Methods 222 (2024) 122–132

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Interpreting the effect of mutations to protein binding sites from large-scale genomic screens

Sara Jamshidi Parvar [a], Benjamin A Hall [b], David Shorthouse [a,*]

[a] UCL School of Pharmacy, 29-39 Brunswick Square, London WC1N 1AX, UK
[b] UCL Department of Medical Physics and Biomedical Engineering, Mallet Place Engineering Building, University College London, Gower Street, London WC1E 6BT, UK

ARTICLE INFO

Keywords:
Binding calculations
FoldX
Protein structure
Missense mutation

ABSTRACT

Predicting the functionality of missense mutations is extremely difficult. Large-scale genomic screens are commonly performed to identify mutational correlates or drivers of disease and treatment resistance, but interpretation of how these mutations impact protein function is limited. One such consequence of mutations to a protein is to impact its ability to bind and interact with partners or small molecules such as ATP, thereby modulating its function. Multiple methods exist for predicting the impact of a single mutation on protein–protein binding energy, but it is difficult in the context of a genomic screen to understand if these mutations with large impacts on binding are more common than statistically expected. We present a methodology for taking mutational data from large-scale genomic screens and generating functional and statistical insights into their role in the binding of proteins both with each other and their small molecule ligands. This allows a quantitative and statistical analysis to determine whether mutations impacting protein binding or ligand interactions are occurring more or less frequently than expected by chance. We achieve this by calculating the potential impact of any possible mutation and comparing an expected distribution to the observed mutations. This method is applied to examples demonstrating its ability to interpret mutations involved in protein–protein binding, protein-DNA interactions, and the evolution of therapeutic resistance.

1. Introduction

With the increasing accessibility of genome sequencing methods, such as Illumina and NanoSeq methods, mutational data for a wide range of organisms and conditions are rapidly becoming available[1–3]. Genome sequencing screens are routinely being applied to study evolution of organisms[4], population diversity[5], and complex diseases such as cancer[6]. As the amount of this data increases, new methods are needed to enable interpretation and understanding of this high dimensional and functionally heterogeneous data. When a mutation is observed within a gene compared to a reference sequence, this mutation can be synonymous (does not change the amino acid composition of the protein encoded by the gene), or non-synonymous (somehow changes the amino acid composition of the protein encoded by the gene). Within non-synonymous mutations, missense mutations (which switch the amino acid at a specific site of the protein with another one) are some of the hardest to functionally interpret[7–9]. These mutations may induce any of a myriad of effects on a protein, including having no effect at all. Missense mutations may modify protein folding, alter its ability to interact with protein partners or complexes, damage a functionally important region such as an enzymatic binding site, or have other impacts. Understanding how these mutations change a protein has large implications for comprehension of protein structure biophysics, treatment decisions, and design of drugs.

Calculations can be performed on a protein structure to predict the energetic effects of a missense mutation, thereby predicting quantitative changes in the stability and folding of a protein (referred to as the Gibbs free energy or ΔΔG). We have previously applied these calculations to understand mutational function in disease[10–13]. These calculations can additionally be applied to the interaction energy between a protein and another molecule or protein, generating the ΔΔG of binding, a measure of how much a specific missense mutation energetically impacts the interface between the two molecules. Many methods exist for calculating the ΔΔG of binding, ranging from machine-learning based calculators[14,15], to "force-field" type methods that use chemical descriptions of atoms and their bonds to calculate energies[16], through to dynamical and "alchemical" methods that use computationally intensive dynamics-based methods[17,18]. In the same field of study, Brownian

* Corresponding author at: UCL School of Pharmacy, 29-39 Brunswick Square, London WC1N 1AX, UK.
E-mail address: d.shorthouse@ucl.ac.uk (D. Shorthouse).

https://doi.org/10.1016/j.ymeth.2023.12.008
Received 31 August 2023; Received in revised form 27 November 2023; Accepted 22 December 2023
Available online 5 January 2024

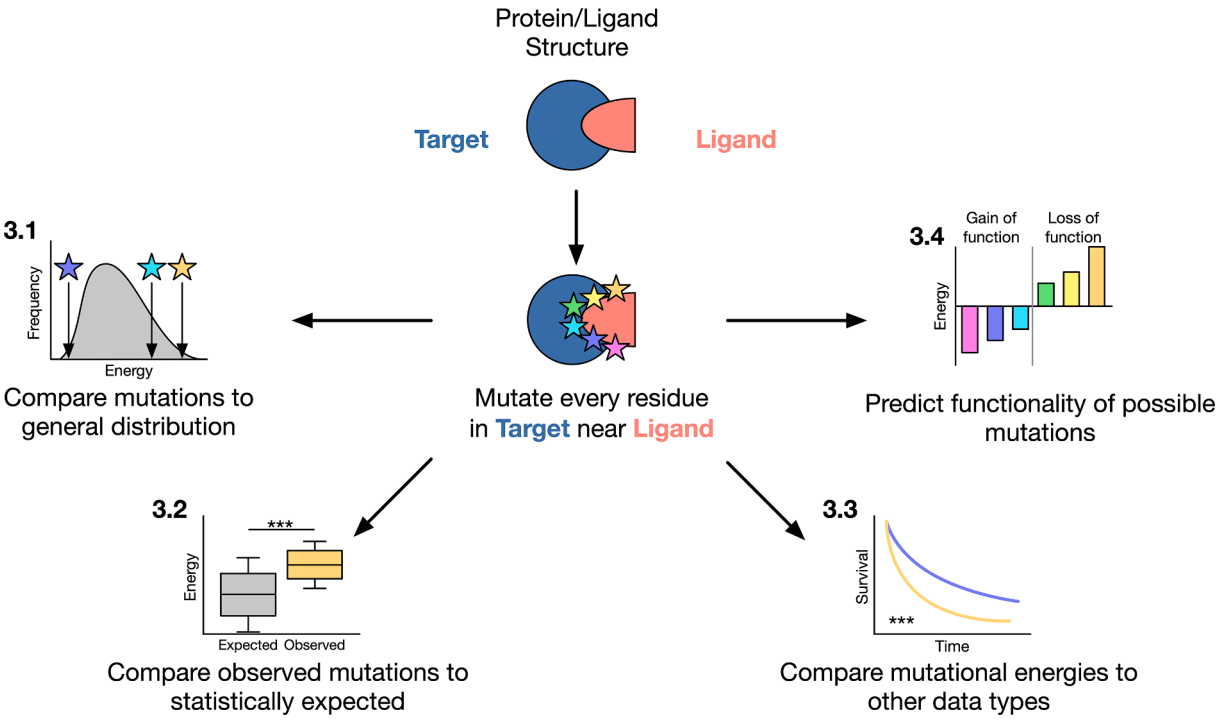

Fig. 1. Workflow for performing binding energy mutagenesis saturation screens. Results can then be used to study single mutations, groups of mutations, compare mutational energies to other data types, and predict the functionality of any potential mutation.

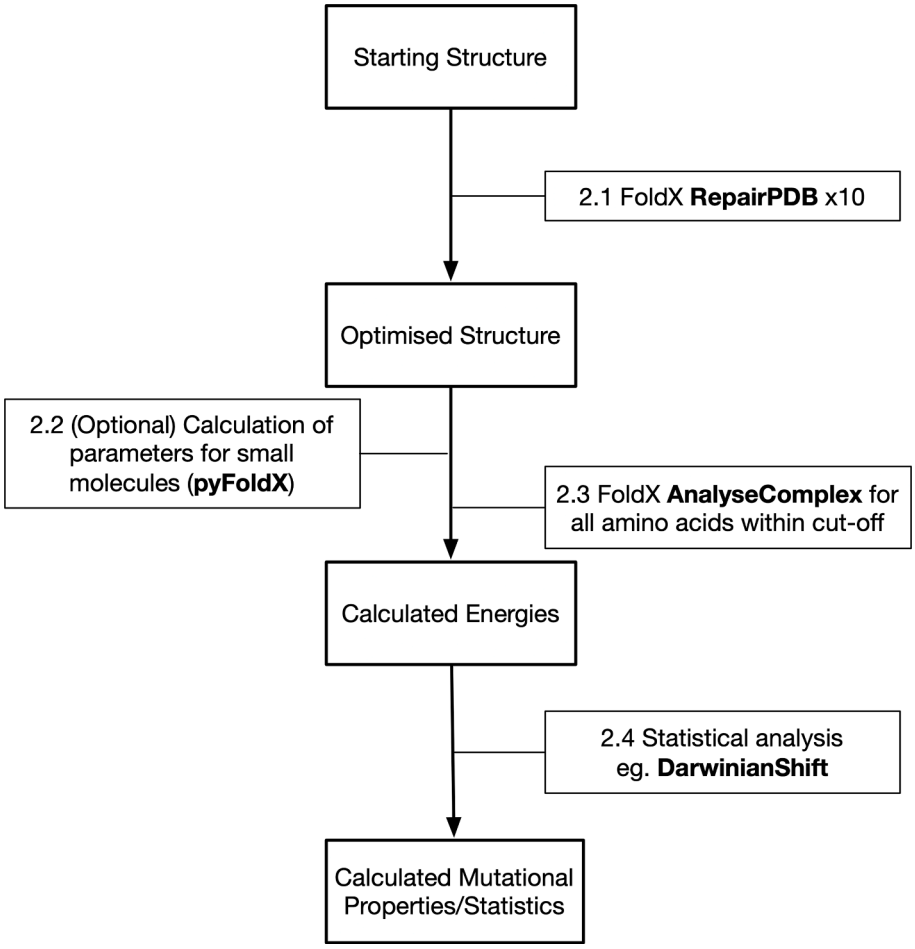

Fig. 2. Flow diagram for workflow presented in this manuscript. Numbers represent sections of this manuscript where the methods are explained.

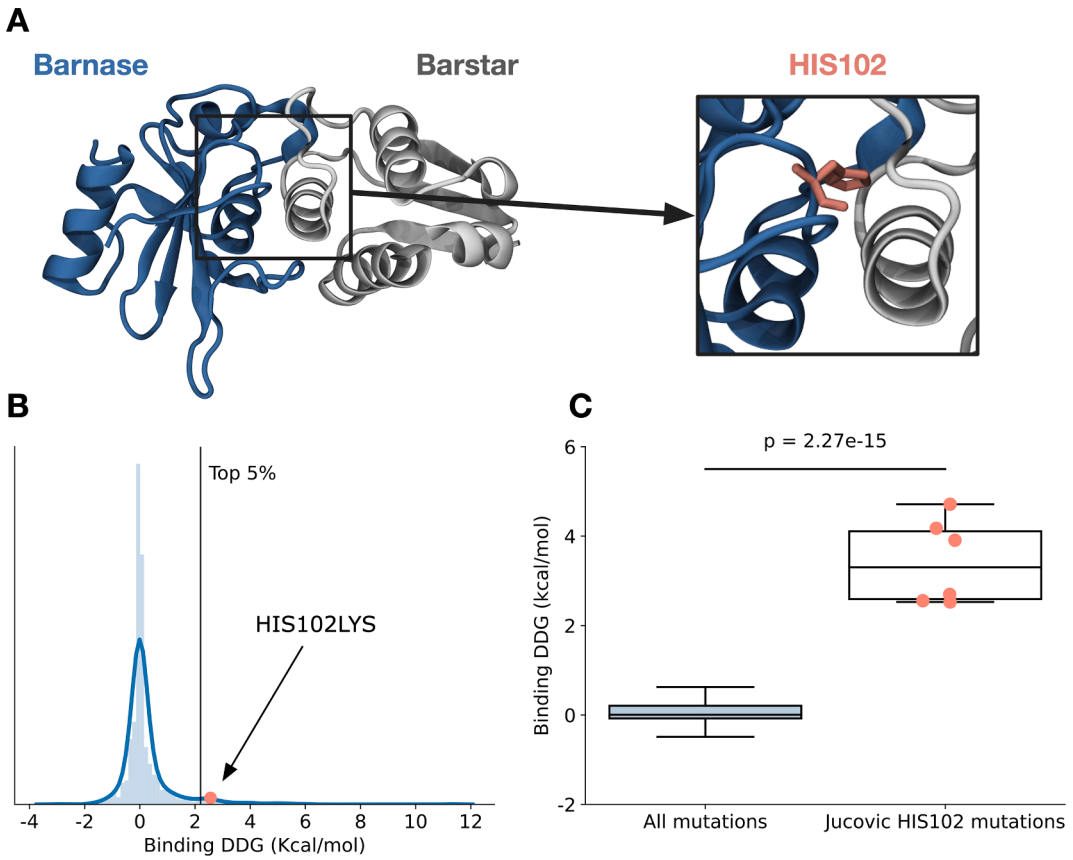

Fig. 3. Comparing single and groups of mutations. A - Structure of barnase-barstar complex with HIS102 highlighted in red (insert – right). B - Binding energy distribution showing the top 5% of mutations, and highlighting the location of HIS102LYS. C - Binding energy distributions for all mutations, and a subset of HIS102 mutations known to disrupt binding of barnase-barstar. P value represents independent *t*-test. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

dynamics calculations can also be used to estimate binding affinity of two biophysical objects, and provide the advantage of incorporating changes to environmental conditions and interactions with surfaces [19,20].

