UNIVERSITÀ DEGLI STUDI DI TORINO

Hotspot activating *PRKD1* somatic mutations in polymorphous low-grade adenocarcinomas of the salivary glands

Ilan Weinreb[1,26], Salvatore Piscuoglio[2,26], Luciano G Martelotto[2,26], Daryl Waggott[3–5,26], Charlotte K Y Ng[2], Bayardo Perez-Ordonez[1], Nicholas J Harding[3], Javier Alfaro[3–6], Kenneth C Chu[3], Agnes Viale[7], Nicola Fusco[2,8], Arnaud da Cruz Paula[2,9], Caterina Marchio[2], Rita A Sakr[10], Raymond Lim[2], Lester D R Thompson[11], Simion I Chiosea[12], Raja R Seethala[12], Alena Skalova[13], Edward B Stelow[14], Isabel Fonseca[15,16], Adel Assaad[17], Christine How[4,5], Jianxin Wang[3], Richard de Borja[3], Michelle Chan-Seng-Yue[3], Christopher J Howlett[18], Anthony C Nichols[18], Y Hannah Wen[2], Nora Katabi[2], Nicholas Buchner[19], Laura Mullen[19], Thomas Kislinger[4–6], Bradly G Wouters[4–6], Fei-Fei Liu[4–6,20], Larry Norton[21], John D McPherson[6,17], Brian P Rubin[22,23], Blaise A Clarke[1,27], Britta Weigelt[2,27], Paul C Boutros[3,6,24,27] & Jorge S Reis-Filho[2,25,27]

[1]Department of Pathology, University Health Network, Toronto, Ontario, Canada.

[2]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York, USA.

[3]Informatics and Bio-Computing Program, Ontario Institute for Cancer Research, Toronto, Ontario, Canada.

[4]Ontario Cancer Institute, Princess Margaret Cancer Centre, University Health Network, Toronto, Onatrio, Canada.

[5]Campbell Family Institute for Cancer Research, Princess Margaret Cancer Centre, University Health Network, Toronto, Ontario, Canada.

[6]Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada.

[7]Integrated Genomics Operation, Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center, New York, New York, USA.

[8]School of Pathology, University of Milan, Milan, Italy.

[9]Instituto Português de Oncologia, Oporto, Portugal.

[10]Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York, USA.

[11]Department of Pathology, Kaiser Permanente, Woodland Hills Medical Center, Woodland Hills, California, USA.

[12]Department of Pathology, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, USA.

[13]Department of Pathology and Laboratory Medicine, Charles University in Prague, Plzen, Czech Republic.

[14]Department of Pathology, University of Virginia Medical Center, Charlottesville, Virginia, USA.

[15]Instituto Português de Oncologia Francisco Gentil, Lisbon, Portugal. [16]Faculdade de Medicina de Lisboa, Lisbon, Portugal.

[17]Department of Pathology, Virginia Mason Hospital and Seattle Medical Center, Seattle, Washington, USA.

[18]Western University, London, Ontario, Canada.

[19]Cancer Genomics Platform, Ontario Institute for Cancer Research, Toronto, Ontario, Canada.

[20]Department of Radiation Oncology, Princess Margaret Hospital and University of Toronto, Toronto, Ontario, Canada.

[21]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA.

[22]Department of Molecular Genetics, Lerner Research Institute, Cleveland, Ohio, USA. [23]Robert J. Tomsich Pathology and Laboratory Medicine Institute, Taussig Cancer Center, Cleveland Clinic, Cleveland, Ohio, USA.

[24]Department of Pharmacology and Toxicology, University of Toronto, Toronto, Ontario, Canada.

[25]Affiliated member of the Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, New York, USA.

[26]These authors contributed equally to this work. [27]These authors jointly directed this work. Correspondence should be addressed to J.S.R.-F. (reisfilj@mskcc.org).

Polymorphous low-grade adenocarcinoma (PLGA) is the second most frequent type of malignant tumor of the minor salivary glands. We identified *PRKD1* hotspot mutations encoding p.Glu710Asp in 72.9% of PLGAs but not in other salivary gland tumors. Functional studies demonstrated that this kinase-activating alteration likely constitutes a driver of PLGA.

PLGA is the second most common intraoral malignant salivary gland tumor, preferentially affecting the minor salivary glands1. This tumor occasionally originates in other anatomical sites, including the lungs and breast2. PLGAs are characterized by cytological uniformity and histological diversity1. Despite their infiltrative growth pattern and tendency for perineural invasion, PLGAs have a relatively indolent clinical course, with lymph node metastasis in up to 29% of cases and

rare distant metastasis3,4. Owing to the large spectrum of histological appearances of PLGAs, the differentiation of these tumors from more aggressive forms of salivary gland cancer is challenging, especially in small diagnostic specimens.

Recent studies have demonstrated that distinct types of malignant salivary gland tumors, including adenoid cystic carcinoma (AdCC), mucoepidermoid carcinoma and secretory carcinoma, are driven by specific highly recurrent somatic genetic alterations (for example, fusion genes)5. Like PLGA, these tumor types can also originate in other anatomical sites, including the lungs and breast2. We therefore reasoned that PLGA might be driven by a highly recurrent somatic genetic alteration.

To test this hypothesis, we subjected three consecutive PLGAs to massively parallel RNA sequencing and whole-exome sequencing and a separate set of three PLGAs to whole-exome sequencing (**Supplementary Table 1**). RNA sequencing did not identify any recurrent in-frame fusion genes (**Supplementary Table 2**). By contrast, whole-exome sequencing identified two somatic heterozygous single-nucleotide variants (SNVs) in the *PRKD1* gene, c.2130A>T and c.2130A>C, affecting a highly conserved amino acid in the catalytic loop (**Fig. 1a**) and resulting in the same amino acid substitution, p.Glu710Asp, in five of six cases (**Fig. 1b**, **Supplementary Fig. 1** and **Supplementary Table 1**). RNA sequencing demonstrated that the mutant alleles were expressed (**Supplementary Fig. 2** and **Supplementary Table 1**). No additional recurrent somatic SNVs or insertion-deletions (indels) were detected (**Supplementary Table 3**). We therefore subjected a validation cohort of 53 PLGAs to Sanger sequencing and targeted amplicon sequencing of *PRKD1* exon 15 and confirmed the presence of the somatic c.2130A>T and c.2130A>C mutations in 41.5% and 30.2% of PLGAs, respectively (**Fig. 1b**, **Supplementary Fig. 3** and **Supplementary Table 4**). In total, 43 of 59 (72.9%) PLGAs tested harbored a *PRKD1* mutation encoding p.Glu710Asp. Detailed pathological analysis of PLGAs with mutated and wild-type *PRKD1*, however, did not identify distinct histological features that would differentiate the two groups (data not shown). PLGAs with and without the *PRKD1* mutation encoding p.Glu710Asp were found to express PRKD1 protein **(Fig. 1c** and **Supplementary Fig. 2b**,**c**), suggesting that, as with other mutation-activated oncogenes6, the expression of PRKD1 protein is more pervasive than the presence of a *PRKD1* hotspot mutation.

