KMT2D*, *KMT2C* and *KMT2B* in bladder, lung and endometrial cancers, whereas the *KMT2A* is mostly mutated in non-small cell lung cancer and colon cancer. Mutations in *ARID1A* are frequent in non-small cell lung cancer, colon cancer, bladder cancer and breast cancer, whereas mutations in *KEAP1* and *STK11* was predominate in non-small cell lung cancer (8.62% and 11.75% respectively) (Fig. 5B). *KRAS* mutations are typically mutually exclusive, with recurrent activating mutations (*KRAS*(Gly 12) and *KRAS*(Gly 13) common in colon cancer, non-small cell lung cancer and pancreatic cancer. We compared the most common hotspot mutations in *KRAS* between CUP and other *KRAS* mutation enriched cancer types (Fig. 5C). Comparing hotspot mutations resulted to enrichment of G12D and G12R in pancreatic cancer, G12C, G12F and G13C in non-small cell lung cancer and CUP samples. These data highlight similarity of *KRAS* hotspot mutations between CUP and NSCLC.

Targetable mutations and drug candidates

To identify or predict possible therapeutics based on genomic alterations identified in SMG in CUP samples, we performed a gene-drug association analysis using PanDrugs platforms¹⁷. The gene-drug associations classified into two groups called “Drug targets” in which drugs can directly target genes that contribute in disease phenotype, and “Biomarkers” where genes are representing a drug-response associated status while its protein products are not targetable¹⁷. From 262 identified interactions, 8.7 % (23/262) was classified as a direct drug target, while 91.3 % (239/262) of gene-drug interactions identified as Biomarker (Fig. 5D). Interestingly, we found five FDA approved drugs, Crizotinib (GScore: 0.76. Dscore: 0.95) and Copanlisib (GScore: 0.76. Dscore: 0.92), Debrafenib, Sorafenib, Vemurafenib, and Regorafenib as best candidate for targeting ALK/MET, PIK3CA, and BRAF inhibitors respectively (Fig. 5D. Supplementary Table 5). Moreover, various off-label and clinically investigating compounds for targeting mutated KRAS were identified, although the GScore and DScore of these compounds did not reach a high score (Supplementary Table-5). Everolimus (mTOR inhibitor), Bortezomib (26S proteasome inhibitor), and Pemetrexed (chemotherapy agent), were identified with the highest GScore and DScore compared to the other drugs candidates in this group (Fig. 5D. Supplementary Table-5). Taken together, these data highlight presence of at least one druggable variants and potential of using genomic alteration guided targeted therapy in CUP patients.

Discussion

Currently, combination chemotherapy regimens have been considered as first-line of therapy for CUP patients¹⁸. Personalized cancer therapy using the identification of druggable mutations has encouraged mutational profiling of various types of tumours including metastasis tumours, including CUP¹⁹⁻²¹. This study analyzed the most significant mutated genes and identified the most prevalence variants in 1709 CUP samples. The gene-drug association studies suggested that at least one of the identified variants is linked to the known oncogenic driver mutations and approved targeted therapy or therapeutics are currently in clinical trial studies highlighting the potential of genomic alteration-based treatment

approach for a patient with CUP. In line with this concept, numerous clinical studies have been reported durable treatment responses using mutation-matched targeted therapies drugs, including EGFR, BRAF, KIT, and MET^{12,22-24}.

Currently, targeted therapy agents Crizotinib and Copanlisib approved for the treatment of tumours that harbour mutations in ROS1/MET/ALK and PIK3CA, while therapeutic agents for the other identified variants including FGFR family, MYC, MET, and KRAS are currently being investigated in active and ongoing clinical trials. A large proportion of the mutations detected in this study are known to be associated with various signal transduction pathways, apoptotic regulation, cell cycle progression, and receptor tyrosine kinase signalling regulations. These results can be promising because the majority of available targeted drugs act through targeting one of these pathways, which are commonly altered in various types of known tumours²⁵⁻²⁹. The most mutated gene identified in this study was TP53 (43%, 743/1709) with numerous different non-synonymous coding region variants. Similar to these data, previous studies demonstrated the association of TP53 mutations in metastatic progression in multiple cancer types, supporting the presence of high mutation load on TP53 reported in CUP^{30,31}.