Interpretation of these energy predictions however can be challenging. Individual calculations may denote a large change in interaction energy, but knowing whether they are larger than an "expected" change due to randomness or the mutational background is necessary to evidence that they are under evolutionary selection, and are therefore highly likely to be functionally impactful. Mutation sets can be compared to all possible mutations to study whether they are enriched. Previous studies have developed statistical tests for analysing mutational data applied to protein structures, producing an "expected distribution" for a biophysical property[21]. In this work, we demonstrate how binding site saturation mutagenesis screens can be performed in silico using the tool FoldX[16], and how the data from these screens can be statistically analysed to understand mutational selection and functional impact. Similar results would be expected to be obtained using any method for calculating the estimated mutational ΔΔG such as Rosetta, as previous studies have reported high similarities between results from a range of methods[22–25]. We provide all code and analysis in the form of scripts and notebooks available at https://github.com/shorthouse-lab/binding_ddg, and apply our methods to a range of biological systems including protein–protein, protein-DNA, and protein–ligand interactions (Fig. 1).

2. Methods

We present methods for calculating binding energies – here applied to the example case of the barnase-barstar complex 1BRS (https://doi.org/10.2210/pdb1BRS/pdb). All code required to replicate this workflow is provided at https://github.com/shorthouse-lab/binding_ddg. We explore analysis of the generated data for this and other complexes in the results section. This workflow involves the energetic optimisation of a target structure, the calculation of mutational energies for every possible amino acid mutation within a defined cutoff, and then statistical analysis of the results (Fig. 2). For this study, calculations were performed on a virtual machine hosted by NMRbox[26].

2.1. Structure preparation

Our workflow utilises the energy calculation methods included as part of FoldX[16]. Prior to performing calculations, structures need to undergo cleaning and energy minimisation. The structure is cleaned by removing excess molecules and complexes from the.pdb file. In the case of barnase-barstar complex, we used the.pdb file 1BRS. This file contains the structures for three barnase-barstar complexes; we opted to retain only chains A and D, which corresponds to one complex. Subsequently, this complex is energy minimised using the FoldX command 'repairpdb', performed ten times in sequence. As 'repairpdb' optimises each amino acid in the protein sequentially, conducting multiple rounds of repair result in a more optimised structure, determined by a lower overall calculated energy. We chose to repair the structure ten times to ensure the structure has reached a local minima, as for large structures multiple rounds of repair will result in a more optimal energy (An example of this for 3KMD can we found in the appendix). We used the following command to repair the pdb, a python script to do this ten times automatically is available at https://github.com/shorthouse-lab/binding_ddg:

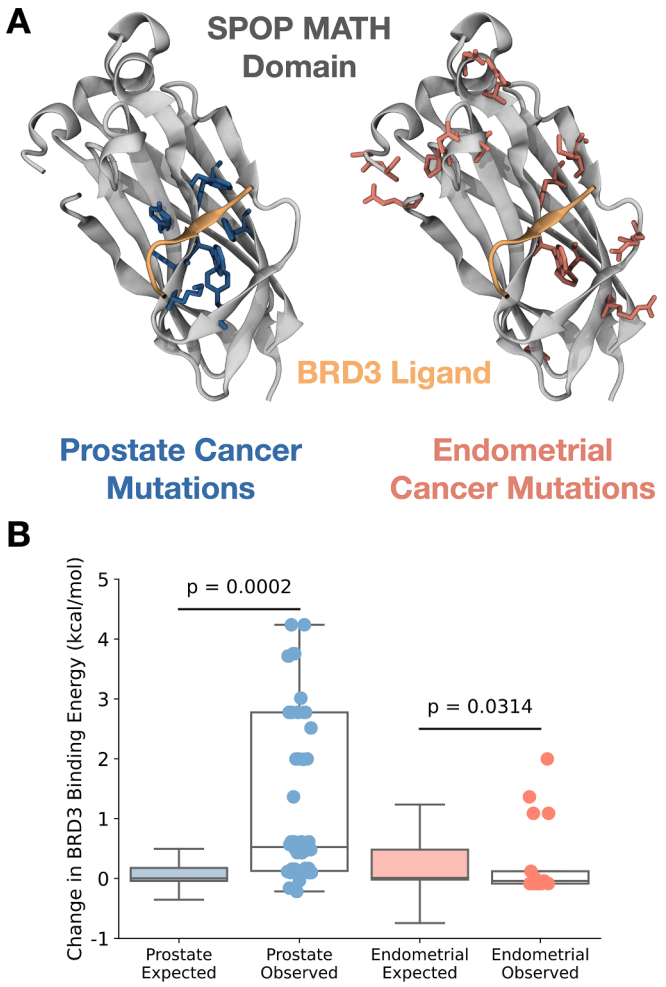

Fig. 4. Comparing mutations to an expected distribution. A - Structure of SPOP MATH domain bound to BRD3 ligand. Mutations observed in the TCGA in prostate cancer are shown in blue (left), mutations observed in the TCGA in endometrial cancer are shown in red (right). B - SPOP-BRD3 binding energy distributions for an expected mutational distribution generated from the mutational signature of the cancer, and the observed mutations for prostate and endometrial cancer in the TCGA. P value represents Monte Carlo Cumulative Distribution Function. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

```
$foldx --command=RepairPDB --pdb=1brs.pdb --water=CRYSTAL
--pH=7 --vdwDesign=2 --ionStrength=0.05
```

.

Inclusion of the "–water = CRYSTAL" command allows waters in the interface of the complex to be retained and included in energy calculations. "–pH" and "–ionStrength" flags enable to tool to perform calculations under physiologically relevant pH and ionic conditions.

2.2. Molecule parameterisation

FoldX recognises certain molecules, such as ATP, using prebuilt parameters. This allows the calculation of energetic effects of mutations on binding, potentially indicating mutations that increase or reduce binding affinity and consequently altering enzymatic activity. These calculations can be performed with a recognised molecule through setting the molecule as the "ligand", and mutating all residues within a defined radius in the "target" protein. For other molecules however, parameters may not be available in FoldX, and so we must parameterise them. We used the existing method incorporated into pyFoldX[27] to perform parameterisation, an example of which is included at https://github.com/shorthouse-lab/binding_ddg.

2.3. Binding DDG calculation

Binding DDG calculations are then performed on every residue of the "target" protein within a defined radius of the "ligand". We chose a cut-off of 10 Å for calculations, and use Biopython[28] to calculate these, as FoldX is only expected to capture local changes in protein structures, and does not allow large conformational or backbone shifts. For each residue within the cut-off, we mutate it to each other amino acid using the FoldX command BuildModel:

```
$foldx --command=BuildModel --pdb=1brs_repair.pdb
--numberOfRuns=1 --pH=7 --vdwDesign=2 --ionStrength=0.05
--mutant-file=mutantfile.txt --water=CRYSTAL
```

.

Where "mutantfile.txt" is a file containing the mutation to induce as denoted by FoldX (Wildtype residue, chain, residue number, mutant residue). For example, to mutate the Threonine at position 100 of barnase (chain A) in 1brs.pdb to Alanine, the file will contain: "TA100A;". Energy is then calculated using the 'AnalyseComplex' command applied to either a wildtype or mutant structure:

```
$foldx --command=AnalyseComplex --pdb=pdbfile.pdb
--analyseComplexChains=A,D --pH=7 --vdwDesign=2
--ionStrength=0.05 --water=CRYSTAL
```

.

By comparing the energies of the wildtype and mutant structures for each mutation, we can calculate the change in binding energy induced by each mutation (Binding ΔΔG). Results are then aggregated to generate a single file containing the binding ΔΔG of every possible mutation to the "target" protein within the defined radius of the "ligand". Scripts to automatically perform and summarise these calculations for any complex of interest are included at https://github.com/shorthouse-lab/binding_ddg.

2.4. Data analysis

Analysis is performed with python3, using the libraries pandas, numpy[29], and scipy[30]. Plotting is performed with matplotlib[31] and seaborn[32]. Survival analysis was performed using the lifelines python package[33]. To calculate evidence of mutational selection we used the library darwinian_shift (https://github.com/michaelhall28/darwinian_shift)[21]. Jupyter notebooks containing the code for all analysis presented in this paper are available at https://github.com/shorthouse-lab/binding_ddg.

3. Results

3.1. Barnase-barstar – Probing single and small groups of mutations

We applied this workflow to the barnase-barstar complex as a proof-of-principle. The barnase-barstar complex is a classically studied pair of proteins whereby barnase is an extracellularly secreted ribonuclease from *Bacillus Amyloliquefaciens*. Barnase is lethal to a bacterial cell that does not also express its inhibitor barstar, which tightly binds it and prevents activity[34,35]. The barnase-barstar structure 1BRS[36] was downloaded from the pdb, repaired, and binding ΔΔG calculated for every possible mutation in barnase. A conserved histidine (HIS102) in barnase is known to destabilise the protein complex when mutated to Lysine[37] (Fig. 3**A**). When plotting the distribution of binding ΔΔG for all potential mutations to barnase within 10 Å of barstar, we find that HIS102LYS, with a ΔΔG of 2.56 kcal/mol, is in the top 5 % (>2.2 kcal/

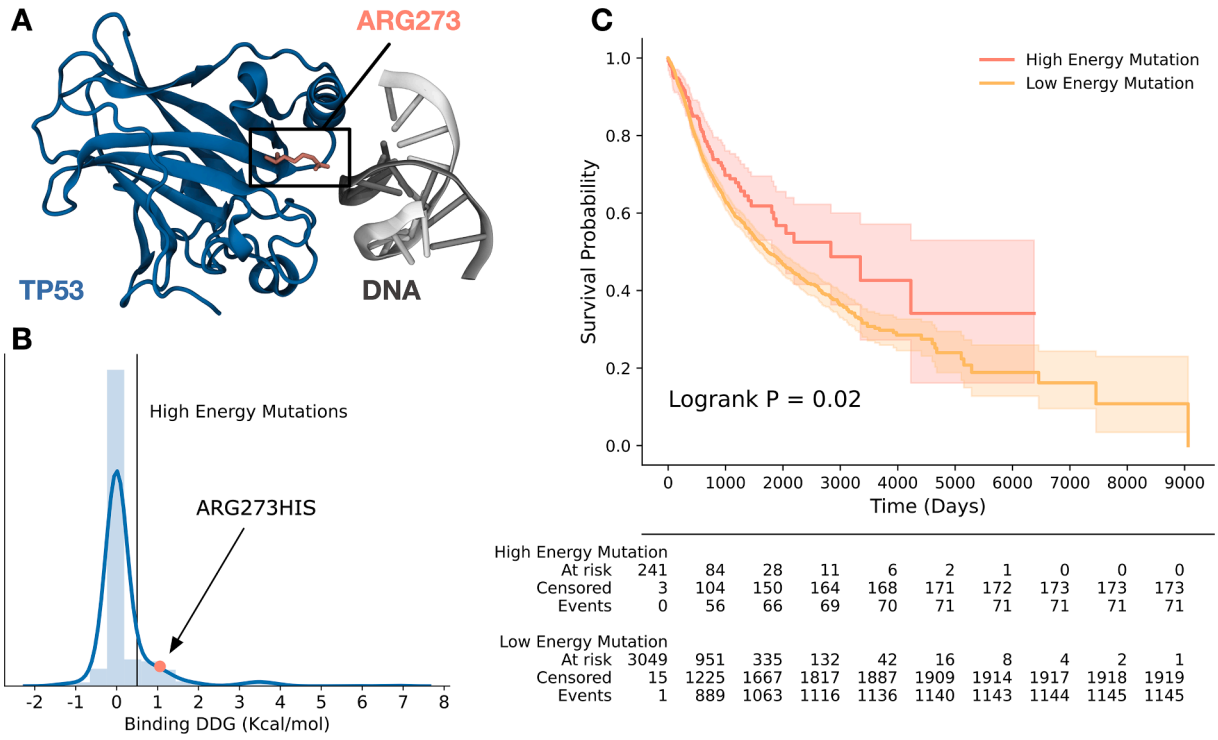

Fig. 5. Assessing mutational impact through comparison with other data. A – Structure of TP53 bound to DNA with ARG273 highlighted. B – TP53-DNA binding energy distribution for all mutations in TP53, highlighting ARG273HIS. C – Kaplan-Meier analysis of patients with high energy (ΔΔG > 0.5 kcal/mol – red) and low energy (ΔΔG < 0.5 kcal/mol – orange) mutations in TP53 across the pan-cancer TCGA dataset. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

mol) of all possible mutations (Fig. 3**B**). Furthermore, a set of other mutations to HIS102 have been experimentally determined to reduce binding affinity (HIS102ASP, HIS102GLN, HIS102GLY, HIS102LEU, HIS102ALA). By extracting the predicted binding ΔΔG for this group of mutations, we can compare them to the total population, and therefore calculate statistics to determine if they are outliers. We find that for these six mutations in barnase, they have a statistically significantly higher energy than the "base" distribution (independent *t*-test p = 2.27e-15) and therefore are higher energy than would be expected (Fig. 3**C**). This example illustrates how our method of calculating binding ΔΔG for all possible mutations allows simple statistical comparison of individual or groups of observed mutations.