PRKD1 is a serine-threonine kinase that has a role in cell adhesion, cell migration, vesicle transport and cell survival7. We found that the *PRKD1* gene was mutated in 2.2% of the cancer samples included in The Cancer Genome Atlas (TCGA) and other published data sets available on cBioPortal8, but none of these cancers harbored the *PRKD1* c.2130A>T or c.2130A>C (p.Glu710Asp) mutations, suggesting that these SNVs might be unique to PLGA (**Fig. 1d**,**e**). In addition, these new *PRKD1* mutations have not been reported in the Catalog of Somatic Mutations in Cancer9 (COSMIC, v68; data not shown). To determine whether the *PRKD1* mutations encoding p.Glu710Asp would be pathognomonic for PLGA, we investigated the presence of these mutations in a series of 311 tumors of the salivary glands (186 malignant and 125 benign lesions), of which 12

were PLGAs. The *PRKD1* mutations encoding p.Glu710Asp were found in eight of the PLGAs but not in any of the other salivary gland tumors studied (**Fig. 1f** and **Supplementary Table 5**). Therefore, our findings demonstrate that the *PRKD1* hotspot mutations encoding p.Glu710Asp, akin to the *CRTC1-MAML2* and *MYB-NFIB* fusion genes defining subsets of mucoepidermoid carcinoma and AdCC, respectively5, define a subset of PLGAs of the salivary glands and might be useful as an ancillary diagnostic marker to differentiate PLGAs from other forms of salivary gland tumors. In addition, the presence of a *PRKD1* mutation encoding p.Glu710Asp was significantly associated with the metastasis-free survival of individuals with malignant salivary gland tumors ($P = 0.027$; **Fig. 1g**, **Supplementary Table 6** and **Supplementary Note**), suggesting that these mutations might be useful in the molecular prognostic work-up of salivary gland cancers.

The p.Glu710Asp hotspot alteration affects the catalytic loop HCDLKPEN of the PRKD1 protein kinase domain (**Fig. 1a**) and was predicted to be pathogenic by independent mutation function assessment algorithms (**Supplementary Table 3**). Although the crystal structure of PRKD1 is not yet solved, homology modeling suggested that the p.Glu710Asp alteration could either interfere with ADP binding by altering coordination with Mg2+ or affect enzyme kinetics, as it is positioned in between the ATP donor binding site and the putative proton acceptor (**Fig. 2a** and **Supplementary Fig. 4a**). The glutamic acid residue is highly conserved among the CAMK kinases; however, an aspartic acid residue is frequently observed within the native structures of AGC kinases, suggesting that the identified alteration is likely active but might induce a functional shift (**Supplementary Fig. 4b**). Consistent with these protein structure predictions, in a cell-free *in vitro* kinase assay, p.Glu710Asp-mutant PRKD1 displayed significantly increased transphosphorylation of the serine-threonine substrate CREBtide and elevated autocatalytic activity in comparison to wild-type PRKD1 ($P < 0.0001$; **Fig. 2b** and **Supplementary Fig. 5a**). Forced expression of wild-type and p.Glu710Asp-mutant PRKD1 in embryonic kidney epithelial and non-malignant breast epithelial cell lines (**Supplementary Fig. 5b**) resulted in increased phosphorylation of Ser738/Ser742 within the activation loop and Ser910 in the autocatalytic site of the mutant PRKD1 protein relative to wild-type protein in all subcellular compartments (**Fig. 2c** and **Supplementary Fig. 5c**), in agreement with the higher trans- and autocatalytic kinase activity detected in the cell-free kinase assay.

Previous studies have suggested that PRKD1 activation reduces cell migration and invasion but promotes the growth of epithelial cells7. Consistent with this notion, forced expression of p.Glu710Asp-mutant PRKD1 in conventional *in vitro* models resulted in reduced cell migration, despite an increase in cell viability (**Supplementary Fig. 6**). To ascertain whether the *PRKD1* hotspot mutations encoding p.Glu710Asp would affect the growth and glandular architecture of epithelial cells, we forced expression of wild-type and p.Glu710Asp-mutant PRKD1 in two non-malignant breast epithelial cell lines, MCF10A and MCF12A, grown in three-dimensional cultures10. These cells were chosen because PLGA cell lines and non-malignant models of minor salivary gland cells are not available and because, albeit rarely, primary PLGAs have been reported in the breast2. Forced expression of p.Glu710Asp-mutant PRKD1 changed the hollow spheroid, acinar-like structures formed by parental or empty vector–transfected MCF10A and MCF12A cells into larger, coalescent structures with filled lumens and irregular contours not uncommonly displaying infiltrating edges, a phenotype consistent with that induced by the forced expression of other oncogenes in this model system10,11 (**Fig. 2d**). Taken together, our findings demonstrate that

the somatic *PRKD1* hotspot mutations encoding p.Glu710Asp are likely activating and might confer a neoplastic advantage to epithelial cells.

We have identified somatic *PRKD1* activating hotspot mutations encoding p.Glu710Asp in the majority of PLGAs, which might be employed as an ancillary molecular marker to differentiate PLGA from its more aggressive mimics, including AdCC. *PRKD1* emerges as a new cancer-related gene, which likely constitutes a driver of PLGA and might be exploited as a diagnostic aid and therapeutic target. Our study emphasizes that investigating the genetic underpinnings of rare cancer types can lead to the identification of new oncogenic drivers and help in the development of a molecular taxonomy of cancers12.