Other common variants detected in this cohort were observed in genes involved in the activation and regulation of key signal transduction pathways, including BRAF and KRAS. This is the first study to report various codon 12 variants of KRAS in CUP samples. The detection of codon 12 mutations in this cohort is consistent with the highly aggressive behaviours of CUP tumours^{19,23}. Furthermore, characterizing the mutational status of KRAS has become clinically relevant in some malignancies, because the presence of a KRAS mutation is known to stimulate resistance to some tyrosine kinase inhibitors³²⁻³⁵. Although, currently no approved therapeutic agent to target and inhibit mutant KRAS activity available, however, recent clinical studies reported a partial response in CUP patients with a KRAS(G12D) mutation treated with trametinib (MEK inhibitor)^{24,36}. In this study, we also observed KRAS(G12C) variant in 25% of CUP samples. Recent promising results from Sotorasib (AMG-510), a specific covalent inhibitor of K-RAS(G12C) in NSCLC suggest the detection of these variants of KRAS as a possible druggable target in CUP patients³⁷. Moreover, targeting KRAS(G12C) using Sotorasib in advanced solid tumours showed an encouraging anticancer activity which might be useful in CUP³⁸.

Similar to other studies, we also identified an activating BRAF V600E mutations in 4.3% (74/1709) cases¹⁸⁻²⁰. This offers the potential of using BRAF inhibitors such as Vemurafenib and Dabrafenib for CUP with BRAF (V600E) mutations. In line with these, through the gene-drug association analysis, we also observed a high GScore and DScore of BRAF inhibitors Dabrafenib and Vemurafenib for targeting V600E variant identified in CUP samples. Moreover, a clinical study showed a complete clinical response of CUP patients treated with BRAF(V600E) targeted therapy Vemurafenib in combination with immunotherapy agent Ipilimumab³⁹.

Mutations in MET and ERBB2 (HER2) were detected in 30 and 55 of cases respectively, suggesting the possibility of targeting these receptor tyrosine kinases²². Targeting MET using Crizotinib combined with

HER2 inhibitor Trastuzumab has been shown with success in CUP patients. The current success of HER2 and MET targeted therapies using Trastuzumab and Crizotinib in a combination manner in advanced and metastatic tumours including HER2- and MET-mutant CUP tumour suggest the future evaluation of these genes as druggable targets in patients with CUP⁴⁰. Our results support those of other CUP studies highlighting the value of sequencing techniques, particularly gene mutation detection, to identify actionable targets^{10,18-21}.

Taken together, these data highlight the molecular heterogeneity of CUP tumours. The mutations detected across the majority of CUP cases included in this study highlights not only the genomic instability present in these tumours but also potential application of targeted therapies for a significant proportion of patients with CUP which might improve the prognosis and therapeutic decisions for the patient with CUP¹¹.

Material And Methods

Data collection

GENIE v5.0 provided the mutation, copy number variation, gene fusion and clinical information of 59,442 tumour samples¹⁵. Most onco-types were classified into 17 categories according to Oncotree (<http://oncotree.mskcc.org/oncotree/>). The onco-types not included in these 17 categories were excluded from our analysis. Raw data were downloaded from Synapse (syn17112456, <https://www.synapse.org/>) and provided by the GENIE project using either R commands or cBioportal (<https://genie.cbioperl.org/>)^{41,42}. The preprocessing protocols for these data are described in the GENIE-provided data guide.

Significantly mutated genes (SMG) analysis

The SMG analysis performed according to the previously developed criteria and protocols^{20,21}. We used the MuSiC suite⁴³ to identify significant genes for CUP samples and also for Pan-Cancer tumours according. This test assigns mutations to seven categories: AT transition, AT transversion, CG transition, CG transversion, CpG transition, CpG transversion and indel, and then uses statistical methods based on convolution, the hypergeometric distribution (Fisher's test P-value < 0.05), and likelihood to combine the category-specific binomials to obtain overall P-values. Notably, the genes with a cohort level alteration frequency of ≥ 5% or a tumour type-specific alteration frequency of ≥ 30% were included in our analysis, while tumours having no mutation, or more than 500 mutations were excluded in this study. Differentially mutated sites were plotted using Mutation-Mapper module in cBioportal. (http://www.cbioperl.org/mutation_mapper.jsp).

Copy number variation analysis

Copy number alteration data were available at AACR Project GENIE, in cBioportal. In the present study, we selected the 17 most common cancer types for comparing their copy number variation frequencies with

CUP samples. We calculated the changes in the average frequency of copy number variation (amplification and deletion) of CUP and Pan-cancer samples using provided R code in cbioportal.

Mutual exclusivity and co-occurrence analysis

We used Fisher's exact test to identify pairs of SMGs with significant (P-value < 0.001 by Benjamini–Hochberg) exclusivity and co-occurrence. We identified significant pairs by analyzing all CUP samples together. Then we used a de novo driver exclusivity algorithm known as Dendrix⁴⁴ to identify sets of approximately mutually exclusive mutations on all samples together. The plotting for mutual exclusivity and co-occurrence was performed using Gitools software (version 2.3.1)⁴⁵.