3.2. SPOP – Evidencing mutational selection

We next applied this workflow to an example of mutations to protein–protein binding in large-scale genomics datasets to identify evolutionary selection. SPOP is an E3 ligase involved in degradation of protein targets, and is known to be under selection in a number of cancers[38]. Notably, SPOP missense mutations occurring in prostate cancer are thought to be loss of function (LoF), decreasing SPOP activity, whilst missense mutations occurring in endometrial cancer are thought to be gain of function (GoF), increasing protein activity[39] (Fig. 4**A**). We used the structure of SPOP bound to its protein–ligand target BRD3 (pdb id: 6I41) to calculate every possible mutation within 10 Å of the ligand. We downloaded the mutational data for a large number of patients with prostate[40] (n = 494) and endometrial[40] (n = 447) cancers from The Cancer Genome Atlas. Both cancer cohorts have numerous observed mutations within the ligand binding domain, and we have used a previously published method known as "Darwinian shift"[21] to generate expected distributions - distributions of binding ΔΔGs that would be expected if mutations to the binding site were only subject to the mutational signature of the cancer. If our observed mutations are significantly different to this expected distribution, then we can conclude that there is evidence of evolutionary selection. We find that mutations near the ligand of SPOP in prostate cancer are statistically significantly (CDF Monte Carlo p = 0.0002) higher energy than expected, and therefore would likely decrease the ability of the protein to bind BRD3, whilst the mutations in endometrial cancer are significantly (CDF Monte Carlo p = 0.0314) lower energy than expected, and therefore do not significantly increase binding energy (Fig. 4**B**). This demonstrates how binding ΔΔG calculations can be used to explore evolutionary selection in large-scale genomics datasets, with the example of understanding selection of ligand binding mutations in SPOP.

3.3. TP53 – Calculating binding to non-protein molecules

TP53 is a DNA binding protein playing a role in a majority of human cancers[41]. Most human cancers mutate TP53 to some extent, and evidence suggests that different types of mutations have different effects on cell phenotype, tumour progression, and treatment response[42]. Many mutations to TP53 impact its ability to bind DNA, whilst others damage its ability to tetramerise, and potentially have numerous other effects. To illustrate how our methodology can explore mutational effects and their relation to other data sources, we first calculated the binding ΔΔG for all mutations in the TP53 structure 3KMD within 10 Å of the cocrystallised DNA fragment (Fig. 5**A**). Energies follow a broadly normal distribution, and the known DNA contact disrupting hotspot mutation ARG273HIS is found to have a comparatively high mutational energy (1.059 kcal/mol, within the top 8 % of mutations) (Fig. 5**B**). We next collected the mutational and clinical data from the TCGA pan-cancer atlas[6], and extracted the clinical data for all patients with a missense mutation in TP53. We subset mutations into those expected to have any (even minor) destabilising effect on TP53-DNA binding (denoted "high energy" - those with a ΔΔG > 0.5 kcal/mol), and those not expected to reduce the ability of TP53 to bind DNA (denoted "low energy" - those with a ΔΔG < 0.5 kcal/mol). We find a statistically

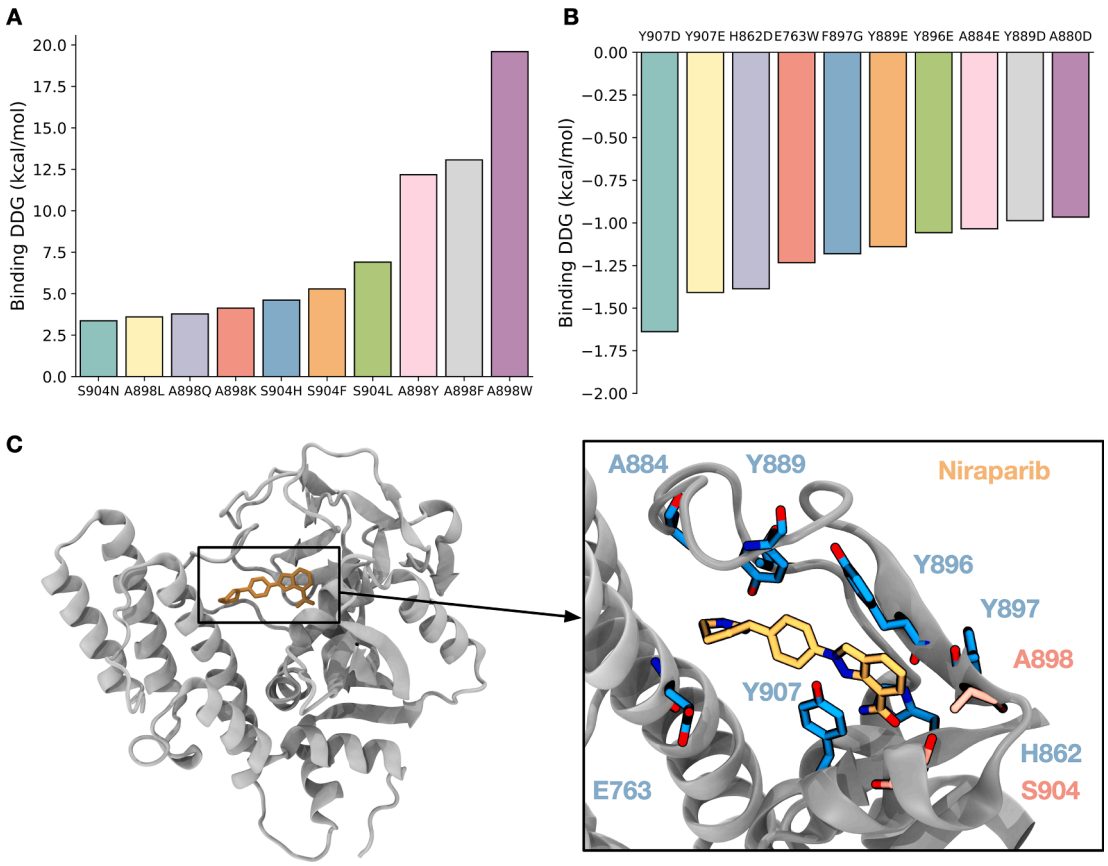

Fig. 6. Predicting mutational effects from binding ΔΔG. A – Highest ten mutations by ΔΔG for the interaction between PARP1 and the inhibitor niraparib. B – Lowest ten mutations by ΔΔG for the interaction between PARP1 and the inhibitor niraparb. C – Structure of PARP1-niraparib (Orange), highlighting the residues with high (blue) and low (red) energy mutations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

significant (logrank P = 0.02) difference in overall survival for our pan-cancer cohort (n = 3290), and patients with high energy TP53-DNA binding mutations have a better overall survival (Fig. 5**C**). This suggests that mutations that do not disrupt DNA binding lead to poorer clinical outcomes, possibly because their biophysical effects are more potent, or they are more likely to respond to treatment. With this example we demonstrate how to apply our method to correlate biophysical properties of mutations in protein structures with other information such as clinical or disease outcome data.

3.4. PARP inhibition by niraparib – Predicting resistance mutations

A final potential use of this workflow is demonstrated by prediction of mutations that might lead to resistance to drugs. Mutations in proteins that reduce the ability of a drug to bind are likely to lead to reduced toxicity to those cells, and so these mutations will be selected for in competitive conditions. This is a large problem for antibiotics, where resistance is emerging and becoming increasingly prevalent, but also for chemotherapy, whereby introduction of targeted therapies provides evolutionary pressure on the tumour to evolve resistance around them [43]. We chose to demonstrate this workflow with the example of the targeted ovarian cancer treatment, niraparib, in complex with its target protein PARP1. The niraparib-PARP1 structure has been solved (4R6E) [44]. We again repeated the presented methodology – first repairing the structure ten times before calculating all possible mutations within 10 Å of the drug. For this example, we had to generate parameters for niraparib. We chose to use pyFoldX[27] to generate parameters - each atom in the molecule is assigned to an 'atomtype' included in FoldX. This resulted in a parameter file for niraparib that, when included in the directory where calculations are performed, allows FoldX to recognise small molecules and perform energy calculations on them.

We calculated every possible mutation to PARP1 in the vicinity of niraparib, surmising that mutations which increase the binding energy significantly are likely to confer resistance. We find that within the top ten mutations by binding ΔΔG, only two amino acids are represented – residues 898 and 904, both of which have direct contact with the ligand (Fig. 6**A, C**). In particular, mutation ALA898TRP is predicted to have an extremely high ΔΔG value of 19.60 kcal/mol, suggesting that it will significantly reduce the ability of niraparib to bind to PARP1. Similarly, looking at mutations that reduce ΔΔG, we can infer that these mutations may increase niraparib binding affinity and thus sensitise the protein to inhibition. The top two mutations that reduce binding ΔΔG are both to residue TYR907 (Fig. 6**B, C**). Interestingly, this residue is known to be phosphorylated by c-Met, which subsequently reduces the effect of PARP inhibitors[45]. This highlights that whilst these predictions can give some insights into mutational effects on protein activity, biological context is required in order to fully understand and interpret them.

4. Discussion

These methods provide a toolkit for interpreting functionality of mutations to protein binding. Whilst we demonstrate that this method can interpret single or small groups of mutations, its most powerful application is in the interpretation of large-scale missense mutational data, which allows the generation of a mutational "expected" signature to compare against. We chose to use the empirical energy calculation method available in FoldX[16] as we feel that this method is a good trade-off between computational time (readily parallelisable and does

not involve dynamics calculations) and accuracy, as it has been shown to be comparable to other available methods[22–24]. However, it is possible to apply these methods to any distribution of energy calculations calculated by a different method, such as Monte Carlo sampling-based rosetta[17], or significantly more computationally intensive but higher accuracy alchemistry methods[18]. Ultimately however, this method is dependent on the accuracy of the underlying energy calculations, and whilst evidence suggests that average accuracy of FoldX is acceptable[24], it is not necessarily accurate for predicting individual mutations. We also note that this methodology is limited to cases where protein structures are available, and whilst recent advances in AI based structure and protein–protein interface predictions mitigate some of this [46,47], it is unknown exactly how reliable individual predicted structures are, and therefore how much weight can be put on the outputs from their predictions.

5. Conclusions

We present a workflow for performing and analysing calculations to understanding the impact of missense mutations on protein–ligand binding, with a focus on "big data". As increasingly large datasets of mutations are generated, methodologies such as these which enable interpretation of this complex and large information become ever more important. Our methodology involves calculating all possible mutations to an interaction site to generate a base distribution that can be used for statistical comparisons. We apply this method to single and small groups of mutations, data from large-scale genomic screens, use our results to understand clinical and other data types, and use it predictively to study potential effects of unobserved mutations. This workflow provides a toolkit for improving our understanding of notoriously hard to functionally interpret whole genome sequencing and CRISPR screens with reasonable computational cost and space requirements – for the example of 3KMD, this study generated 1 GB of data, taking just 6 h on a 40 CPU virtual machine. The workflow presented involves the use of FoldX, but can be applied to any current or future methods that predict or calculate the binding ΔΔG, including Rosetta, which has been demonstrated to generate comparable results[10,23,24]. This methodology has applications in areas of genomics research such as biomedical study of diseases like cancer, evolutionary biology, and protein or therapeutic engineering. In an effort to make this workflow as accessible as possible, we include code to implement these analyses at https://github.com/shorthouse-lab/binding_ddg.

Funding

SJP is funded through the EPSRC & SFI Centre for Doctoral Training in Transformative Pharmaceutical Technologies, United Kingdom (grant no. EP/S023054/1). BAH acknowledges support from the Royal Society, United Kingdom (grant no. UF130039).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data is included or linked in the methods or available at the associated GitHub: https://github.com/shorthouse-lab/binding_ddg

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ymeth.2023.12.008.