J.S.R.-F. and C.M. interpreted the immunohistochemical results. C.K.Y.N., K.C.C. and P.C.B. performed statistical analyses. B.W. and J.S.R.-F. wrote the first draft of the manuscript, which was initially reviewed by I.W., S.P., L.G.M., D.W., L.N., B.P.R. and P.C.B. All authors edited and approved the final draft.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

ONLINE METHODS

Subjects and samples. After obtaining institutional review board (IRB) approval from the respective institutions of the authors, representative histological blocks of PLGAs were retrieved from the pathology archives of the respective pathology departments and sent to the pathology department of the University Health Network (UHN) (Toronto, Ontario, Canada) or Memorial Sloan-Kettering Cancer Center (MSKCC) (New York, New York, USA). In addition, three frozen samples of PLGAs and matched adjacent normal salivary gland samples were retrieved from the MSKCC Tumor Procurement Service, and three frozen samples of PLGAs were retrieved from the UHN Biospecimen Sciences Program.

Sixty-eight cases were reviewed by the contributing pathologists and centrally reviewed by two pathologists with an interest and expertise in head and neck cancer pathology (I.W. and B.P.-O.). Only cases classified as classic PLGA according to the World Health Organization definition were included in this study1. Cases that fulfilled the diagnostic criteria of cribriform adenocarcinoma of the minor salivary glands or adenocarcinoma not otherwise specified were excluded1. In total, 59 classic PLGAs were included in this study, six of which were fresh frozen and 53 of which were formalin fixed and paraffin embedded. Patient consents were obtained if required by the IRB protocols approved by each author's institution. Follow-up data were obtained from patient charts from the authors' institutions. Before tissue processing, samples were anonymized.

The discovery cohort (**Supplementary Table 1**) comprised six fresh-frozen PLGAs, three of which had matched frozen adjacent normal salivary gland tissue available. These three cases with matched adjacent normal tissue were microdissected with a needle under a stereomicroscope as previously described13. RNA was extracted from microdissected tumor samples as previously described13 and subjected to paired-end massively parallel RNA sequencing. DNA was extracted separately from microdissected tumor samples and normal adjacent tissues as previously described13 and subjected

to whole-exome sequencing. From the remaining three cases, only representative frozen tumor tissues were available; these cases were microdissected and DNA was extracted and subjected to whole-exome sequencing.

The validation cohort (**Supplementary Table 4**) consisted of representative formalin-fixed, paraffin-embedded samples from 53 PLGAs. These samples were microdissected with a needle under a stereomicroscope, and DNA was extracted and quantified and its quality assessed as previously described13.

For power calculations, assuming that, akin to other rare tumor types14,15, PLGAs would be driven by a recurrent hotspot mutation or fusion gene present in ≥70% of cases, a binomial distribution showed that by sequencing six samples we would be able to identify a recurrent event (occurring in two or more cases) with >90% statistical power.

Paired-end massively parallel RNA sequencing and fusion gene identification. The three PLGAs retrieved from MSKCC were subjected to massively parallel RNA sequencing using a validated protocol16 employed in the MSKCC Integrated Genomics Operation (iGO). In brief, paired-end RNA sequencing was performed with 2 × 75-bp cycles on the Illumina HiSeq 2000 platform. Data were aligned to the reference human genome hg19 using Bowtie 2 (ref. 17). For PLGA_1, PLGA_2 and PLGA_3, 121.7, 66.7 and 56.8 million reads were obtained, respectively. Read pairs supporting fusion transcripts were identified using ChimeraScan (v0.4.5)18 and deFuse (v0.6.1)19. To filter out common alignment artifacts and normal transcriptional variation, we removed fusion gene and read-through candidates that were also found in a set of 47 normal samples from TCGA (project 4767). Candidate fusions were annotated using Oncofuse20 (**Supplementary Table 2**).

Whole-exome sequencing. For the MSKCC samples, microdissected tumor and germline DNA were subjected to exome capture with SureSelect Human All Exon v4 (Agilent Technologies) using a validated protocol16,21 employed in the MSKCC iGO. An average of 154 million 75-bp paired-end reads were generated on an Illumina HiSeq 2000 instrument for each sample, equivalent to an average depth of 102× (range of 84.5–181.5×).

For the UHN samples, DNA from the three tumors was subjected to exome capture with Agilent SureSelect All Exon G3362 (203,733 regions; 50,075,646 bp). Each sample was run in a single flow cell lane of an Illumina HiSeq 2000 instrument, generating approximately 220 million 101-bp paired-end reads, to an average depth of 159× (range of 137–187×).

Variant calling. For the MSKCC samples, reads were aligned to the reference human genome hg19 using the Burrows-Wheeler Aligner (BWA, v0.6.2)22. Local realignment and quality score recalibration were performed using the Genome Analysis Toolkit (GATK, v3.1.1)23. Deduplication was performed using Picard (v1.92). Somatic SNVs were called using MuTect (v.1.1.4)24, and small somatic indels were called using VarScan2 (v2.3.6)25 and the GATK HaplotypeCaller (v3.1.1)23. Variants found with >5% global minor allele frequency in dbSNP (Build 137) or that were supported by <5 reads were disregarded. SNVs for which the tumor variant allele fraction was <5 times than that of the normal variant allele fraction were disregarded. All indels were manually inspected using the Integrative Genomics Viewer (IGV)26.

For the UHN samples, reads were aligned to the UCSC build hg19 reference using Novoalign (v2.07.14, Novocraft Technologies). Alignments were optimized by soft clipping low-quality end positions and rescuing poorly aligned reads via high-quality anchoring of their fragment pair. SAMtools (v1.18)27 was used for deduplication and filtering. Local realignment, quality score recalibration and variant detection were performed using GATK (v2.4.7)23, employing recommended best practices (v3 and v4). On-target variants (padded by 10 bp) were considered, and variants that were flagged by GATK base quality, read depth, alternate allele frequency and haplotype inconsistency filters were removed. Variants in dbSNP (Build 137; variants not flagged as somatic or clinical or as having a minor allele frequency of <1%; ref. 28) or in the 1000 Genomes Project (phase 1, release v3, 23 November 2010) were removed. Variants were excluded if they occurred in Fuentes29, ENCODE30 or the Duplicated Gene Database31 or in ten normal blood samples (data not shown). All COSMIC variants (v64) were white-listed9,32.