Declarations

Data Availability

The genomic data from the GENIE dataset used in this study are openly available for download in <https://www.synapse.org>, reference number [syn17112456]¹⁵.

Acknowledgements

The authors would like to thank GENIE for data collection, as well as cBioPortal (<http://www.cbioportal.org>, version v3.2.11) for the provision of data processing and customizable functions.

Author Contributions

HAE participated in the study design and performed data extraction and analysis. He wrote the first draft of the manuscript additionally. HM participated in the study design and critically edited the manuscript. AHA participated as clinical oncologists and scientific advisor in study design, MT also participated in the study design and editing of the manuscript and he has supported the project financially.

Funding

This work was supported in part by a grant from Royan Institute, Tehran, Iran, Royan TuCAGene Ltd. Tehran, Iran, and School of Biological Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran.

Availability of data and materials

All data generated and described in this article are available from the corresponding web servers and portal and are freely available to download for noncommercial purposes, without breaching participant confidentiality.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable

Competing Interests

The authors declare that they have no competing interests.

References

1. Greco, F. A. Cancer of unknown primary site: still an entity, a biological mystery and a metastatic model. *Nat Rev Cancer* **14**, 3-4, doi:10.1038/nrc3646 (2014).
2. Oien, K. A. & Dennis, J. L. Diagnostic work-up of carcinoma of unknown primary: from immunohistochemistry to molecular profiling. *Ann Oncol* **23 Suppl 10**, x271-277, doi:10.1093/annonc/mds357 (2012).
3. Moran, S., Martinez-Cardus, A., Boussios, S. & Esteller, M. Precision medicine based on epigenomics: the paradigm of carcinoma of unknown primary. *Nat Rev Clin Oncol* **14**, 682-694, doi:10.1038/nrclinonc.2017.97 (2017).
4. Hainsworth, J. D. & Greco, F. A. Cancer of Unknown Primary Site: New Treatment Paradigms in the Era of Precision Medicine. *Am Soc Clin Oncol Educ Book* **38**, 20-25, doi:10.1200/EDBK_100014 (2018).
5. Frampton, G. M. *et al.* Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat Biotechnol* **31**, 1023-1031, doi:10.1038/nbt.2696 (2013).
6. Campbell, J. D. *et al.* Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet* **48**, 607-616, doi:10.1038/ng.3564 (2016).
7. Hoadley, K. A. *et al.* Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. *Cell* **173**, 291-304 e296, doi:10.1016/j.cell.2018.03.022 (2018).
8. Zehir, A. *et al.* Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med* **23**, 703-713, doi:10.1038/nm.4333 (2017).
9. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333-339, doi:10.1038/nature12634 (2013).
10. Varghese, A. M. *et al.* Clinical and molecular characterization of patients with cancer of unknown primary in the modern era. *Ann Oncol* **28**, 3015-3021, doi:10.1093/annonc/mdx545 (2017).
11. Rassy, E. & Pavlidis, N. Progress in refining the clinical management of cancer of unknown primary in the molecular era. *Nat Rev Clin Oncol* **17**, 541-554, doi:10.1038/s41571-020-0359-1 (2020).
12. Tan, D. S. *et al.* Molecular profiling for druggable genetic abnormalities in carcinoma of unknown primary. *J Clin Oncol* **31**, e237-239, doi:10.1200/JCO.2012.44.3937 (2013).