References

[1] E. Turro, W.J. Astle, K. Megy, S. Gräf, D. Greene, O. Shamardina, H.L. Allen, A. Sanchis-Juan, M. Frontini, C. Thys, J. Stephens, R. Mapeta, O.S. Burren, K. Downes, M. Haimel, S. Tuna, S.V.V. Deevi, T.J. Aitman, D.L. Bennett, P. Calleja, K. Carss, M.J. Caulfield, P.F. Chinnery, P.H. Dixon, D.P. Gale, R. James, A. Koziell, M.A. Laffan, A.P. Levine, E.R. Maher, H.S. Markus, J. Morales, N.W. Morrell, A. D. Mumford, E. Ormondroyd, S. Rankin, A. Rendon, S. Richardson, I. Roberts, N.B. A. Roy, M.A. Saleem, K.G.C. Smith, H. Stark, R.Y.Y. Tan, A.C. Themistocleous, A. J. Thrasher, H. Watkins, A.R. Webster, M.R. Wilkins, C. Williamson, J. Whitworth, S. Humphray, D.R. Bentley, S. Abbs, L. Abulhoul, J. Adlard, M. Ahmed, H. Alachkar, D.J. Allsup, J. Almeida-King, P. Ancliff, R. Antrobus, R. Armstrong, G. Arno, S. Ashford, A. Attwood, P. Aurora, C. Babbs, C. Bacchelli, T. Bakchoul, S. Banka, T. Bariana, J. Barwell, J. Batista, H.E. Baxendale, P.L. Beales, D. R. Bentley, A. Bierzynska, T. Biss, M.A.K. Bitner-Glindzicz, G.C. Black, M. Bleda, I. Blesneac, D. Bockenhauer, H. Bogaard, C.J. Bourne, S. Boyce, J.R. Bradley, E. Bragin, G. Breen, P. Brennan, C. Brewer, M. Brown, A.C. Browning, M. J. Browning, R.J. Buchan, M.S. Buckland, T. Bueser, C.B. Diz, J. Burn, S.O. Burns, O.S. Burren, N. Burrows, C. Campbell, G. Carr-White, K. Carss, R. Casey, J. Chambers, J. Chambers, M.M.Y. Chan, C. Cheah, F. Cheng, P.F. Chinnery, M. Chitre, M.T. Christian, C. Church, J. Clayton-Smith, M. Cleary, N.C. Brod, G. Coghlan, E. Colby, T.R.P. Cole, J. Collins, P.W. Collins, C. Colombo, C. J. Compton, R. Condliffe, S. Cook, H.T. Cook, N. Cooper, P.A.A. Corris, A. Furnell, F. Cunningham, N.S. Curry, A.J. Cutler, M.J. Daniels, M. Dattani, L.C. Daugherty, J. Davis, A. De Soyza, S.V.V. Deevi, T. Dent, C. Deshpande, E.F. Dewhurst, P. H. Dixon, S. Douzgou, K. Downes, A.M. Drazyk, E. Drewe, D. Duarte, T. Dutt, J.D. M. Edgar, K. Edwards, W. Egner, M.N. Ekani, P. Elliott, W.N. Erber, M. Erwood, M. C. Estiu, D.G. Evans, G. Evans, T. Everington, M. Eyries, H. Fassihi, R. Favier, J. Findhammer, D. Fletcher, F.A. Flinter, R.A. Floto, T. Fowler, J. Fox, A.J. Frary, C. E. French, K. Freson, M. Frontini, D.P. Gale, H. Gall, V. Ganesan, M. Gattens, C. Geoghegan, T.S.A. Gerighty, A.G. Gharavi, S. Ghio, H.A. Ghofrani, J.S.R. Gibbs, K. Gibson, K.C. Gilmour, B. Girerd, N.S. Gleadall, S. Goddard, D.B. Goldstein, K. Gomez, P. Gordins, D. Gosal, S. Gräf, J. Graham, L. Grassi, D. Greene, L. Greenhalgh, A. Greinacher, P. Gresele, P. Griffiths, S. Grigoriadou, R.J. Grocock, D. Grozeva, M. Gurnell, S. Hackett, C. Hadinnapola, W.M. Hague, R. Hague, M. Haimel, M. Hall, H.L. Hanson, E. Haque, K. Harkness, A.R. Harper, C.L.L. Harris, D. Hart, A. Hassan, G. Hayman, A. Henderson, A. Herwadkar, J. Hoffman, S. Holden, R. Horvath, H. Houlden, A.C.C. Houweling, L.S. Howard, F. Hu, G. Hudson, J. Hughes, A.P. Huissoon, M. Humbert, S. Humphray, S. Hunter, M. Hurles, M. Irving, L. Izatt, S.A. Johnson, S. Jolles, J. Jolley, D. Josifova, N. Jurkute, T. Karten, J. Karten, M.A. Kasanicki, H. Kazkaz, R. Kazmi, P. Kelleher, A.M. Kelly, W. Kelsall, C. Kempster, D.G. Kiely, N. Kingston, R. Klima, N. Koelling, M. Kostadima, G. Kovacs, A. Koziell, R. Kreuzhuber, T.W. Kuijpers, A. Kumar, D. Kumararatne, M.A. Kurian, M.A. Laffan, F. Lalloo, M. Lambert, A. Lawrie, D. M. Layton, N. Lench, C. Lentaigne, T. Lester, A.P. Levine, R. Linger, H. Longhurst, L.E. Lorenzo, E. Louka, P.A. Lyons, R.D. Machado, R.V. MacKenzie Ross, B. Madan, E.R. Maher, J. Maimaris, S. Malka, S. Mangles, R. Mapeta, K.J. Marchbank, S. Marks, H.U. Marschall, A. Marshall, J. Martin, M. Mathias, E. Matthews, H. Maxwell, P. McAlinden, M.I. McCarthy, H. McKinney, A. McMahon, S. Meacham, A.J. Mead, I.M. Castello, K. Megy, S.G.G. Mehta, M. Michaelides, C. Millar, S.N. Mohammed, S. Moledina, D. Montani, A.T. Moore, J. Morales, N. W. Morrell, M. Mozere, K.W. Muir, A.D. Mumford, A.H. Nemeth, W.G. Newman, M. Newnham, S. Noorani, P. Nurden, J. O'Sullivan, S. Obaji, C. Odhams, S. Okoli, A. Olschewski, H. Olschewski, K.R. Ong, S.H. Oram, E. Ormondroyd, W. H. Ouwehand, C. Palles, S. Papadia, S.M. Park, D. Parry, S. Patel, J. Paterson, A. Peacock, S.H.H. Pearce, J. Peden, K. Peerlinck, C.J. Penkett, J. Pepke-Zaba, R. Petersen, C. Pilkington, K.E.S. Poole, R. Prathalingam, B. Psaila, A. Pyle, R. Quinton, S. Rahman, S. Rankin, A. Rao, F.L. Raymond, P.J. Rayner-Matthews, C. Rees, T. Renton, C.J. Rhodes, A.S.C. Rice, S. Richardson, A. Richter, L. Robert, I. Roberts, A. Rogers, S.J. Rose, R. Ross-Russell, C. Roughley, N.B.A. Roy, D. M. Ruddy, O. Sadeghi-Alavijeh, M.A. Saleem, N. Samani, C. Samarghitean, A. Sanchis-Juan, R.B. Sargur, R.N. Sarkany, S. Satchell, S. Savic, J.A. Sayer, G. Sayer, L. Scelsi, A.M. Schaefer, S. Schulman, R. Scott, M. Scully, C. Searle, W. Seeger, A. Sen, W.A.C. Sewell, D. Seyres, N. Shah, O. Shamardina, S.E. Shapiro, A.C. Shaw, P.J. Short, K. Sibson, L. Side, I. Simeoni, M.A.A. Simpson, M.C. Sims, S. Sivapalaratnam, D. Smedley, K.R. Smith, K. Snape, N. Soranzo, F. Soubrier, L. Southgate, O. Spasic-Boskovic, S. Staines, E. Staples, H. Stark, J. Stephens, C. Steward, K.E. Stirrups, A. Stuckey, J. Suntharalingam, E.M. Swietlik, P. Syrris, R. C. Tait, K. Talks, R.Y.Y. Tan, K. Tate, J.M. Taylor, J.C. Taylor, J.E. Thaventhiran, A. C. Themistocleous, E. Thomas, D. Thomas, M.J. Thomas, P. Thomas, K. Thomson, A.J. Thrasher, G. Threadgold, C. Thys, T. Tilly, M. Tischkowitz, C. Titterton, J. A. Todd, C.H. Toh, B. Tolhuis, I.P. Tomlinson, M. Toshner, M. Traylor, C. Treacy, P. Treadaway, R. Trembath, S. Tuna, W. Turek, E. Turro, P. Twiss, T. Vale, C. Van Geet, N. van Zuydam, M. Vandekuilen, A.M. Vandersteen, M. Vazquez-Lopez, J. von Ziegenweidt, A.V. Noordegraaf, A. Wagner, Q. Waisfisz, S.M. Walker, N. Walker, K. Walter, J.S. Ware, H. Watkins, C. Watt, A.R. Webster, L. Wedderburn, W. Wei, S.B. Welch, J. Wessels, S.K. Westbury, J.P. Westwood, J. Wharton, D. Whitehorn, J. Whitworth, A.O.M. Wilkie, M.R. Wilkins, C. Williamson, B. T. Wilson, E.K.S. Wong, N. Wood, Y. Wood, C.G. Woods, E.R.R. Woodward, S. J. Wort, A. Worth, M. Wright, K. Yates, P.F.K. Yong, T. Young, P. Yu, P. Yu-Wai-Man, E. Zlamalova, N. Kingston, N. Walker, C.J. Penkett, K. Freson, K.E. Stirrups, F. L. Raymond, Whole-genome sequencing of patients with rare diseases in a national health system, Nature 583 (2020), https://doi.org/10.1038/s41586-020-2434-2.

[2] D.M. Altshuler, R.A. Gibbs, L. Peltonen, S.F. Schaffner, F. Yu, E. Dermitzakis, P. E. Bonnen, P.I.W. De Bakker, P. Deloukas, S.B. Gabriel, R. Gwilliam, S. Hunt, M. Inouye, X. Jia, M. Aarno Palotie, P. Parkin, K. Whittaker, A. Chang, L.R. Hawes,

Y. Lewis, D. Ren, D.M. Wheeler, C. Muzny, K. Barnes, M. Darvishi, J.M. Hurles, K. Korn, C. Kristiansson, S.A. Lee, J. McCarroll, A. Nemesh, S.B.M. Keinan, A. L. Samuela Pollack, N. Price, C. Soranzo, V. Gonzaga-Jauregui, W. Anttila, M. J. Brodeur, S. Daly, G. Leslie, L. McVean, H. Moutsianas, Q. Nguyen, M.J.R. Zhang, R. Ghori, W. McGinnis, S. McLaren, F. Pollack, S.R. Takeuchi, I. Grossman, E. B. Shlyakhter, P.C. Hostetter, C.A. Sabeti, M.W. Adebamowo, D.R. Foster, J. Gordon, M.C. Licinio, P.A. Manca, I. Marshall, D. Matsuda, V.O. Ngare, D. Wang, C.N. Reddy, C.D. Rotimi, R.R. Royal, C. Sharp, D.B. Zeng, J.E.M. Brooks, Integrating common and rare genetic variation in diverse human populations, Nature 467 (2010), https://doi.org/10.1038/nature09298.

[3] F. Abascal, L.M.R. Harvey, E. Mitchell, A.R.J. Lawson, S.V. Lensing, P. Ellis, A.J. C. Russell, R.E. Alcantara, A. Baez-Ortega, Y. Wang, E.J. Kwa, H. Lee-Six, A. Cagan, T.H.H. Coorens, M.S. Chapman, S. Olafsson, S. Leonard, D. Jones, H.E. Machado, M. Davies, N.F. Øbro, K.T. Mahubani, K. Allinson, M. Gerstung, K. Saeb-Parsy, D. G. Kent, E. Laurenti, M.R. Stratton, R. Rahbari, P.J. Campbell, R.J. Osborne, I. Martincorena, Somatic mutation landscapes at single-molecule resolution, Nature 593 (2021), https://doi.org/10.1038/s41586-021-03477-4.

[4] Z.N. Kronenberg, I.T. Fiddes, D. Gordon, S. Murali, S. Cantsilieris, O.S. Meyerson, J. G. Underwood, B.J. Nelson, M.J.P. Chaisson, M.L. Dougherty, K.M. Munson, A.R. Hastie, M. Diekhans, F. Hormozdiari, N. Lorusso, K. Hoekzema, R. Qiu, K. Clark, A. Raja, A.M.E. Welch, M. Sorensen, C. Baker, R.S. Fulton, J. Armstrong, T.A. Graves-Lindsay, A.M. Denli, E.R. Hoppe, P.H. Hsieh, C.M. Hill, A.W.C. Pang, J. Lee, E.T. Lam, S.K. Dutcher, F.H. Gage, W.C. Warren, J. Shendure, D. Haussler, V.A. Schneider, H. Cao, M. Ventura, R.K. Wilson, B. Paten, A. Pollen, E.E. Eichler, High-resolution comparative analysis of great ape genomes, Science (1979). 360 (2018). https://doi.org/10.1126/science.aar6343.

[5] P.A. Audano, A. Sulovari, T.A. Graves-Lindsay, S. Cantsilieris, M. Sorensen, A.M. E. Welch, M.L. Dougherty, B.J. Nelson, A. Shah, S.K. Dutcher, W.C. Warren, V. Magrini, S.D. McGrath, Y.I. Li, R.K. Wilson, E.E. Eichler, Characterizing the major structural variant alleles of the human genome, Cell 176 (2019), https://doi.org/10.1016/j.cell.2018.12.019.