Prediction of the functional effect of the mutations was performed using PolyPhen-2 (ref. 33), MutationTaster34, Mutation Assessor35, FATHMM36, CHASM (breast, head and neck, and other classifiers)37, CanDrA (breast classifier)38 and CADD39.

PCR amplification, Sanger sequencing and targeted amplicon sequencing. The presence of somatic *PRKD1* mutations encoding p.Glu710Asp was validated in all PLGAs from the discovery cohort and investigated in all PLGAs from the validation cohort by Sanger sequencing analysis of tumor and germline samples as previously described40. For this analysis, primers (**Supplementary Table 7**) to amplify a 338-bp fragment covering exon 15 of the *PRKD1* gene (chromosomal position: chr. 14: 30,068,073–30,068,410) were designed using OligoPerfect (Life Technologies). PCR fragments were then purified using ExoSAP-IT (USB) and sequenced as previously described40.

For targeted amplicon sequencing, 5 ng of genomic DNA was amplified with primers specific for exon 15 of the *PRKD1* gene (**Supplementary Table 7**) using a high-fidelity Taq polymerase (Roche). PCR amplicons were sequenced on a MiSeq instrument (Illumina) using 150 × 150 chemistry. Reads were aligned to the reference human genome hg19 using BWA (v0.6.2)22. Local realignment and quality score recalibration were performed using GATK (v3.1.1)23. Mutations were called using the GATK HaplotypeCaller (v3.1.1)23 and were manually inspected using IGV26.

Immunohistochemistry. Representative sections from 3 frozen PLGAs included in the discovery set from MSKCC and from 18 formalin-fixed, paraffin-embedded PLGAs from the validation cohort were cut at 4 μm and mounted on Superfrost Plus slides (Fisher Scientific). Frozen sections of each of the three tumors were briefly fixed in acetone and subjected to PRKD1 immunohisto-chemical analysis. Formalin-fixed, paraffin-embedded sections were subjected to heat-induced antigen retrieval using the pH 9 Target Retrieval Solution (Dako) at 92.8 °C for 20 min in a water bath, and immunohistochemical analysis was then performed using a polyclonal antibody to PRKD1 (Sigma-Aldrich, HPA029834; 1:250 dilution). Detection of bound antibody was performed using the ABC-diaminobenzidine method (Vector Laboratories). Negative controls (omission of the primary antibody and substitution of the primary antibody with IgG-matched control) and positive controls (formalin-fixed, paraffin-embedded pellets41 of MCF12A cells transfected with empty vector or with vector for wild-type or p.Glu710Asp-mutant PRKD1) were included in each slide

run. Sections were analyzed by two pathologists (C.M. and J.S.R.-F.) blinded to the *PRKD1* mutation status of the cases using the H-score system, whereby cytoplasmic expression in histologically unequivocal neoplastic cells was semiquantitatively assessed. Nuclear expression was semiquantitatively assessed using the Allred scoring system as previously described42; an Allred score of >2 was considered positive. Unequivocal membranous expression was not observed in any case. No statistically significant difference in the expression levels of cytoplasmic and nuclear PRKD1 was observed between PLGAs with mutant and wild-type *PRKD1* (**Supplementary Fig. 2c**).

***PRKD1* c.2130A>T and c.2130A>C (p.Glu710Asp) mutation screening in salivary gland tumors.** We retrieved 311 primary salivary gland lesions from the pathology departments of UHN, Cleveland Clinic and MSKCC. The inclusion criteria for the cases were as follows. Consecutive cases of salivary gland lesions diagnosed and managed at UHN between January 2002 and December 2010, at Cleveland Clinic between January 1999 and December 2013, and at MSKCC between January 2008 and December 2013 were retrieved from the respective hospital files; referral cases and cases where insufficient tumor material was available were excluded (**Supplementary Table 5**). DNA was extracted from all cases as described above. In 96-well plates, the DNA samples from these cases and 8 randomly selected PLGAs previously included in the validation series described above were assembled by 3 of the authors and subjected to Sanger sequencing as described above by another author, who was blinded to the identity and diagnosis of the samples and prior sequencing results.

To define the impact of the *PRKD1* mutations encoding p.Glu710Asp on the outcome of patients with malignant salivary gland tumors, a cohort of malignant salivary gland tumors was retrospectively accrued by combining (i) all cases of malignant salivary gland tumors subjected to screening for the *PRKD1* c.2130A>T and c.2130A>C (p.Glu710Asp) mutations with (ii) all cases of PLGA analyzed in this study. This combined data set comprised 237 individual patients. Follow-up data were obtained from patient charts reviewed by three of the authors. After follow-up data were obtained, samples were anonymized. Follow-up information was available for 149 patients with malignant salivary gland tumors, ranging from 1 to 180 months (median of 34 months; **Supplementary Table 6**). Metastasis-free survival was expressed as the number of months from diagnosis to the occurrence of distant relapse. Survival analysis was performed using SPSS version 20 (IBM). Cumulative survival probabilities were calculated using the Kaplan-Meier method. Differences between metastasis-free survival rates were tested with the log-rank test. A P value of <0.05 was considered significant.

Homology modeling. A homology model of the kinase domain of PRKD1 and p.Glu710Asp-mutant PRKD1 was generated using SWISS-MODEL43–46 on the structure of CHEK2 (Protein Data Bank (PDB) 2YCR) with which it has 40% sequence identity (**Supplementary Fig. 4a**,**b**). The model for the ADP molecule was obtained from a related CHEK2 structure (2CN5) with overlap generated by PyMOl (v1.5.0.4, Schrödinger). The homology models were evaluated using QMEAN47–50 (**Supplementary Fig. 4a**). The QMEANnorm score and QMEAN z score49 estimate the absolute quality of a homology model in comparison to X-ray crystallography reference structures. The reference structures are a non-redundant subset of PDB with less than 30% pairwise sequence identity among themselves and resolved at a resolution of <2 Å. PRKD1 models

have QMEAN *z* scores of −1.80 (wild type) and −1.71 (mutant), resulting in a QMEANnorm score of 0.60 for both (**Supplementary Fig. 4a**) and placing the models well within the range of scores for similarly sized resolved structures in the reference set. Residue-by-residue examination showed that the active site and mutation region (**Supplementary Fig. 4a**) were modeled with high accuracy. Model error was dominated by regions near flexible loops and toward the extremities of the homology model. All homology model visualizations are for one chain in the dimer and were rendered using the Visual Molecular Dynamics software suite51.