13. Robinson, D. R. *et al.* Integrative clinical genomics of metastatic cancer. *Nature***548**, 297-303, doi:10.1038/nature23306 (2017).
14. Micheel, C. M. *et al.* American Association for Cancer Research Project Genomics Evidence Neoplasia Information Exchange: From Inception to First Data Release and Beyond-Lessons Learned and Member Institutions' Perspectives. *JCO Clin Cancer Inform***2**, 1-14, doi:10.1200/CCI.17.00083 (2018).
15. Consortium, A. P. G. AACR Project GENIE: Powering Precision Medicine through an International Consortium. *Cancer Discov***7**, 818-831, doi:10.1158/2159-8290.CD-17-0151 (2017).
16. Litchfield, K., Turajlic, S. & Swanton, C. The GENIE Is Out of the Bottle: Landmark Cancer Genomics Dataset Released. *Cancer Discov***7**, 796-798, doi:10.1158/2159-8290.CD-17-0604 (2017).
17. Pineiro-Yanez, E. *et al.* PanDrugs: a novel method to prioritize anticancer drug treatments according to individual genomic data. *Genome Med***10**, 41, doi:10.1186/s13073-018-0546-1 (2018).
18. Gatalica, Z. *et al.* Comprehensive tumor profiling identifies numerous biomarkers of drug response in cancers of unknown primary site: analysis of 1806 cases. *Oncotarget***5**, 12440-12447, doi:10.18632/oncotarget.2574 (2014).
19. Ross, J. S. *et al.* Comprehensive Genomic Profiling of Carcinoma of Unknown Primary Site: New Routes to Targeted Therapies. *JAMA Oncol***1**, 40-49, doi:10.1001/jamaoncol.2014.216 (2015).
20. Loffler, H. *et al.* Molecular driver alterations and their clinical relevance in cancer of unknown primary site. *Oncotarget***7**, 44322-44329, doi:10.18632/oncotarget.10035 (2016).
21. Tothill, R. W. *et al.* Massively-parallel sequencing assists the diagnosis and guided treatment of cancers of unknown primary. *J Pathol***231**, 413-423, doi:10.1002/path.4251 (2013).
22. Stella, G. M. *et al.* MET mutations in cancers of unknown primary origin (CUPs). *Hum Mutat***32**, 44-50, doi:10.1002/humu.21374 (2011).
23. Palma, N. A. *et al.* Durable Response to Crizotinib in a MET-Amplified, KRAS-Mutated Carcinoma of Unknown Primary. *Case Rep Oncol***7**, 503-508, doi:10.1159/000365326 (2014).
24. Kato, S. *et al.* Utility of Genomic Analysis In Circulating Tumor DNA from Patients with Carcinoma of Unknown Primary. *Cancer Res***77**, 4238-4246, doi:10.1158/0008-5472.CAN-17-0628 (2017).
25. Holohan, C., Van Schaeybroeck, S., Longley, D. B. & Johnston, P. G. Cancer drug resistance: an evolving paradigm. *Nat Rev Cancer***13**, 714-726, doi:10.1038/nrc3599 (2013).
26. Ramirez, M. *et al.* Diverse drug-resistance mechanisms can emerge from drug-tolerant cancer persister cells. *Nat Commun***7**, 10690, doi:10.1038/ncomms10690 (2016).
27. Gottesman, M. M. Mechanisms of cancer drug resistance. *Annu Rev Med***53**, 615-627, doi:10.1146/annurev.med.53.082901.103929 (2002).
28. Yuan, T. L. & Cantley, L. C. PI3K pathway alterations in cancer: variations on a theme. *Oncogene***27**, 5497-5510, doi:10.1038/onc.2008.245 (2008).
29. Roberts, P. J. & Der, C. J. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene***26**, 3291-3310, doi:10.1038/sj.onc.1210422 (2007).

30. Powell, E., Piwnica-Worms, D. & Piwnica-Worms, H. Contribution of p53 to metastasis. *Cancer Discov***4**, 405-414, doi:10.1158/2159-8290.CD-13-0136 (2014).
31. Tang, Q., Su, Z., Gu, W. & Rustgi, A. K. Mutant p53 on the Path to Metastasis. *Trends Cancer***6**, 62-73, doi:10.1016/j.trecan.2019.11.004 (2020).
32. Pao, W. *et al.* KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS Med***2**, e17, doi:10.1371/journal.pmed.0020017 (2005).
33. Del Re, M. *et al.* Contribution of KRAS mutations and c.2369C > T (p.T790M) EGFR to acquired resistance to EGFR-TKIs in EGFR mutant NSCLC: a study on circulating tumor DNA. *Oncotarget***8**, 13611-13619, doi:10.18632/oncotarget.6957 (2017).
34. Ohashi, K. *et al.* Lung cancers with acquired resistance to EGFR inhibitors occasionally harbor BRAF gene mutations but lack mutations in KRAS, NRAS, or MEK1. *Proc Natl Acad Sci U S A***109**, E2127-2133, doi:10.1073/pnas.1203530109 (2012).
35. Misale, S. *et al.* Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature***486**, 532-536, doi:10.1038/nature11156 (2012).
36. Ross, S. J. *et al.* Targeting KRAS-dependent tumors with AZD4785, a high-affinity therapeutic antisense oligonucleotide inhibitor of KRAS. *Sci Transl Med***9**, doi:10.1126/scitranslmed.aal5253 (2017).
37. Canon, J. *et al.* The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature***575**, 217-223, doi:10.1038/s41586-019-1694-1 (2019).
38. Hong, D. S. *et al.* KRAS(G12C) Inhibition with Sotorasib in Advanced Solid Tumors. *N Engl J Med***383**, 1207-1217, doi:10.1056/NEJMoa1917239 (2020).
39. Roe, O. D. & Wahl, S. G. The undifferentiated carcinoma that became a melanoma: Re-biopsy of a cancer of an unknown primary site: a case report. *J Med Case Rep***11**, 82, doi:10.1186/s13256-017-1238-y (2017).
40. Clynick, B. *et al.* Genetic characterisation of molecular targets in carcinoma of unknown primary. *J Transl Med***16**, 185, doi:10.1186/s12967-018-1564-x (2018).
41. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov***2**, 401-404, doi:10.1158/2159-8290.CD-12-0095 (2012).
42. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal***6**, pl1, doi:10.1126/scisignal.2004088 (2013).
43. Dees, N. D. *et al.* MuSiC: identifying mutational significance in cancer genomes. *Genome Res***22**, 1589-1598, doi:10.1101/gr.134635.111 (2012).
44. Vandin, F., Upfal, E. & Raphael, B. J. De novo discovery of mutated driver pathways in cancer. *Genome Res***22**, 375-385, doi:10.1101/gr.120477.111 (2012).
45. Perez-Llamas, C. & Lopez-Bigas, N. Gitools: analysis and visualisation of genomic data using interactive heat-maps. *PLoS One***6**, e19541, doi:10.1371/journal.pone.0019541 (2011).