[6] P.J. Campbell, G. Getz, J.O. Korbel, J.M. Stuart, J.L. Jennings, L.D. Stein, M.D. Perry, H.K. Nahal-Bose, B.F.F. Ouellette, C.H. Li, E. Rheinbay, G.P. Nielsen, D.C. Sgroi, C.L. Wu, W.C. Faquin, V. Deshpande, P.C. Boutros, A.J. Lazar, K.A. Hoadley, D.N. Louis, L.J. Dursi, C.K. Yung, M.H. Bailey, G. Saksena, K.M. Raine, I. Buchhalter, K. Kleinheinz, M. Schlesner, J. Zhang, W. Wang, D.A. Wheeler, L. Ding, J.T. Simpson, B.D. O'Connor, S. Yakneen, K. Ellrott, N. Miyoshi, A.P. Butler, R. Royo, S.I. Shorser, M. Vazquez, T. Rausch, G. Tiao, S.M. Waszak, B. Rodriguez-Martin, S. Shringarpure, D.Y. Wu, G.M. Demidov, O. Delaneau, S. Hayashi, S. Imoto, N. Habermann, A. V. Segre, E. Garrison, A. Cafferkey, E.G. Alvarez, J.M. Heredia-Genestar, F. Muyas, O. Drechsel, A.L. Bruzos, J. Temes, J. Zamora, A. Baez-Ortega, H.L. Kim, R.J. Mashl, K. Ye, A. DiBiase, K. lin Huang, I. Letunic, M.D. McLellan, S.J. Newhouse, T. Shmaya, S. Kumar, D.C. Wedge, M.H. Wright, V.D. Yellapantula, M. Gerstein, E. Khurana, T. Marques-Bonet, A. Navarro, C.D. Bustamante, R. Siebert, H. Nakagawa, D.F. Easton, S. Ossowski, J.M.C. Tubio, F.M. De La Vega, X. Estivill, D. Yuen, G.L. Mihaiescu, L. Omberg, V. Ferretti, R. Sabarinathan, O. Pich, A. Gonzalez-Perez, A. Taylor-Weiner, M.W. Fittall, J. Demeulemeester, M. Tarabichi, N.D. Roberts, P. Van Loo, I. Cortés-Ciriano, L. Urban, P. Park, B. Zhu, E. Pitkänen, Y. Li, N. Saini, L.J. Klimczak, J. Weischenfeldt, N. Sidiropoulos, L.B. Alexandrov, R. Rabionet, G. Escaramis, M. Bosio, A.Z. Holik, H. Susak, A. Prasad, S. Erkek, C. Calabrese, B. Raeder, E. Harrington, S. Mayes, D. Turner, S. Juul, S.A. Roberts, L. Song, R. Koster, L. Mirabello, X. Hua, T.J. Tanskanen, M. Tojo, J. Chen, L.A. Aaltonen, G. Rätsch, R.F. Schwarz, A.J. Butte, A. Brazma, S.J. Chanock, N. Chatterjee, O. Stegle, O. Harismendy, G.S. Bova, D.A. Gordenin, D. Haan, L. Sieverling, L. Feuerbach, D. Chalmers, Y. Joly, B.M. Knoppers, F. Molnár-Gábor, M. Phillips, A. Thorogood, D. Townend, M. Goldman, N.A. Fonseca, Q. Xiang, B. Craft, E. Piñeiro-Yáñez, A. Muñoz, R. Petryszak, A. Füllgrabe, F. Al-Shahrour, M. Keays, D. Haussler, J.N. Weinstein, W. Huber, A. Valencia, I. Papatheodorou, J. Zhu, Y. Fan, D. Torrents, M. Bieg, K. Chen, Z. Chong, K. Cibulskis, R. Eils, R.S. Fulton, J.L. Gelpi, S. Gonzalez, I.G. Gut, F. Hach, M. Heinold, T. Hu, V. Huang, B. Hutter, N. Jäger, J. Jung, Y. Kumar, C. Lalansingh, I. Leshchiner, D. Livitz, E.Z. Ma, Y.E. Maruvka, A. Milovanovic, M.M. Nielsen, N. Paramasivam, J.S. Pedersen, M. Puiggròs, S.C. Sahinalp, I. Sarrafi, C. Stewart, M.D. Stobbe, J.A. Wala, J. Wang, M. Wendl, J. Werner, Z. Wu, H. Xue, T.N. Yamaguchi, V. Yellapantula, B.N. Davis-Dusenbery, R.L. Grossman, Y. Kim, M.C. Heinold, J. Hinton, D.R. Jones, A. Menzies, L. Stebbings, J.M. Hess, M. Rosenberg, A.J. Dunford, M. Gupta, M. Imielinski, M. Meyerson, R. Beroukhim, J. Reimand, P. Dhingra, F. Favero, S. Dentro, J. Wintersinger, V. Rudneva, J.W. Park, E.P. Hong, S. G. Heo, A. Kahles, K. Van Lehmann, C.M. Soulette, Y. Shiraishi, F. Liu, Y. He, D. Demircioğlu, N.R. Davidson, L. Greger, S. Li, D. Liu, S.G. Stark, F. Zhang, S.B. Amin, P. Bailey, A. Chateigner, M. Frenkel-Morgenstern, Y. Hou, M.R. Huska, H. Kilpinen, F.C. Lamaze, C. Li, X. Li, X. Li, X. Liu, M.G. Marin, J. Markowski, T. Nandi, A.I. Ojesina, Q. Pan-Hammarström, P.J. Park, C.S. Pedamallu, H. Su, P. Tan, B.T. Teh, J. Wang, H. Xiong, C. Ye, C. Yung, X. Zhang, L. Zheng, S. Zhu, P. Awadalla, C.J. Creighton, K. Wu, H. Yang, J. Göke, Z. Zhang, A.N. Brooks, M.W. Fittall, I. Martincorena, C. Rubio-Perez, M. Juul, S. Schumacher, O. Shapira, D. Tamborero, L. Mularoni, H. Hornshøj, J. Deu-Pons, F. Muiños, J. Bertl, Q. Guo, A. Gonzalez-Perez, Q. Xiang, W. Bazant, E. Barrera, S.T. Al-Sedairy, A. Aretz, C. Bell, M. Betancourt, C. Buchholz, F. Calvo, C. Chomienne, M. Dunn, S. Edmonds, E. Green, S. Gupta, C.M. Hutter, K. Jegalian, N. Jones, Y. Lu, H. Nakagama, G. Nettekoven, L. Planko, D. Scott, T. Shibata, K. Shimizu, M.R. Stratton, T. Yugawa, G. Tortora, K. VijayRaghavan, J.C. Zenklusen, D. Townend, B. Aminou, J. Bartolome, K.A. Boroevich, R. Boyce, A. Buchanan, N.J. Byrne, Z. Chen, S. Cho, W. Choi, P. Clapham, M.T. Dow, L.J. Dursi, J. Eils, C. Farcas, N. Fayzullaev, P. Flicek, A.P. Heath, O. Hofmann, J.H. Hong, T.J. Hudson, D. Hübschmann, S. Ivkovic, S.H. Jeon, W. Jiao, R. Kabbe, A. Kahles, J.N.A. Kerssemakers, H. Kim, J. Kim, M. Koscher, A. Koures, M. Kovacevic, C. Lawerenz, J. Liu, S. Mijalkovic, A.M. Mijalkovic-Lazic, S. Miyano, M. Nastic, J. Nicholson, D. Ocana, K. Ohi, L. Ohno-Machado, T.D. Pihl, M. Prinz, P. Radovic, C. Short, H.J. Sofia, J. Spring, A.J. Struck, N. Tijanic, D. Vicente, Z. Wang, A. Williams, Y. Woo, A.J. Wright, L. Yang, M.P. Hamilton, T.A. Johnson, A. Kahraman, M. Kellis, P. Polak, R. Sallari, N. Sinnott-Armstrong, C. von Mering, S. Beltran, D.S. Gerhard, M. Gut, J.R. Trotta, J.P. Whalley, B. Niu, S.M.G. Espiritu, S. Gao, Y. Huang, C.M. Lalansingh, J.W. Teague, M.C. Wendl, F. Abascal, G.D. Bader, P. Bandopadhayay, J. Barenboim, S. Brunak, J. Carlevaro-Fita, D. Chakravarty, C.W.Y. Chan, J.K. Choi, K. Diamanti, J.L. Fink, J. Frigola, C. Gambacorti-Passerini, D.W. Garsed, N.J. Haradhvala, A.O. Harmanci, M. Helmy, C. Herrmann, A. Hobolth, E. Hodzic, C. Hong, K. Isaev, J.M.G. Izarzugaza, R. Johnson, R.I. Juul, J. Kim, J.K. Kim, Jan Komorowski, A. Lanzós, E. Larsson, D. Lee, S. Li, X. Li, Z. Lin, E.M. Liu, L. Lochovsky, S. Lou, T. Madsen, K. Marchal, A. Martinez-Fundichely, P.D. McGillivray, W. Meyerson, M. Paczkowska, K. Park, K. Park, T. Pons, S. Pulido-Tamayo, I. Reyes-Salazar, M.A. Reyna, M.A. Rubin, L. Salichos, C. Sander, S.E. Schumacher, M. Shackleton, C. Shen, R. Shrestha, S. Shuai, T. Tsunoda, H.M. Umer, L. Uusküla-Reimand, L.P.C. Verbeke, C. Wadelius, L. Wadi, J. Warrell, G. Wu, J. Yu, J. Zhang, X. Zhang, Y. Zhang, Z. Zhao, L. Zou, M.S. Lawrence, B.J. Raphael, P.J. Bailey, D. Craft, M.J. Goldman, H. Aburatani, H. Binder, H.Q. Dinh, S.C. Heath, S. Hoffmann, C.D. Imbusch, H. Kretzmer, P.W. Laird, J.I. Martin-Subero, G. Nagae, H. Shen, Q. Wang, D. Weichenhan, W. Zhou, B.P. Berman, B. Brors, C. Plass, K.C. Akdemir, D.D.L. Bowtell, K.H. Burns, J. Busanovich, K. Chan, A. Dueso-Barroso, P.A. Edwards, D. Etemadmoghadam, J.E. Haber, D.T.W. Jones, Y.S. Ju, M.D. Kazanov, Y. Koh, K. Kumar, E.A. Lee, J.J.K. Lee, A.G. Lynch, G. Macintyre, F. Markowetz, F.C.P. Navarro, J. V. Pearson, K. Rippe, R. Scully, I. Villasante, N. Waddell, L. Yang, X. Yao, S.S. Yoon, C.Z. Zhang, E.N. Bergstrom, A. Boot, K. Covington, A. Fujimoto, M. N. Huang, S.M.A. Islam, J.R. McPherson, S. Morganella, V. Mustonen, A.W.T. Ng, S. D. Prokopec, I. Vázquez-García, Y. Wu, F. Yousif, W. Yu, S.G. Rozen, V.A. Rudneva, S.S. Shringarpure, D.J. Turner, T. Xia, G. Atwal, D.K. Chang, S.L. Cooke, B.M. Faltas, S. Haider, V.B. Kaiser, R. Karlić, M. Kato, K. Kübler, A. Margolin, S. Martin, S. Nik-Zainal, C. P'ng, C.A. Semple, J. Smith, R.X. Sun, K. Thai, D.W. Wright, K. Yuan, A. V. Biankin, L. Garraway, S.M. Grimmond, D.J. Adams, P. Anur, S. Cao, E. L. Christie, M. Cmero, Y. Cun, K.J. Dawson, S.C. Dentro, A.G. Deshwar, N. Donmez, R.M. Drews, M. Gerstung, G. Ha, K. Haase, L. Jerman, Y. Ji, C. Jolly, J. Lee, H. Lee-Six, S. Malikic, T.J. Mitchell, Q.D. Morris, L. Oesper, M. Peifer, M. Peto, D. Rosebrock, Y. Rubanova, A. Salcedo, S. Sengupta, R. Shi, S.J. Shin, O. Spiro, S. Vembu, J.A. Wintersinger, T.P. Yang, K. Yu, H. Zhu, P.T. Spellman, Y. Chen, M. Fujita, L. Han, T. Hasegawa, M. Komura, J. Li, S. Mizuno, E. Shimizu, Y. Wang, Y. Xu, R. Yamaguchi, F. Yang, Y. Yang, C.J. Yoon, Y. Yuan, H. Liang, M. Alawi, I. Borozan, D.S. Brewer, C.S. Cooper, N. Desai, A. Grundhoff, M. Iskar, X. Su, M. Zapatka, P. Lichter, K. Alsop, T.J.C. Bruxner, A.N. Christ, S.M. Cordner, P.A. Cowin, R. Drapkin, S. Fereday, J. George, A. Hamilton, O. Holmes, J.A. Hung, K.S. Kassahn, S.H. Kazakoff, C.J. Kennedy, C.R. Leonard, L. Mileshkin, D.K. Miller, G.M. Arnau, C. Mitchell, F. Newell, K. Nones, A.M. Patch, M.C. Quinn, D.F. Taylor, H. Thorne, N. Traficante, R. Vedururu, N.M. Waddell, P.M. Waring, S. Wood, Q. Xu, A. deFazio, M.J. Anderson, D. Antonello, A.P. Barbour, C. Bassi, S. Bersani, I. Cataldo, L.A. Chantrill, Y.E. Chiew, A. Chou, S. Cingarlini, N. Cloonan, V. Corbo, M.V. Davi, F.R. Duthie, A.J. Gill, J.S. Graham, I. Harliwong, N.B. Jamieson, A.L. Johns, J.G. Kench, L. Landoni, R.T. Lawlor, A. Mafficini, N.D. Merrett, M. Miotto, E.A. Musgrove, A.M. Nagrial, K.A. Oien, M. Pajic, M. Pinese, A.J. Robertson, I. Rooman, B.C. Rusev, J.S. Samra, M. Scardoni, C.J. Scarlett, A. Scarpa, E. Sereni, K.O. Sikora, M. Simbolo, M.L. Taschuk, C.W. Toon, C. Vicentini, J. Wu, N. Zeps, A. Behren, H. Burke, J. Cebon, R.A. Dagg, R. De Paoli-Iseppi, K. Dutton-Regester, M.A. Field, A. Fitzgerald, P. Hersey, V. Jakrot, P.A. Johansson, H. Kakavand, R.F. Kefford, L.M.S. Lau, G. V. Long, H.A. Pickett, A.L. Pritchard, G.M. Pupo, R.P.M. Saw, S.J. Schramm, C.A. Shang, P. Shang, A.J. Spillane, J.R. Stretch, V. Tembe, J.F. Thompson, R.E. Vilain, J.S. Wilmott, J.Y. Yang, N.K. Hayward, G.J. Mann, R.A. Scolyer, J. Bartlett, P. Bavi, D.E. Chadwick, M. Chan-Seng-Yue, S. Cleary, A.A. Connor, K. Czajka, R.E. Denroche, N.C. Dhani, J. Eagles, S. Gallinger, R.C. Grant, D. Hedley, M.A. Hollingsworth, G.H. Jang, J. Johns, S. Kalimuthu, S. Ben Liang, I. Lungu, X. Luo, F. Mbabaali, T.A. McPherson, J.K. Miller, M.J. Moore, F. Notta, D. Pasternack, G.M. Petersen, M.H.A. Roehrl, M. Sam, I. Selander, S. Serra, S. Shahabi, S.P. Thayer, L.E. Timms, G.W. Wilson, J.M. Wilson, B.G. Wouters, J.D. McPherson, T.A. Beck, V. Bhandari, C.C. Collins, N.E. Fleshner, N.S. Fox, M. Fraser, L.E. Heisler, E. Lalonde, J. Livingstone, A. Meng, V.Y. Sabelnykova, Y.J. Shiah, T. Van der Kwast, R.G. Bristow, S. Ding, D. Fan, L. Li, Y. Nie, X. Xiao, R. Xing, S. Yang, Y. Yu, Y. Zhou, R.E. Banks, G. Bourque, P. Brennan, L. Letourneau, Y. Riazalhosseini, G. Scelo, N. Vasudev, J. Viksna, M. Lathrop, J. Tost, S.M. Ahn, S. Aparicio, L. Arnould, M.R. Aure, S.G. Bhosle, E. Birney, A. Borg, S. Boyault, A.B. Brinkman, J.E. Brock, A. Broeks, A.L. Børresen-Dale, C. Caldas, S.F. Chin, H. Davies, C. Desmedt, L. Dirix, S. Dronov, A. Ehinger, J.E. Eyfjord, A. Fatima, J.A. Foekens, P.A. Futreal, Ø. Garred, D.D. Giri, D. Glodzik, D. Grabau, H. Hilmarsdottir, G.K. Hooijer, J. Jacquemier, S.J. Jang, J.G. Jonasson, J. Jonkers, H.Y. Kim, T.A. King, S. Knappskog, G. Kong, S. Krishnamurthy, S.R. Lakhani, A. Langerød, D. Larsimont, H.J. Lee, J.Y. Lee, M.T.M. Lee, O.C. Lingjærde, G. MacGrogan, J.W.M. Martens, S. O'Meara, I. Pauporté, S. Pinder, X. Pivot, E. Provenzano, C.A. Purdie, M. Ramakrishna, K. Ramakrishnan, J. Reis-Filho, A.L. Richardson, M. Ringnér, J.B. Rodriguez, F.G. Rodríguez-González, G. Romieu, R. Salgado, T. Sauer, R. Shepherd, A.M. Sieuwerts, P.T. Simpson, M. Smid, C. Sotiriou, P.N. Span, Ó.A. Stefánsson, A. Stenhouse, H.G. Stunnenberg, F. Sweep, B.K.T. Tan, G. Thomas, A.M. Thompson, S. Tommasi, I. Treilleux, A. Tutt, N.T. Ueno, S. Van Laere, G.G. Van den Eynden, P. Vermeulen, A. Viari, A. Vincent-Salomon, B.H. Wong, L. Yates, X. Zou, C.H.M. van Deurzen, M.J. van de Vijver, L. van't Veer, O. Ammerpohl, S. Aukema, A.K. Bergmann, S.H. Bernhart, A. Borkhardt, C. Borst, B. Burkhardt, A. Claviez, M.E. Goebler, A. Haake, S. Haas, M. Hansmann, J.I. Hoell, M. Hummel, D. Karsch, W. Klapper, M. Kneba, M. Kreuz, D.