Sequence alignment and analysis of the PRKD1 mutant site against the human protein kinase complement. Multiple-Sequence Comparison by Log Expectation (MUSCLE)52 was used to align the mutated kinase domain of PRKD1 against an alignment of the eukaryotic protein kinase (ePK) domains used in Manning *et al.*53 describing the major kinase groups. All sequences observed to display a glutamic acid or aspartic acid at the homologous site of the PRKD1 alteration were identified and used in conjunction with Kinome-render54 to annotate the tree shown in **Supplementary Figure 4b**.

Amino acid sequence alignment of PRKD1 across species. The amino acid sequences of the PRKD1 protein from *Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Pan troglodytes*, *Equus ferus caballus*, *Bos taurus*, *Gallus gallus*, *Xenopus tropicalis*, *Danio rerio* and *Drosophila melanogaster* were retrieved from Ensembl and aligned using ClustaW in DNASTAR Lasergene.

Cell lines. MCF10A, MCF12A, HEK293T and NIH3T3 cells were purchased from the American Type Culture Collection (ATCC), authenticated using short tandem repeat profiling as previously described40 and tested for mycoplasma using the PCR-based Universal Mycoplasma Detection kit (ATCC). All cell lines were maintained in an atmosphere with 5% CO_2 at 37 °C. MCF10A and MCF12A cells were cultured in DMEM/F12 supplemented with 5% horse serum, 20 ng/ml epidermal growth factor (EGF), 10 μg/ml insulin, 0.5 μg/ml hydrocortisone and 1% penicillin-streptomycin as described in ref. 10, and HEK293T and NIH3T3 cells were cultured in DMEM high glucose supplemented with 10% FBS and 1% penicillin-streptomycin.

Vector construction, mutagenesis, transformation and plasmid preparation. The human *PRKD1* (NM_002742) cDNA clone was purchased from Origene (RG211676; C-terminal TurboGFP tag). The mutation encoding p.Glu710Asp was introduced in pCMV6:PRKD1-tGFP using the Q5 Site-Directed Mutagenesis kit (New England BioLabs). The presence of the correct mutation and the absence of rearrangements were confirmed by Sanger sequencing (primer sequences available in **Supplementary Table 7**).

Transfections of mammalian cells and analysis of transgene expression. Transfections with plasmids were performed using Lipofectamine LTX and PLUS Reagent (Life Technologies) following the manufacturer's guidelines. The expression of transgenes was visually evaluated 48 h after transfection using a Nikon Eclipse Ti fluorescence inverted microscope and by protein blotting using antibodies to PRKD1/PKCμ (see below).

Protein blotting. Standard protein blotting was conducted as previously described40 using antibodies against PRKD1/PKCμ (2025), phosphorylated PRKD1/PKCμ (Ser744/Ser748; 2054), phosphorylated PRKD1/PKCμ (Ser916; 2051), MEK1/MEK2, AIF, vimentin and histone H3 (Cell Fractionation Antibody Sampler kit, 11843; all from Cell Signaling Technology; 1:1,000 dilution).

Conjugated IRDye680RD/800CW secondary antibodies were used at a 1:15,000 dilution and detected using the Odyssey Infrared Imaging System (LI-COR Biosciences)40. Quantification and analysis were performed using LI-COR Image Studio Software. When required, stripping of membranes was performed using Restore PLUS Western Blot Stripping Buffer (Thermo Scientific).

Confocal microscopy of PRKD1 subcellular localization. Cells growing on coverslips expressing PRKD1-tGFP, PRKD1(Glu710Asp)-tGFP and tGFP (control) proteins were fixed in 10% buffered formalin (Fisher Scientific), washed with PBS, counterstained with 4′,6-diamidino-2-phenylindole (DAPI; Life Technologies) and mounted using ProLong Gold Antifade Reagent (Life Technologies). Photomicrographs were acquired using a Leica TCS SP5-II Upright microscope. All experiments were performed in triplicate.

Cell fractionation. Protein fractions enriched for cytoplasmic, organellular/membranous and nuclear/cytoskeletal components from MCF10A and MCF12A cells expressing the PRKD1-tGFP, PRKD1(Glu710Asp)-tGFP and tGFP proteins were prepared using the Cell Fractionation kit (9038; Cell Signaling Technology) following the manufacturer's protocol. The efficiency of separation was evaluated using the Cell Fractionation Antibody Sampler kit (11843, Cell Signaling Technology).

Three-dimensional cell cultures and confocal microcopy. Overlay three-dimensional cultures of MCF10A and MCF12A cells expressing the PRKD1-tGFP, PRKD1(Glu710Asp)-tGFP and tGFP proteins were carried out as previously described10. At day 14, phase-contrast images of acinar structures were acquired (EVOS XL Imaging System, Life Technologies); cells were fixed, stained with DAPI, extracted from the three-dimensional cultures and mounted as described in Lee *et al.*55. Confocal images were acquired at the acinar-like spheroid midsection using a Leica TCS SP5-II Upright microscope. All experiments were performed in triplicate.

PRKD1-tGFP and *PRKD1*(Glu710Asp)-tGFP immunoprecipitation and serine-threonine kinase assays. Protein lysates of HEK293T cells were prepared 48 h after transfection. The PRKD1-tGFP, PRKD1(Glu710Asp)-tGFP and tGFP proteins were immunoprecipitated using anti-tGFP(2H8)-conjugated magnetic beads (Origene). All steps were carried out at 4 °C. All experiments were performed in triplicate.