Figures

Figure 1

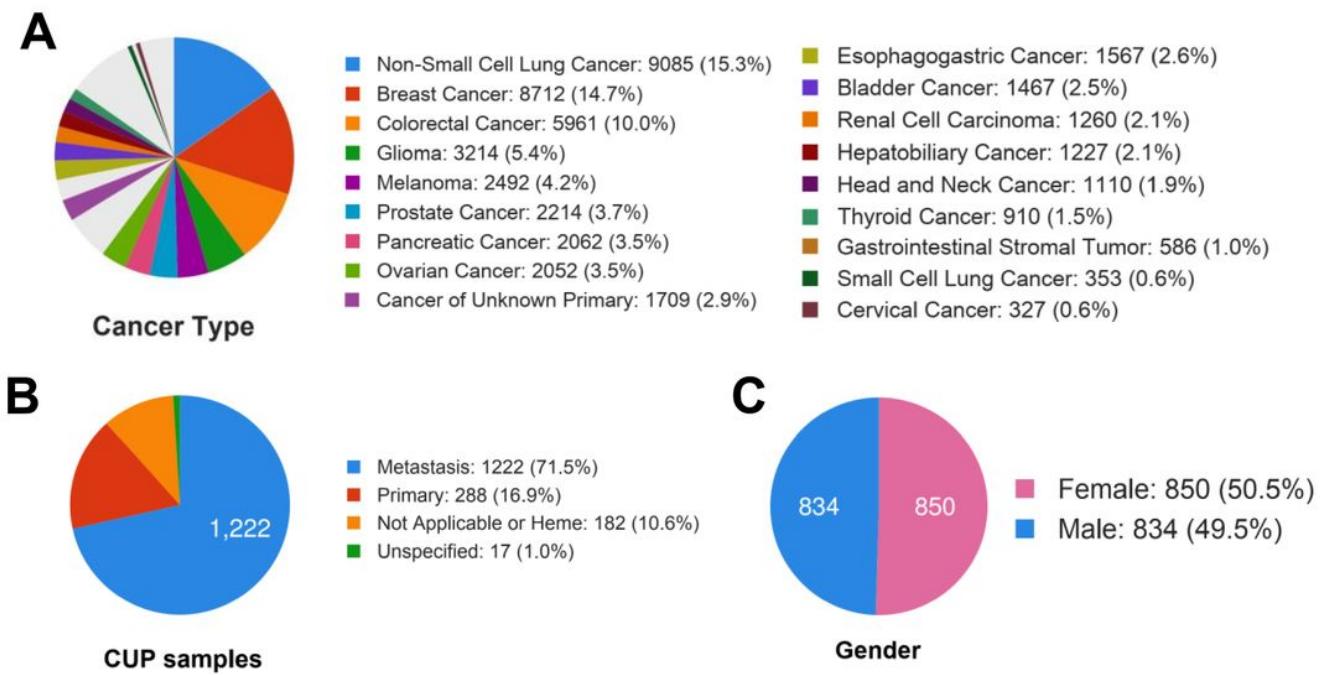


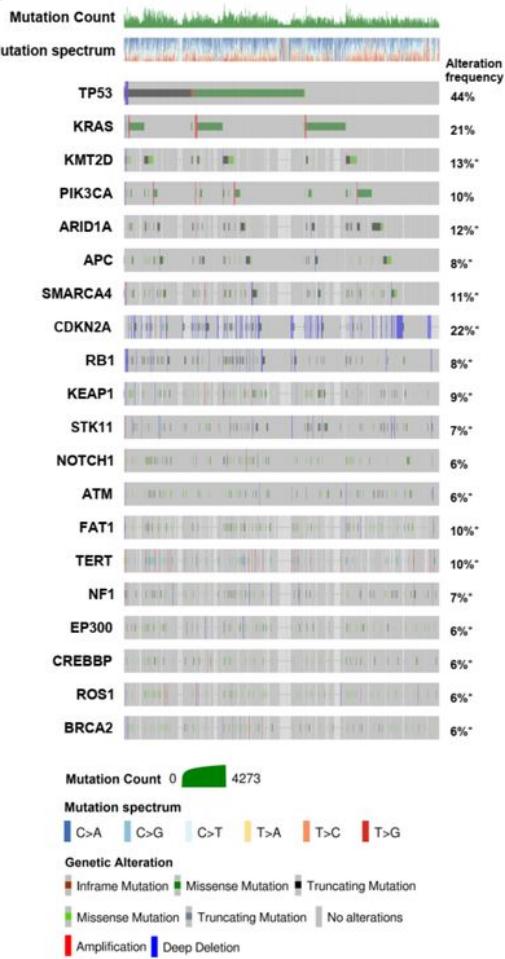
Figure 1

Overview of the GENIE database. Distribution of tumour types among cases successfully sequenced and analyzed in this cohort.

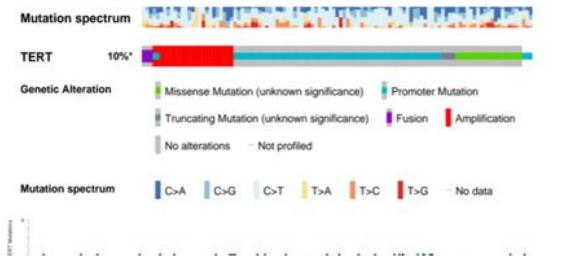
Figure 2 A

	# Mutation	Samples	Freq
TP53	803	735	43.00%
KRAS	346	341	19.90%
KMT2D	214	163	12.60%
PIK3CA	165	152	8.90%
ARID1A	187	151	11.60%
APC	162	140	8.20%
SMARCA4	141	135	10.40%
CDKN2A	134	131	10.30%
RB1	131	119	7.00%
KEAP1	115	106	8.30%
STK11	107	104	6.10%
NOTCH1	122	103	6.00%
ATM	107	99	5.80%