Kube, R. Küppers, D. Lenze, M. Loeffler, C. López, L. Mantovani-Löffler, P. Möller, G. Ott, B. Radlwimmer, J. Richter, M. Rohde, P.C. Rosenstiel, A. Rosenwald, M.B. Schilhabel, S. Schreiber, P.F. Stadler, P. Staib, S. Stilgenbauer, S. Sungalee, M. Szczepanowski, U.H. Toprak, L.H.P. Trümper, R. Wagener, T. Zenz, V. Hovestadt, C. von Kalle, M. Kool, A. Korshunov, P. Landgraf, H. Lehrach, P.A. Northcott, S.M. Pfister, G. Reifenberger, H.J. Warnatz, S. Wolf, M.L. Yaspo, Y. Assenov, C. Gerhauser, S. Minner, T. Schlomm, R. Simon, G. Sauter, H. Sültmann, N.K. Biswas, A. Maitra, P.P. Majumder, R. Sarin, S. Barbi, G. Bonizzato, C. Cantù, A.P. Dei Tos, M. Fassan, S. Grimaldi, C. Luchini, G. Malleo, G. Marchegiani, M. Milella, S. Paiella, A. Pea, P. Pederzoli, A. Ruzzenente, R. Salvia, N. Sperandio, Y. Arai, N. Hama, N. Hiraoka, F. Hosoda, H. Nakamura, H. Ojima, T. Okusaka, Y. Totoki, T. Urushidate, M. Fukayama, S. Ishikawa, H. Katai, H. Katoh, D. Komura, H. Rokutan, M. Saito-Adachi, A. Suzuki, H. Taniguchi, K. Tatsuno, T. Ushiku, S. Yachida, S. Yamamoto, H. Aikata, K. Arihiro, S. ichi Ariizumi, K. Chayama, M. Furuta, K. Gotoh, S. Hayami, S. Hirano, Y. Kawakami, K. Maejima, T. Nakamura, K. Nakano, H. Ohdan, A. Sasaki-Oku, H. Tanaka, M. Ueno, M. Yamamoto, H. Yamaue, S.P. Choo, I. Cutcutache, N. Khuntikeo, C.K. Ong, C. Pairojkul, I. Popescu, K.S. Ahn, M. Aymerich, A. Lopez-Guillermo, C. López-Otín, X.S. Puente, E. Campo, F. Amary, D. Baumhoer, S. Behjati, B. Bjerkehagen, P.A. Futreal, O. Myklebost, N. Pillay, P. Tarpey, R. Tirabosco, O. Zaikova, A.M. Flanagan, J. Boultwood, D.T. Bowen, M. Cazzola, A.R. Green, E. Hellstrom-Lindberg, L. Malcovati, J. Nangalia, E. Papaemmanuil, P. Vyas, Y. Ang, H. Barr, D. Beardsmore, M. Eldridge, J. Gossage, N. Grehan, G.B. Hanna, S.J. Hayes, T.R. Hupp, D. Khoo, J. Lagergren, L.B. Lovat, S. MacRae, M. O'Donovan, J.R. O'Neill, S.L. Parsons, S.R. Preston, S. Puig, T. Roques, G. Sanders, S. Sothi, S. Tavaré, O. Tucker, R. Turkington, T.J. Underwood, I. Welch, R.C. Fitzgerald, D.M. Berney, J.S. De Bono, D. Cahill, N. Camacho, N.M. Dennis, T. Dudderidge, S.E. Edwards, C. Fisher, C.S. Foster, M. Ghori, P. Gill, V.J. Gnanapragasam, G. Gundem, F.C. Hamdy, S. Hawkins, S. Hazell, W. Howat, W.B. Isaacs, K. Karaszi, J.D. Kay, V. Khoo, Z. Kote-Jarai, B. Kremeyer, P. Kumar, A. Lambert, D.A. Leongamornlert, N. Livni, Y.J. Lu, H.J. Luxton, L. Marsden, C.E. Massie, L. Matthews, E. Mayer, U. McDermott, S. Merson, D.E. Neal, A. Ng, D. Nicol, C. Ogden, E.W. Rowe, N.C. Shah, S. Thomas, A. Thompson, C. Verrill, T. Visakorpi, A.Y. Warren, H.C. Whitaker, H. Zhang, N. van As, R.A. Eeles, A. Abeshouse, N. Agrawal, R. Akbani, H. Al-Ahmadie, M. Albert, K. Aldape, A. Ally, E. L. Appelbaum, J. Armenia, S. Asa, J.T. Auman, M. Balasundaram, S. Balu, J. Barnholtz-Sloan, O.F. Bathe, S.B. Baylin, C. Benz, A. Berchuck, M. Berrios, D. Bigner, M. Birrer, T. Bodenheimer, L. Boice, M.S. Bootwalla, M. Bosenberg, R. Bowlby, J. Boyd, R.R. Broaddus, M. Brock, D. Brooks, S. Bullman, S.J. Caesar-Johnson, T.E. Carey, R. Carlsen, R. Cerfolio, V.S. Chandan, H.W. Chen, A.D. Cherniack, J. Chien, J. Cho, E. Chuah, C. Cibulskis, L. Cope, M.G. Cordes, E. Curley, B. Czerniak, L. Danilova, I.J. Davis, T. Defreitas, J.A. Demchok, N. Dhalla, R. Dhir, H.V. Doddapaneni, A. El-Naggar, I. Felau, M.L. Ferguson, G. Finocchiaro, K.M. Fong, S. Frazer, W. Friedman, C.C. Fronick, L.A. Fulton, S.B. Gabriel, J. Gao, N. Gehlenborg, J.E. Gershenwald, R. Ghossein, N.H. Giama, R.A. Gibbs, C. Gomez, R. Govindan, D.N. Hayes, A.M. Hegde, D.I. Heiman, Z. Heins, A.J. Hepperla, A. Holbrook, R.A. Holt, A.P. Hoyle, R.H. Hruban, J. Hu, M. Huang, D. Huntsman, J. Huse, C.A. Iacobuzio-Donahue, M. Ittmann, J.C. Jayaseelan, S.R. Jefferys, C.D. Jones, S.J.M. Jones, H. Juhl, K.J. Kang, B. Karlan, K. Kasaian, E. Kebebew, H.K. Kim, V. Korchina, R. Kundra, P.H. Lai, E. Lander, X. Le, D. Lee, D.A. Levine, L. Lewis, T. Ley, H.I. Li, P. Lin, W.M. Linehan, F.F. Liu, Y. Lu, L. Lype, Y. Ma, D.T. Maglinte, E.R. Mardis, J. Marks, M.A. Marra, T.J. Matthew, M. Mayo, K. McCune, S. R. Meier, S. Meng, P.A. Mieczkowski, T. Mikkelsen, C.A. Miller, G.B. Mills, R.A. Moore, C. Morrison, L.E. Mose, C.D. Moser, A.J. Mungall, K. Mungall, D. Mutch, D. M. Muzny, J. Myers, Y. Newton, M.S. Noble, P. O'Donnell, B.P. O'Neill, A. Ochoa, J.W. Park, J.S. Parker, H. Pass, A. Pastore, N.A. Pennell, C.M. Perou, N. Petrelli, O. Potapova, J.S. Rader, S. Ramalingam, W.K. Rathmell, V. Reuter, S.M. Reynolds, M. Ringel, J. Roach, L.R. Roberts, A.G. Robertson, S. Sadeghi, C. Saller, F. Sanchez-Vega, D. Schadendorf, J.E. Schein, H.K. Schmidt, N. Schultz, R. Seethala, Y. Senbabaoglu, T. Shelton, Y. Shi, J. Shih, I. Shmulevich, C. Shriver, S. Signoretti, J. V. Simons, S. Singer, P. Sipahimalani, T.J. Skelly, K. Smith-McCune, N.D. Socci, M. G. Soloway, A.K. Sood, A. Tam, D. Tan, R. Tarnuzzer, N. Thiessen, R.H. Thompson, L.B. Thorne, M. Tsao, C. Umbricht, D.J. Van Den Berg, E.G. Van Meir, U. Veluvolu, D. Voet, L. Wang, P. Weinberger, D.J. Weisenberger, D. Wigle, M.D. Wilkerson, R. K. Wilson, B. Winterhoff, M. Wiznerowicz, T. Wong, W. Wong, L. Xi, C. Yau, H. Zhang, H. Zhang, J. Zhang, Pan-cancer analysis of whole genomes, Nature. 578 (2020). https://doi.org/10.1038/s41586-020-1969-6.

[7] J.R. Pon, M.A. Marra, Driver and passenger mutations in cancer, Annu. Rev. Pathol. 10 (2015), https://doi.org/10.1146/annurev-pathol-012414-040312.

[8] M.S. Lawrence, P. Stojanov, P. Polak, G.V. Kryukov, K. Cibulskis, A. Sivachenko, S. L. Carter, C. Stewart, C.H. Mermel, S.A. Roberts, A. Kiezun, P.S. Hammerman, A. McKenna, Y. Drier, L. Zou, A.H. Ramos, T.J. Pugh, N. Stransky, E. Helman, J. Kim, C. Sougnez, L. Ambrogio, E. Nickerson, E. Shefler, M.L. Cortés, D. Auclair, G. Saksena, D. Voet, M. Noble, D. DiCara, P. Lin, L. Lichtenstein, D.I. Heiman, T. Fennell, M. Imielinski, B. Hernandez, E. Hodis, S. Baca, A.M. Dulak, J. Lohr, D.-A. Landau, C.J. Wu, J. Melendez-Zajgla, A. Hidalgo-Miranda, A. Koren, S. A. McCarroll, J. Mora, B. Crompton, R. Onofrio, M. Parkin, W. Winckler, K. Ardlie, S.B. Gabriel, C.W.M. Roberts, J.A. Biegel, K. Stegmaier, A.J. Bass, L.A. Garraway, M. Meyerson, T.R. Golub, D.A. Gordenin, S. Sunyaev, E.S. Lander, G. Getz, Mutational heterogeneity in cancer and the search for new cancer-associated genes, Nature 499 (2013) 214–218, https://doi.org/10.1038/nature12213.

[9] L. Gerasimavicius, B.J. Livesey, J.A. Marsh, Loss-of-function, gain-of-function and dominant-negative mutations have profoundly different effects on protein structure, Nat Commun. 13 (2022), https://doi.org/10.1038/s41467-022-31686-6.

[10] D. Shorthouse, M.W.J. Hall, B.A. Hall, Computational saturation screen reveals the landscape of mutations in human fumarate hydratase, J. Chem. Inf. Model. 61 (2021) 1970–1980, https://doi.org/10.1021/acs.jcim.1c00063.

[11] D. Shorthouse, B.A. Hall, SARS-CoV-2 variants are selecting for spike protein mutations that increase protein stability, J. Chem. Inf. Model. (2021), https://doi.org/10.1021/acs.jcim.1c00990.