Serine-threonine kinase activity was evaluated using the ADP Hunter HS Assay (DiscoveRx) following the procedure recommended by the manufacturer. The immunocomplexes bound to beads were resuspended in 40 µl of ADP Assay Buffer supplemented with 25 µM ATP and with 100 µM CREBtide as a substrate (KRREILSRRPSYR; Enzo Life Sciences); autophosphorylation was evaluated by omitting CREBtide in the reaction. The reactions proceeded for 30 min at 30 °C with vigorous agitation, and were terminated by placing the tubes on a magnetic stand and removing the supernatant for the ADP detection step. Fluorescence readings (excitation/emission: 530/590 nm) were taken after a 20-min incubation using a Victor X4 Multimode Plate Reader (PerkinElmer).

The immunocomplexes bound to beads were mixed with 2× NuPAGE LDS Sample Buffer (Life Technologies) and heated for 5 min at 95 °C, and eluted proteins were resolved by electrophoresis. Standard protein blotting, using antibodies against PRKD1/PKCµ and phosphorylated PRKD1/PKCµ, was carried out as described above.

Growth curves. MCF10A and MCF12A cells expressing PRKD1-tGFP, PRKD1(Glu710Asp)-tGFP and tGFP were seeded in 96-well plates (1,500 cells/well, in triplicate), and proliferation was monitored over 5 d using the CellTiter-Blue Cell Viability Assay (Promega). Fluorescence readings were performed using the Victor X4 Multimode Plate Reader (PerkinElmer). Growth curves were plotted and analyzed (multiple *t* tests, corrected for multiple comparisons using the Holm-Šídák method as the default as recommended, $\alpha = 0.05$) using Prism v6.0c (GraphPad).

Transwell assays. Transwell assays were performed by plating 1 × 105 serum-starved and mytomicin C–treated (Sigma-Aldrich) MCF12A and NIH3T3 cells expressing PRKD1-tGFP, PRKD1(Glu710Asp)-tGFP and tGFP in low-serum growth medium (with 0.5% horse serum and 0.5% FBS, respectively) into the upper chamber of a 24-well format Falcon Cell Culture Insert (8-μm pore size; Corning); the lower compartment was filled with the corresponding complete growth medium. The assay was carried out for 18 h, and migratory cells on the lower side of the insert filter were fixed and stained with 1% crystal violet in 2% ethanol (Sigma-Aldrich). The entire well for each condition was photographed, and the various photos were merged using Adobe Photoshop (Photomerge tool). Experiments were performed in triplicate. Quantification of migrated cells was performed by automatic counting using the ImageJ software (v1.48b), and numbers were compared using an unpaired *t* test (GraphPad Prism v6.0c). A two-tailed *P* value of <0.05 was considered significant.

Scratch wound healing assays. Serum-starved and mytomicin C–treated (Sigma-Aldrich) NIH3T3 and MCF12A cells expressing PRKD1-tGFP, PRKD1(Glu710Asp)-tGFP and tGFP (control) were seeded at 90–95% confluency. After cell attachment, a p200 pipet tip was used to create a scratch in the cell monolayer. Representative phase-contrast images were taken at 0 h and 18 h after scratch wounding. Experiments were performed in triplicate.

1. Barnes, L., Eveson, J., Reichart, P. & Sidransky, D. *World Health Organization Classification of Tumours. Pathology And Genetics of Head and Neck Tumours* (IARC Press, Lyon, France, 2005).