FAT1	113	95	9.70%
TERT	101	91	7.70%
NF1	100	88	6.60%
EP300	94	83	6.40%
CREBBP	91	80	6.20%
ROS1	89	73	5.40%
BRCA2	74	70	5.30%

B



D



C

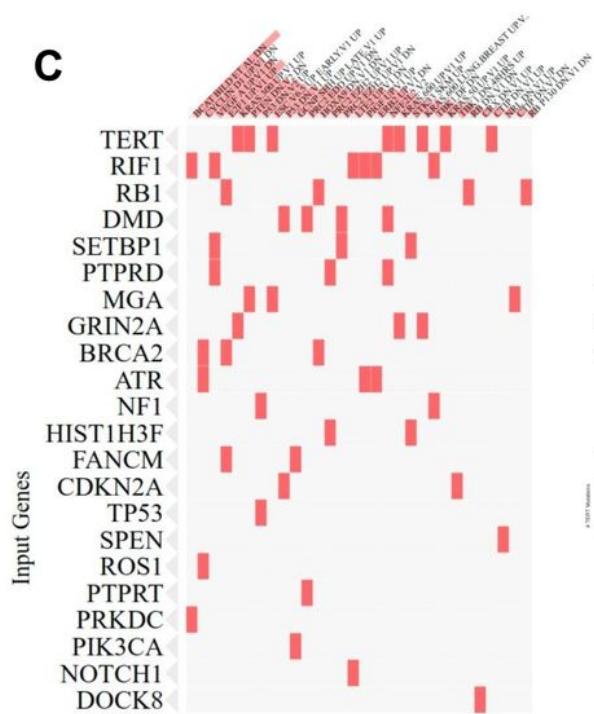


Figure 2

The most significant mutated genes in CUP samples. (A) Mutation frequency of SMGs. Genes with a cohort-level alteration frequency of > 5% or a tumour type-specific alteration frequency of > 30% are displayed. (B) Genomic alterations of 52 SMGs within CUP samples. (C) Pathway enrichment analysis of SMGs from MSigDB. (D) Genomic alterations identified in TERT promoter among CUP samples.