[12] M. Lee, D. Shorthouse, R. Mahen, B.A. Hall, A.R. Venkitaraman, Cancer-causing *BRCA2* missense mutations disrupt an intracellular protein assembly mechanism to disable genome maintenance, Nucleic Acids Res. 49 (2021) 5588–5604, https://doi.org/10.1093/nar/gkab308.

[13] J.C. Fowler, C. King, C. Bryant, M.W.J. Hall, R. Sood, S.H. Ong, E. Earp, D. Fernandez-Antoran, J. Koeppel, S.C. Dentro, D. Shorthouse, A. Durrani, K. Fife, E. Rytina, D. Milne, A. Roshan, K. Mahububani, K. Saeb-Parsy, B.A. Hall, M. Gerstung, P.H. Jones, Selection of oncogenic mutant clones in normal human skin varies with body site, Cancer Discov. 11 (2021) 340–361, https://doi.org/10.1158/2159-8290.CD-20-1092.

[14] M. Li, F.L. Simonetti, A. Goncearenco, A.R. Panchenko, MutaBind estimates and interprets the effects of sequence variants on protein–protein interactions, Nucleic Acids Res. 44 (2016), https://doi.org/10.1093/nar/gkw374.

[15] A. Strokach, T.Y. Lu, P.M. Kim, ELASPIC2 (EL2): Combining contextualized language models and graph neural networks to predict effects of mutations, J. Mol. Biol. 433 (2021), https://doi.org/10.1016/j.jmb.2021.166810.

[16] J. Delgado, L.G. Radusky, D. Cianferoni, L. Serrano, FoldX 5.0: working with RNA, small molecules and a new graphical interface, Bioinformatics 35 (2019) 4168–4169, https://doi.org/10.1093/bioinformatics/btz184.

[17] R.F. Alford, A. Leaver-Fay, J.R. Jeliazkov, M.J. O'Meara, F.P. DiMaio, H. Park, M. V. Shapovalov, P.D. Renfrew, V.K. Mulligan, K. Kappel, J.W. Labonte, M.S. Pacella, R. Bonneau, P. Bradley, R.L. Dunbrack, R. Das, D. Baker, B. Kuhlman, T. Kortemme, J.J. Gray, J.J. Gray, The Rosetta All-atom energy function for macromolecular modeling and design, J. Chem. Theory Comput. 13 (2017) 3031–3048, https://doi.org/10.1021/acs.jctc.7b00125.

[18] D. Seeliger, B.L. de Groot, Protein thermostability calculations using alchemical free energy simulations, Biophys. J. 98 (2010), https://doi.org/10.1016/j.bpj.2010.01.051.

[19] A. Muñiz-Chicharro, L.W. Votapka, R.E. Amaro, R.C. Wade, Brownian dynamics simulations of biomolecular diffusional association processes, Wiley Interdiscip. Rev. Comput. Mol. Sci. 13 (2023), https://doi.org/10.1002/wcms.1649.

[20] R.R. Gabdoulline, R.C. Wade, Protein-protein association: Investigation of factors influencing association rates by Brownian dynamics simulations, J. Mol. Biol. 306 (2001), https://doi.org/10.1006/jmbi.2000.4404.

[21] M.W.J. Hall, D. Shorthouse, R. Alcraft, P.H. Jones, B.A. Hall, Mutations observed in somatic evolution reveal underlying gene mechanisms, Commun. Biol. 6 (2023) 753, https://doi.org/10.1038/s42003-023-05136-y.

[22] A. Strokach, C. Corbi-Verge, P.M. Kim, Predicting changes in protein stability caused by mutation using sequence-and structure-based methods in a CAGI5 blind challenge, Hum. Mutat. 40 (2019) 1414–1423, https://doi.org/10.1002/humu.23852.

[23] C. Geng, L.C. Xue, J. Roel-Touris, A.M.J.J. Bonvin, Finding the ΔΔG spot: Are predictors of binding affinity changes upon mutations in protein–protein interactions ready for it? Wiley Interdiscip. Rev. Comput. Mol. Sci. 9 (2019) https://doi.org/10.1002/wcms.1410.

[24] O. Buß, J. Rudat, K. Ochsenreither, FoldX as protein engineering tool: Better than random based approaches? Comput. Struct. Biotechnol. J. 16 (2018) 25–33, https://doi.org/10.1016/J.CSBJ.2018.01.002.

[25] R. Yin, J.D. Guest, G. Taherzadeh, R. Gowthaman, I. Mittra, J. Quackenbush, B. G. Pierce, Structural and energetic profiling of SARS-CoV-2 receptor binding domain antibody recognition and the impact of circulating variants, PLoS Comput. Biol. 17 (2021), https://doi.org/10.1371/journal.pcbi.1009380.

[26] M.W. Maciejewski, A.D. Schuyler, M.R. Gryk, I.I. Moraru, P.R. Romero, E.L. Ulrich, H.R. Eghbalnia, M. Livny, F. Delaglio, J.C. Hoch, NMRbox: A resource for biomolecular NMR computation, Biophys. J. 112 (2017), https://doi.org/10.1016/j.bpj.2017.03.011.

[27] L.G. Radusky, L. Serrano, PyFoldX: enabling biomolecular analysis and engineering along structural ensembles, Bioinformatics 38 (2022), https://doi.org/10.1093/bioinformatics/btac072.

[28] P.J.A. Cock, T. Antao, J.T. Chang, B.A. Chapman, C.J. Cox, A. Dalke, I. Friedberg, T. Hamelryck, F. Kauff, B. Wilczynski, M.J.L. de Hoon, Biopython: freely available Python tools for computational molecular biology and bioinformatics, Bioinformatics 25 (2009) 1422–1423, https://doi.org/10.1093/bioinformatics/btp163.

[29] C.R. Harris, K.J. Millman, S.J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N.J. Smith, R. Kern, M. Picus, S. Hoyer, M.H. van Kerkwijk, M. Brett, A. Haldane, J.F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T.E. Oliphant, Array programming with NumPy, Nature 585 (2020), https://doi.org/10.1038/s41586-020-2649-2.

[30] P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K.J. Millman, N. Mayorov, A.R.J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E.W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E.A. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, A. Vijaykumar, A. Pietro Bardelli, A. Rothberg, A. Hilboll, A. Kloeckner, A. Scopatz, A. Lee, A. Rokem, C.N. Woods, C. Fulton, C. Masson, C. Häggström, C. Fitzgerald, D.A. Nicholson, D.R. Hagen, D. V. Pasechnik, E. Olivetti, E. Martin, E. Wieser, F. Silva, F. Lenders, F. Wilhelm, G. Young, G.A. Price, G.L. Ingold, G.E. Allen, G.R. Lee, H. Audren, I. Probst, J. P. Dietrich, J. Silterra, J.T. Webber, J. Slavič, J. Nothman, J. Buchner, J. Kulick, J. L. Schönberger, J.V. de Miranda Cardoso, J. Reimer, J. Harrington, J.L. C. Rodríguez, J. Nunez-Iglesias, J. Kuczynski, K. Tritz, M. Thoma, M. Newville,

M. Kümmerer, M. Bolingbroke, M. Tartre, M. Pak, N.J. Smith, N. Nowaczyk, N. Shebanov, O. Pavlyk, P.A. Brodtkorb, P. Lee, R.T. McGibbon, R. Feldbauer, S. Lewis, S. Tygier, S. Sievert, S. Vigna, S. Peterson, S. More, T. Pudlik, T. Oshima, T.J. Pingel, T.P. Robitaille, T. Spura, T.R. Jones, T. Cera, T. Leslie, T. Zito, T. Krauss, U. Upadhyay, Y.O. Halchenko, Y. Vázquez-Baeza, SciPy 1.0: fundamental algorithms for scientific computing in Python, Nat. Methods 17 (2020), https://doi.org/10.1038/s41592-019-0686-2.

[31] J.D. Hunter, Matplotlib: A 2D graphics environment, Comput. Sci. Eng. 9 (2007) 90–95, https://doi.org/10.1109/MCSE.2007.55.

[32] M. Waskom, seaborn: statistical data visualization, J. Open Source Softw. 6 (2021), https://doi.org/10.21105/joss.03021.

[33] C. Davidson-Pilon, lifelines: survival analysis in Python, J. Open Source Softw. 4 (2019), https://doi.org/10.21105/joss.01317.

[34] G. Schreiber, A.R. Fersht, Energetics of protein-protein interactions: Analysis ofthe Barnase-Barstar interface by single mutations and double mutant cycles, J. Mol. Biol. 248 (1995), https://doi.org/10.1016/S0022-2836(95)80064-6.

[35] G. Schreiber, A.R. Fersht, Interaction of Barnase with its polypeptide inhibitor barstar studied by protein engineering, Biochemistry 32 (1993), https://doi.org/10.1021/bi00070a025.

[36] A.M. Buckle, G. Schreiber, A.R. Fersht, Protein-protein recognition: Crystal structural analysis of a barnase-barstar complex at 2.0-Å resolution, Biochemistry 33 (1994), https://doi.org/10.1021/bi00196a004.

[37] M. Jucovic, R.W. Hartley, Protein-protein interaction: A genetic selection for compensating mutations at the barnase-barstar interface, Proc. Natl. Acad. Sci. U. S. A. 93 (1996), https://doi.org/10.1073/pnas.93.6.2343.

[38] A. Clark, M. Burleson, SPOP and cancer: a systematic review, Am. J. Cancer Res. 10 (2020).

[39] H. Janouskova, G. El Tekle, E. Bellini, N.D. Udeshi, A. Rinaldi, A. Ulbricht, T. Bernasocchi, G. Civenni, M. Losa, T. Svinkina, C.M. Bielski, G.V. Kryukov, L. Cascione, S. Napoli, R.I. Enchev, D.G. Mutch, M.E. Carney, A. Berchuck, B.J. N. Winterhoff, R.R. Broaddus, P. Schraml, H. Moch, F. Bertoni, C.V. Catapano, M. Peter, S.A. Carr, L.A. Garraway, P.J. Wild, J.P.P. Theurillat, Opposing effects of cancer-Type-specific SPOP mutants on BET protein degradation and sensitivity to BET inhibitors, Nat. Med. 23 (2017) 1046–1054, https://doi.org/10.1038/nm.4372.