2. Foschini, M.P. & Eusebi, V. *Pathology* **41**, 48–56 (2009).

3. Evans, H.L. & Luna, M.A. *Am. J. Surg. Pathol.* **24**, 1319–1328 (2000).

4. Perez-Ordonez, B., Linkov, I. & Huvos, A.G. *Histopathology* **32**, 521–529 (1998).

5. Weinreb, I. *Adv. Anat. Pathol.* **20**, 367–377 (2013).

6. Cappuzzo, F. *et al. J. Natl. Cancer Inst.* **97**, 643–655 (2005).

7. Rozengurt, E. *Physiology (Bethesda)* **26**, 23–33 (2011).

8. Gao, J. *et al. Sci. Signal.* **6**, pl1 (2013).

9. Forbes, S.A. *et al. Nucleic Acids Res.* **39**, D945–D950 (2011).

10. Debnath, J., Muthuswamy, S.K. & Brugge, J.S. *Methods* **30**, 256–268 (2003).

11. Debnath, J. & Brugge, J.S. *Nat. Rev. Cancer* **5**, 675–688 (2005).

12. Weigelt, B., Geyer, F.C. & Reis-Filho, J.S. *Mol. Oncol.* **4**, 192–208 (2010).

13. Natrajan, R. *et al. J. Pathol.* **232**, 553–565 (2014).

14. Shah, S.P. *et al. N. Engl. J. Med.* **360**, 2719–2729 (2009).

15. Tognon, C. *et al. Cancer Cell* **2**, 367–376 (2002).

16. Kohsaka, S. *et al. Nat. Genet.* **46**, 595–600 (2014).

17. Langmead, B. & Salzberg, S.L. *Nat. Methods* **9**, 357–359 (2012).

18. Iyer, M.K., Chinnaiyan, A.M. & Maher, C.A. *Bioinformatics* **27**, 2903–2904 (2011).

19. McPherson, A. *et al. PLoS Comput. Biol.* **7**, e1001138 (2011).

20. Shugay, M., Ortiz de Mendibil, I., Vizmanos, J.L. & Novo, F.J. *Bioinformatics* **29**, 2539–2546 (2013).

21. Ho, A.S. *et al. Nat. Genet.* **45**, 791–798 (2013).

22. Li, H. & Durbin, R. *Bioinformatics* **26**, 589–595 (2010).

23. McKenna, A. *et al. Genome Res.* **20**, 1297–1303 (2010).

24. Cibulskis, K. *et al. Nat. Biotechnol.* **31**, 213–219 (2013).

25. Koboldt, D.C. *et al. Genome Res.* **22**, 568–576 (2012).

26. Thorvaldsdóttir, H., Robinson, J.T. & Mesirov, J.P. *Brief. Bioinform.* **14**, 178–192 (2013).

27. Li, H. *et al. Bioinformatics* **25**, 2078–2079 (2009).

28. Landrum, M.J. *et al. Nucleic Acids Res.* **42**, D980–D985 (2014).

29. Fuentes Fajardo, K.V. *et al. Hum. Mutat.* **33**, 609–613 (2012).

30. ENCODE Project Consortium. *Nature* **489**, 57–74 (2012).

31. Ouedraogo, M. *et al. PLoS ONE* **7**, e50653 (2012).

32. Forbes, S.A. *et al. Nucleic Acids Res.* **38**, D652–D657 (2010).

33. Adzhubei, I.A. *et al. Nat. Methods* **7**, 248–249 (2010).

34. Schwarz, J.M., Rodelsperger, C., Schuelke, M. & Seelow, D. *Nat. Methods* **7**, 575–576 (2010).

35. Reva, B., Antipin, Y. & Sander, C. *Nucleic Acids Res.* **39**, e118 (2011).

36. Shihab, H.A. *et al. Hum. Mutat.* **34**, 57–65 (2013).

37. Carter, H. *et al. Cancer Res.* **69**, 6660–6667 (2009).

38. Mao, Y. *et al. PLoS ONE* **8**, e77945 (2013).

39. Kircher, M. *et al. Nat. Genet.* **46**, 310–315 (2014).

40. Weigelt, B., Warne, P.H. & Downward, J. *Oncogene* **30**, 3222–3233 (2011).

41. Marchiò, C., Dowsett, M. & Reis-Filho, J.S. *BMC Med.* **9**, 41 (2011).

42. Geyer, F.C. *et al. Mod. Pathol.* **24**, 209–231 (2011).

43. Biasini, M. *et al. Nucleic Acids Res.* **42**, W252–W258 (2014).

44. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. *Bioinformatics* **22**, 195–201 (2006).

45. Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L. & Schwede, T. *Nucleic Acids Res.* **37**, D387–D392 (2009).