Figure 3



Figure 3

Mutual exclusivity and co-occurrence between identified SMG in CUP

Figure 4

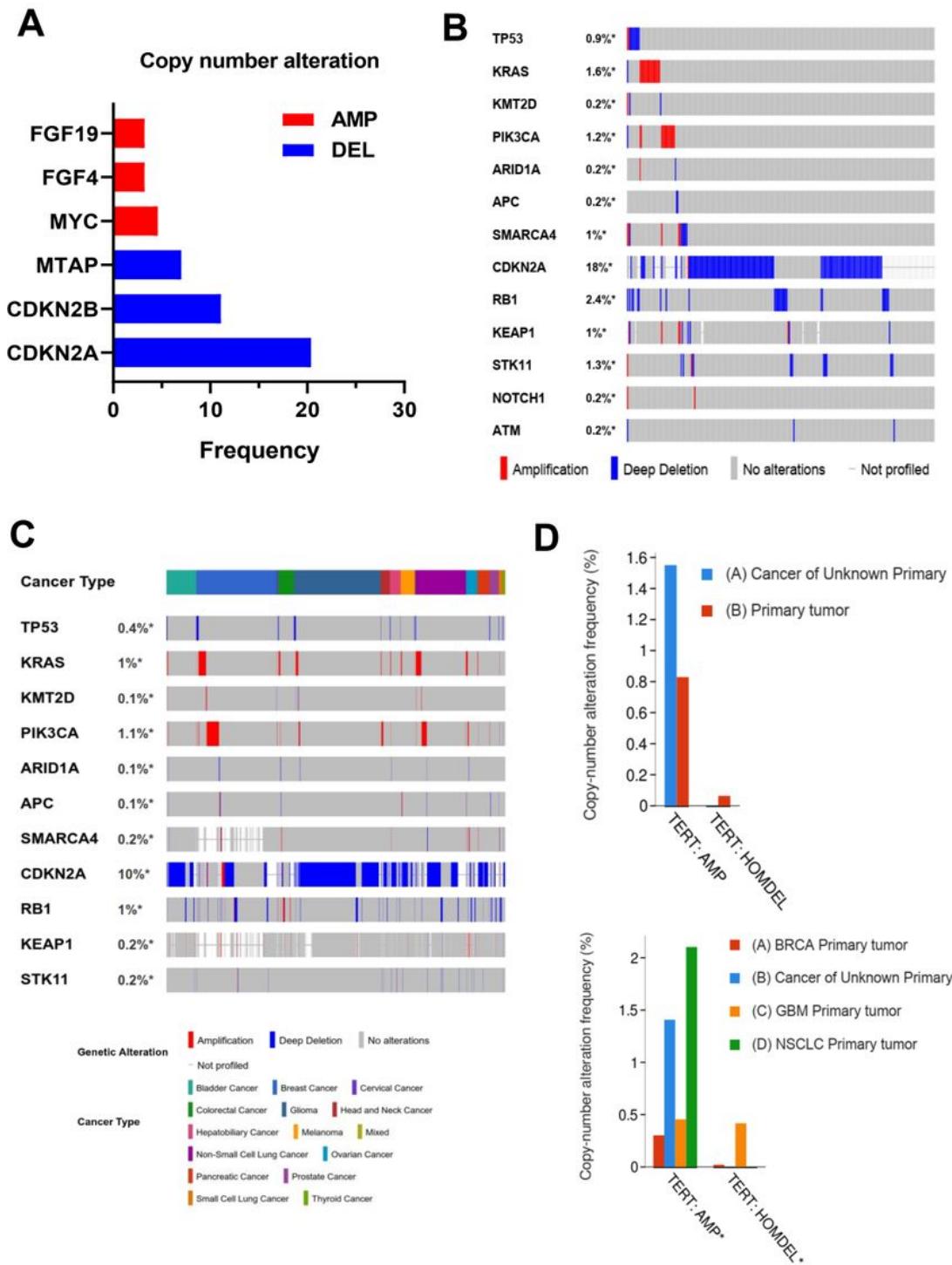


Figure 4

Copy number variation among CUP and other cancer types. (A) major copy number alteration detected in CUP samples. (B) Amplification and deletion status of identified CUP-SMGs within CUP samples and (C) other known primary tumours registered in GENIE database. (D) copy number variation analysis of TERT between CUP and all primary tumours (up panel) and NSCLC, GBM and BRCA (bottom panel).