[40] K.A. Hoadley, C. Yau, T. Hinoue, D.M. Wolf, A.J. Lazar, E. Drill, R. Shen, A.M. Taylor, A.D. Cherniack, V. Thorsson, R. Akbani, R. Bowlby, C.K. Wong, M. Wiznerowicz, F. Sanchez-Vega, A.G. Robertson, B.G. Schneider, M.S. Lawrence, H. Noushmehr, T.M. Malta, S.J. Caesar-Johnson, J.A. Demchok, I. Felau, M. Kasapi, M.L. Ferguson, C.M. Hutter, H.J. Sofia, R. Tarnuzzer, Z. Wang, L. Yang, J.C. Zenklusen, J. (Julia) Zhang, S. Chudamani, J. Liu, L. Lolla, R. Naresh, T. Pihl, Q. Sun, Y. Wan, Y. Wu, J. Cho, T. DeFreitas, S. Frazer, N. Gehlenborg, G. Getz, D.I. Heiman, J. Kim, M.S. Lawrence, P. Lin, S. Meier, M.S. Noble, G. Saksena, D. Voet, H. Zhang, B. Bernard, N. Chambwe, V. Dhankani, T. Knijnenburg, R. Kramer, K. Leinonen, Y. Liu, M. Miller, S. Reynolds, I. Shmulevich, V. Thorsson, W. Zhang, R. Akbani, B.M. Broom, A.M. Hegde, Z. Ju, R.S. Kanchi, A. Korkut, J. Li, H. Liang, S. Ling, W. Liu, Y. Lu, G.B. Mills, K.-S. Ng, A. Rao, M. Ryan, J. Wang, J.N. Weinstein, J. Zhang, A. Abeshouse, J. Armenia, D. Chakravarty, W.K. Chatila, I. de Bruijn, J. Gao, B.E. Gross, Z.J. Heins, R. Kundra, K. La, M. Ladanyi, A. Luna, M.G. Nissan, A. Ochoa, S.M. Phillips, E. Reznik, F. Sanchez-Vega, C. Sander, N. Schultz, R. Sheridan, S.O. Sumer, Y. Sun, B.S. Taylor, J. Wang, H. Zhang, P. Anur, M. Peto, P. Spellman, C. Benz, J.M. Stuart, C.K. Wong, C. Yau, D.N. Hayes, J.S. Parker, M.D. Wilkerson, A. Ally, M. Balasundaram, R. Bowlby, D. Brooks, R. Carlsen, E. Chuah, N. Dhalla, R. Holt, S.J.M. Jones, K. Kasaian, D. Lee, Y. Ma, M.A. Marra, M. Mayo, R. A. Moore, A.J. Mungall, K. Mungall, A.G. Robertson, S. Sadeghi, J.E. Schein, P. Sipahimalani, A. Tam, N. Thiessen, K. Tse, T. Wong, A.C. Berger, R. Beroukhim, A. D. Cherniack, C. Cibulskis, S.B. Gabriel, G.F. Gao, G. Ha, M. Meyerson, S.E. Schumacher, J. Shih, M.H. Kucherlapati, R.S. Kucherlapati, S. Baylin, L. Cope, L. Danilova, M.S. Bootwalla, P.H. Lai, D.T. Maglinte, D.J. Van Den Berg, D.J. Weisenberger, J.T. Auman, S. Balu, T. Bodenheimer, C. Fan, K.A. Hoadley, A.P. Hoyle, S.R. Jefferys, C.D. Jones, S. Meng, P.A. Mieczkowski, L.E. Mose, A.H. Perou, C.M. Perou, J. Roach, Y. Shi, J. V. Simons, T. Skelly, M.G. Soloway, D. Tan, U. Veluvolu, H. Fan, T. Hinoue, P.W. Laird, H. Shen, W. Zhou, M. Bellair, K. Chang, K. Covington, C.J. Creighton, H. Dinh, H. Doddapaneni, L.A. Donehower, J. Drummond, R.A. Gibbs, R. Glenn, W. Hale, Y. Han, J. Hu, V. Korchina, S. Lee, L. Lewis, W. Li, X. Liu, M. Morgan, D. Morton, D. Muzny, J. Santibanez, M. Sheth, E. Shinbrot, L. Wang, M. Wang, D.A. Wheeler, L. Xi, F. Zhao, J. Hess, E.L. Appelbaum, M. Bailey, M.G. Cordes, L. Ding, C.C. Fronick, L.A. Fulton, R.S. Fulton, C. Kandoth, E.R. Mardis, M.D. McLellan, C.A. Miller, H.K. Schmidt, R.K. Wilson, D. Crain, E. Curley, J. Gardner, K. Lau, D. Mallery, S. Morris, J. Paulauskis, R. Penny, C. Shelton, T. Shelton, M. Sherman, E. Thompson, P. Yena, J. Bowen, J.M. Gastier-Foster, M. Gerken, K.M. Leraas, T.M. Lichtenberg, N.C. Ramirez, L. Wise, E. Zmuda, N. Corcoran, T. Costello, C. Hovens, A.L. Carvalho, A.C. de Carvalho, J.H. Fregnani, A. Longatto-Filho, R.M. Reis, C. Scapulatempo-Neto, H.C.S. Silveira, D.O. Vidal, A. Burnette, J. Eschbacher, B. Hermes, A. Noss, R. Singh, M.L. Anderson, P.D. Castro, M. Ittmann, D. Huntsman, B. Kohl, X. Le, R. Thorp, C. Andry, E.R. Duffy, V. Lyadov, O. Paklina, G. Setdikova, A. Shabunin, M. Tavobilov, C. McPherson, R. Warnick, R. Berkowitz, D. Cramer, C. Feltmate, N. Horowitz, A. Kibel, M. Muto, C.P. Raut, A. Malykh, J.S. Barnholtz-Sloan, W. Barrett, K. Devine, J. Fulop, Q.T. Ostrom, K. Shimmel, Y. Wolinsky, A.E. Sloan, A. De Rose, F. Giuliante, M. Goodman, B.Y. Karlan, C.H. Hagedorn, J. Eckman, J. Harr, J. Myers, K. Tucker, L.A. Zach, B. Deyarmin, H. Hu, L. Kvecher, C. Larson, R.J. Mural, S. Somiari, A. Vicha, T. Zelinka, J. Bennett, M. Iacocca, B. Rabeno, P. Swanson, M. Latour, L. Lacombe, B. Têtu, A. Bergeron, M. McGraw, S.M. Staugaitis, J. Chabot, H. Hibshoosh, A. Sepulveda, T. Su, T. Wang, O. Potapova, O. Voronina, L. Desjardins, O. Mariani, S. Roman-Roman, X. Sastre, M.-H. Stern, F. Cheng, S. Signoretti, A. Berchuck, D. Bigner, E. Lipp, J. Marks, S. McCall, R. McLendon, A. Secord, A. Sharp, M. Behera, D.J. Brat, A. Chen, K. Delman, S. Force, F. Khuri, K. Magliocca, S. Maithel, J.J. Olson, T. Owonikoko, A. Pickens, S. Ramalingam, D.M. Shin, G. Sica, E.G. Van Meir, H. Zhang, W. Eijckenboom, A. Gillis, E. Korpershoek, L. Looijenga, W. Oosterhuis, H. Stoop, K.E. van Kessel, E.C. Zwarthoff, C. Calatozzolo, L. Cuppini, S. Cuzzubbo, F. DiMeco, G. Finocchiaro, L. Mattei, A. Perin, B. Pollo, C. Chen, J. Houck, P. Lohavanichbutr, A. Hartmann, C. Stoehr, R. Stoehr, H. Taubert, S. Wach, B. Wullich, W. Kycler, D. Murawa, M. Wiznerowicz, K. Chung, W.J. Edenfield, J. Martin, E. Baudin, G. Bubley, R. Bueno, A. De Rienzo, W.G. Richards, S. Kalkanis, T. Mikkelsen, H. Noushmehr, L. Scarpace, N. Girard, M. Aymerich, E. Campo, E. Giné, A.L. Guillermo, N. Van Bang, P.T. Hanh, B.D. Phu, Y. Tang, H. Colman, K. Evason, P.R. Dottino, J.A. Martignetti, H. Gabra, H. Juhl, T. Akeredolu, S. Stepa, D. Hoon, K. Ahn, K.J. Kang, F. Beuschlein, A. Breggia, M. Birrer, D. Bell, M. Borad, A. H. Bryce, E. Castle, V. Chandan, J. Cheville, J.A. Copland, M. Farnell, T. Flotte, N. Giama, T. Ho, M. Kendrick, J.-P. Kocher, K. Kopp, C. Moser, D. Nagorney, D. O'Brien, B.P. O'Neill, T. Patel, G. Petersen, F. Que, M. Rivera, L. Roberts, R. Smallridge, T. Smyrk, M. Stanton, R.H. Thompson, M. Torbenson, J.D. Yang, L. Zhang, F. Brimo, J.A. Ajani, A.M.A. Gonzalez, C. Behrens, olanta Bondaruk, R. Broaddus, B. Czerniak, B. Esmaeli, J. Fujimoto, J. Gershenwald, C. Guo, A.J. Lazar, C. Logothetis, F. Meric-Bernstam, C. Moran, L. Ramondetta, D. Rice, A. Sood, P. Tamboli, T. Thompson, P. Troncoso, A. Tsao, I. Wistuba, C. Carter, L. Haydu, P. Hersey, V. Jakrot, H. Kakavand, R. Kefford, K. Lee, G. Long, G. Mann, M. Quinn, R. Saw, R. Scolyer, K. Shannon, A. Spillane, J. Stretch, M. Synott, J. Thompson, J. Wilmott, H. Al-Ahmadie, T.A. Chan, R. Ghossein, A. Gopalan, D.A. Levine, V. Reuter, S. Singer, B. Singh, N.V. Tien, T. Broudy, C. Mirsaidi, P. Nair, P. Drwiega, J. Miller, J. Smith, H. Zaren, J.-W. Park, N.P. Hung, E. Kebebew, W.M. Linehan, A.R. Metwalli, K. Pacak, P.A. Pinto, M. Schiffman, L.S. Schmidt, C.D. Vocke, N. Wentzensen, R. Worrell, H. Yang, M. Moncrieff, C. Goparaju, J. Melamed, H. Pass, N. Botnariuc, I. Caraman, M. Cernat, I. Chemencedji, A. Clipca, S. Doruc, G. Gorincioi, S. Mura, M. Pirtac, I. Stancul, D. Tcaciuc, M. Albert, I. Alexopoulou, A. Arnaout, J. Bartlett, J. Engel, S. Gilbert, J. Parfitt, H. Sekhon, G. Thomas, D.M. Rassl, R.C. Rintoul, C. Bifulco, R. Tamakawa, W. Urba, N. Hayward, H. Timmers, A. Antenucci, F. Facciolo, G. Grazi, M. Marino, R. Merola, R. de Krijger, A.-P. Gimenez-Roqueplo, A. Piché, S. Chevalier, G. McKercher, K. Birsoy, G. Barnett, C. Brewer, C. Farver, T. Naska, N.A. Pennell, D. Raymond, C. Schilero, K. Smolenski, F. Williams, C. Morrison, J.A. Borgia, M.J. Liptay, M. Pool, C.W. Seder, K. Junker, L. Omberg, M. Dinkin, G. Manikhas, D. Alvaro, M.C. Bragazzi, V. Cardinale, G. Carpino, E. Gaudio, D. Chesla, S. Cottingham, M. Dubina, F. Moiseenko, R. Dhanasekaran, K.-F. Becker, K.-P. Janssen, J. Slotta-Huspenina, M.H. Abdel-Rahman, D. Aziz, S. Bell, C.M. Cebulla, A. Davis, R. Duell, J.B. Elder, J. Hilty, B. Kumar, J. Lang, N.L. Lehman, R. Mandt, P. Nguyen, R. Pilarski, K. Rai, L. Schoenfield, K. Senecal, P. Wakely, P. Hansen, R. Lechan, J. Powers, A. Tischler, W. E. Grizzle, K.C. Sexton, A. Kastl, J. Henderson, S. Porten, J. Waldmann, M. Fassnacht, S.L. Asa, D. Schadendorf, M. Couce, M. Graefen, H. Huland, G. Sauter, T. Schlomm, R. Simon, P. Tennstedt, O. Olabode, M. Nelson, O. Bathe, P.R. Carroll, J. M. Chan, P. Disaia, P. Glenn, R.K. Kelley, C.N. Landen, J. Phillips, M. Prados, J. Simko, K. Smith-McCune, S. VandenBerg, K. Roggin, A. Fehrenbach, A. Kendler, S. Sifri, R. Steele, A. Jimeno, F. Carey, I. Forgie, M. Mannelli, M. Carney, B. Hernandez, B. Campos, C. Herold-Mende, C. Jungk, A. Unterberg, A. von Deimling, A. Bossler, J. Galbraith, L. Jacobus, M. Knudson, T. Knutson, D. Ma, M. Milhem, R. Sigmund, A.K. Godwin, R. Madan, H.G. Rosenthal, C. Adebamowo, S.N. Adebamowo, A. Boussioutas, D. Beer, T. Giordano, A.-M. Mes-Masson, F. Saad, T. Bocklage, L. Landrum, R. Mannel, K. Moore, K. Moxley, R. Postier, J. Walker, R. Zuna, M. Feldman, F. Valdivieso, R. Dhir, J. Luketich, E.M.M. Pinero, M. Quintero-Aguilo, C.G. Carlotti, J.S. Dos Santos, R. Kemp, A. Sankarankuty, D. Tirapelli, J. Catto, K. Agnew, E. Swisher, J. Creaney, B. Robinson, C.S. Shelley, E.M. Godwin, S. Kendall, C. Shipman, C. Bradford, T. Carey, A. Haddad, J. Moyer, L. Peterson, M. Prince, L. Rozek, G. Wolf, R. Bowman, K.M. Fong, I. Yang, R. Korst, W.K. Rathmell, J.L. Fantacone-Campbell, J.A. Hooke, A.J. Kovatich, C.D. Shriver, J. DiPersio, B. Drake, R. Govindan, S. Heath, T. Ley, B. Van Tine, P. Westervelt, M.A. Rubin, J. Il Lee, N.D. Aredes, A. Mariamidze, J.M. Stuart, C.C. Benz, P.W. Laird, Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer, Cell. 173 (2018) 291-304.e6. https://doi.org/10.1016/J.CELL.2018.03.022.

[41] M. Olivier, M. Hollstein, P. Hainaut, TP53 mutations in human cancers: origins, consequences, and clinical use, Cold Spring Harb. Perspect. Biol. 2 (2010), https://doi.org/10.1101/cshperspect.a001008.

[42] E.H. Baugh, H. Ke, A.J. Levine, R.A. Bonneau, C.S. Chan, Why are there hotspot mutations in the TP53 gene in human cancers? Cell Death Differ. 25 (2018) https://doi.org/10.1038/cdd.2017.180.

[43] J. Foo, F. Michor, Evolution of acquired resistance to anti-cancer therapy, J. Theor. Biol. 355 (2014), https://doi.org/10.1016/j.jtbi.2014.02.025.

[44] A.G. Thorsell, T. Ekblad, T. Karlberg, M. Löw, A.F. Pinto, L. Trésaugues, M. Moche, M.S. Cohen, H. Schüler, Structural basis for potency and promiscuity in poly(ADP-ribose) polymerase (PARP) and Tankyrase inhibitors, J. Med. Chem. 60 (2017), https://doi.org/10.1021/acs.jmedchem.6b00990.

[45] Y. Du, H. Yamaguchi, Y. Wei, J.L. Hsu, H.L. Wang, Y.H. Hsu, W.C. Lin, W.H. Yu, P. G. Leonard, G.R. Lee, M.K. Chen, K. Nakai, M.C. Hsu, C. Te Chen, Y. Sun, Y. Wu, W. C. Chang, W.C. Huang, C.L. Liu, Y.C. Chang, C.H. Chen, M. Park, P. Jones, G. N. Hortobagyi, M.C. Hung, Blocking c-Met-mediated PARP1 phosphorylation

enhances anti-tumor effects of PARP inhibitors, Nat. Med. 22 (2016), https://doi.org/10.1038/nm.4032.

[46] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, A. Bridgland, C. Meyer, S.A. A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021), https://doi.org/10.1038/s41586-021-03819-2.

[47] D.F. Burke, P. Bryant, I. Barrio-Hernandez, D. Memon, G. Pozzati, A. Shenoy, W. Zhu, A.S. Dunham, P. Albanese, A. Keller, R.A. Scheltema, J.E. Bruce, A. Leitner, P. Kundrotas, P. Beltrao, A. Elofsson, Towards a structurally resolved human protein interaction network, Nat. Struct. Mol. Biol. 30 (2023), https://doi.org/10.1038/s41594-022-00910-8.