46. Guex, N., Peitsch, M.C. & Schwede, T. *Electrophoresis* **30** (suppl. 1), S162–S173 (2009).

47. Benkert, P., Tosatto, S.C. & Schomburg, D. *Proteins* **71**, 261–277 (2008).

48. Benkert, P., Schwede, T. & Tosatto, S.C. *BMC Struct. Biol.* **9**, 35 (2009).

49. Benkert, P., Biasini, M. & Schwede, T. *Bioinformatics* **27**, 343–350 (2011).

50. Benkert, P., Kunzli, M. & Schwede, T. *Nucleic Acids Res.* **37**, W510–W514 (2009).

51. Humphrey, W., Dalke, A. & Schulten, K. *J. Mol. Graph.* **14**, 33–38 (1996).

52. Edgar, R.C. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

53. Manning, G., Whyte, D.B., Martinez, R., Hunter, T. & Sudarsanam, S. *Science* **298**, 1912–1934 (2002).

54. Chartier, M., Chenard, T., Barker, J. & Najmanovich, R. *PeerJ* **1**, e126 (2013).

55. Lee, G.Y., Kenny, P.A., Lee, E.H. & Bissell, M.J. *Nat. Methods* **4**, 359–365 (2007).

Figure 1

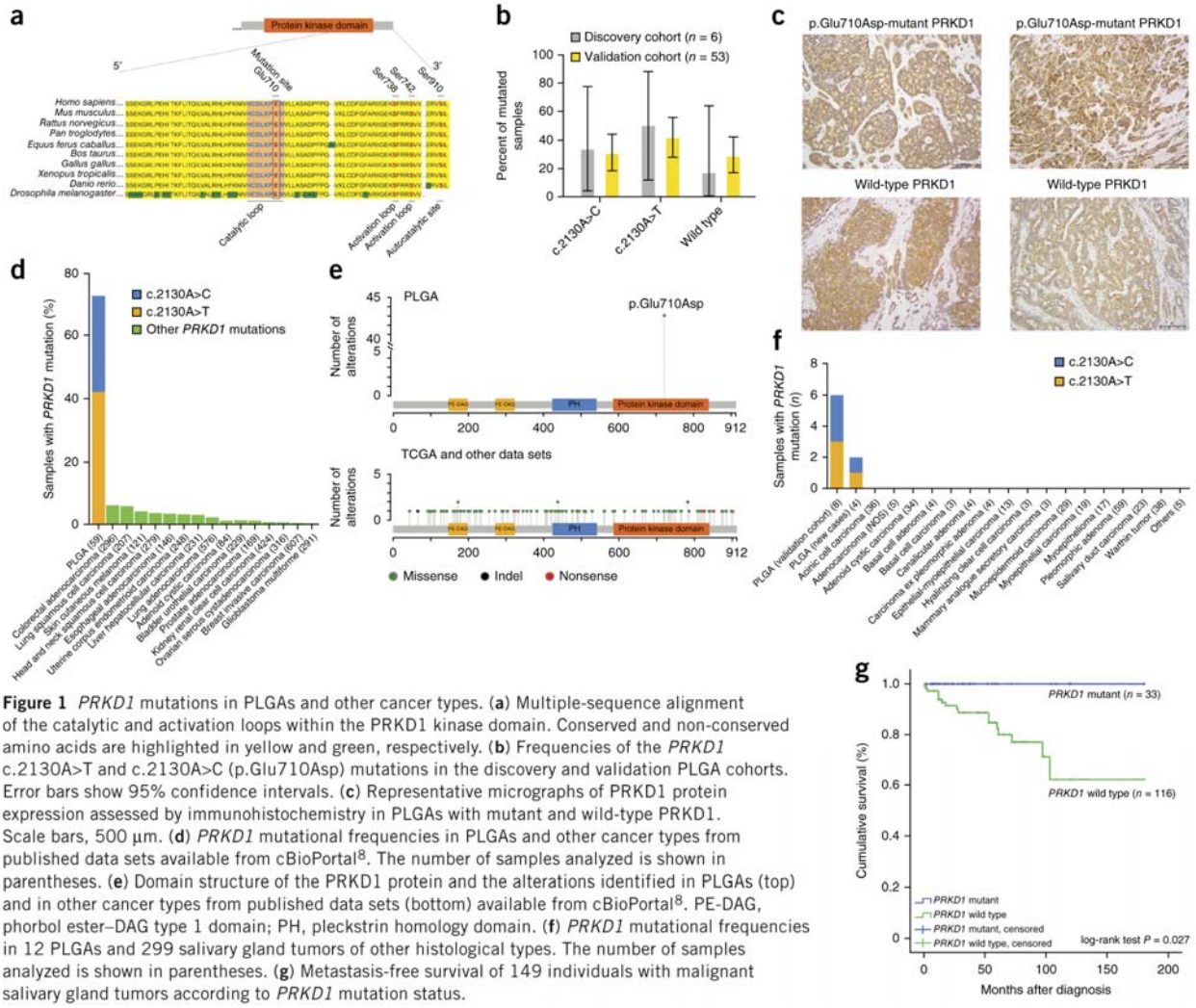

Figure 1 *PRKD1* mutations in PLGAs and other cancer types. (**a**) Multiple-sequence alignment of the catalytic and activation loops within the PRKD1 kinase domain. Conserved and non-conserved amino acids are highlighted in yellow and green, respectively. (**b**) Frequencies of the *PRKD1* c.2130A>T and c.2130A>C (p.Glu710Asp) mutations in the discovery and validation PLGA cohorts. Error bars show 95% confidence intervals. (**c**) Representative micrographs of PRKD1 protein expression assessed by immunohistochemistry in PLGAs with mutant and wild-type PRKD1. Scale bars, 500 µm. (**d**) *PRKD1* mutational frequencies in PLGAs and other cancer types from published data sets available from cBioPortal[8]. The number of samples analyzed is shown in parentheses. (**e**) Domain structure of the PRKD1 protein and the alterations identified in PLGAs (top) and in other cancer types from published data sets (bottom) available from cBioPortal[8]. PE-DAG, phorbol ester–DAG type 1 domain; PH, pleckstrin homology domain. (**f**) *PRKD1* mutational frequencies in 12 PLGAs and 299 salivary gland tumors of other histological types. The number of samples analyzed is shown in parentheses. (**g**) Metastasis-free survival of 149 individuals with malignant salivary gland tumors according to *PRKD1* mutation status.

Figure 2

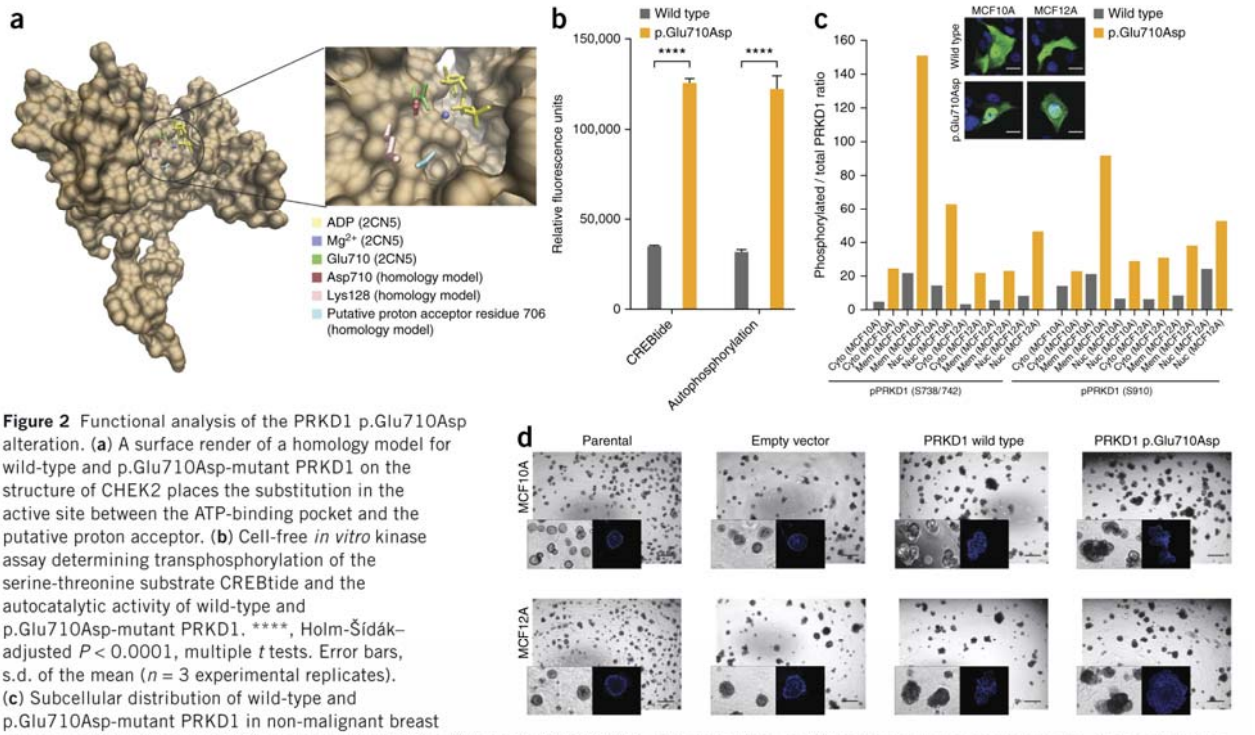

Figure 2 Functional analysis of the PRKD1 p.Glu710Asp alteration. (**a**) A surface render of a homology model for wild-type and p.Glu710Asp-mutant PRKD1 on the structure of CHEK2 places the substitution in the active site between the ATP-binding pocket and the putative proton acceptor. (**b**) Cell-free *in vitro* kinase assay determining transphosphorylation of the serine-threonine substrate CREBtide and the autocatalytic activity of wild-type and p.Glu710Asp-mutant PRKD1. ****, Holm-Šídák–adjusted $P < 0.0001$, multiple *t* tests. Error bars, s.d. of the mean ($n = 3$ experimental replicates). (**c**) Subcellular distribution of wild-type and p.Glu710Asp-mutant PRKD1 in non-malignant breast epithelial MCF10A and MCF12A cells and Ser738/Ser742 and Ser910 PRKD1 phosphorylation in distinct subcellular compartments. Cyto, cytoplasm; mem, membrane; nuc, nuclear. Scale bars, 25 μm. (**d**) Representative micrographs of the impact of wild-type and p.Glu710Asp-mutant PRKD1 expression on the growth and glandular architecture of MCF10A and MCF12A cells grown in three-dimensional basement membrane cultures. Scale bars, 500 μm. Insets, representative micrographs of three-dimensional structures (2.5× magnification relative to the original micrographs) and fluorescence micrographs of individual DAPI-stained three-dimensional structures.