Figure 5

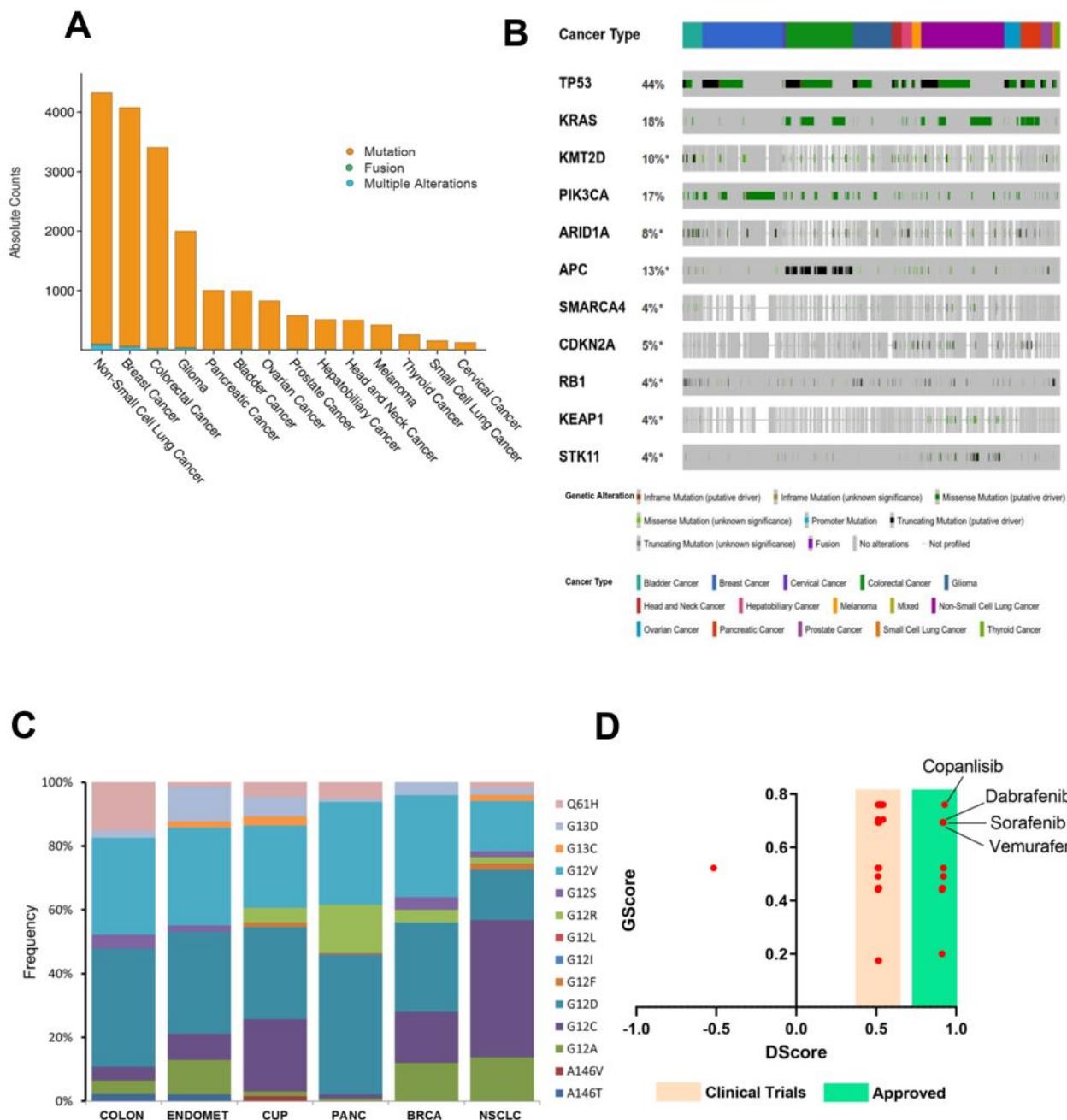


Figure 5

Mutation frequency of CUP-SMGs across 17 cancer types. (A, B) Distribution of genomic alterations in 52 CUP-SMGs across primary tumours of 17 different cancer types in GENIE cohort. (C) Distribution of hotspot mutations identified in KRAS among six different cancer types including CUP. (D) The results of gene-drug association analysis using PanDrug platforms. The best candidate drugs with highest GScore and Dscore are labeled.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryLegends.docx](#)
- [SupplementaryTable1.xlsx](#)
- [SupplementaryTable2.xlsx](#)
- [SupplementaryTable3.xlsx](#)
- [SupplementaryTable4.xlsx](#)
- [SupplementaryTable5.csv](#)