# KCNQ GENE FAMILY MEMBERS ACT AS BOTH TUMOR SUPPRESSORS AND ONCOGENES IN GASTROINTESTINAL CANCERS

*David Shorthouse[1], Eric Rahrmann[2], Cassandra Kosmidou[1], Benedict Greenwood[1], Michael Hall[1], Ginny Devonshire[2], Richard Gilbertson[2], Rebecca C. Fitzgerald[1], Benjamin A Hall[1*]*

[1]MRC Cancer Unit,
University of Cambridge,
Hutchison/MRC Research Centre,
Box 197,
Cambridge Biomedical Campus,
Cambridge,
CB2 0XZ

[2]Cancer Research UK Cambridge Institute
University of Cambridge
Li Ka Shing Centre
Robinson Way
Cambridge CB2 0RE
*to whom correspondence should be addressed.
Email: bh418@mrc-cu.cam.ac.uk

Short title: KCNQ genes in gastrointestinal cancer

## SUMMARY

We present evidence that KCNQ genes are drivers and suppressors of gastrointestinal (GI) cancer in humans. The KCNQ family of genes encode for subunits of a potassium channel complex involved in membrane polarisation and little is known about their role in cancer. We use human cancer data and a multidisciplinary computational-based approach including structural modelling and simulation, coupled with *in vitro* experiments to show that KCNQ1 is a tumor suppressor, and KCNQ3 and KCNQ5 are oncogenic across human GI cancers. We link the expression of KCNQ genes to WNT signalling, EMT, and survival and propose that mutation/copy number alteration of KCNQ genes can significantly alter patient prognosis in GI cancers.

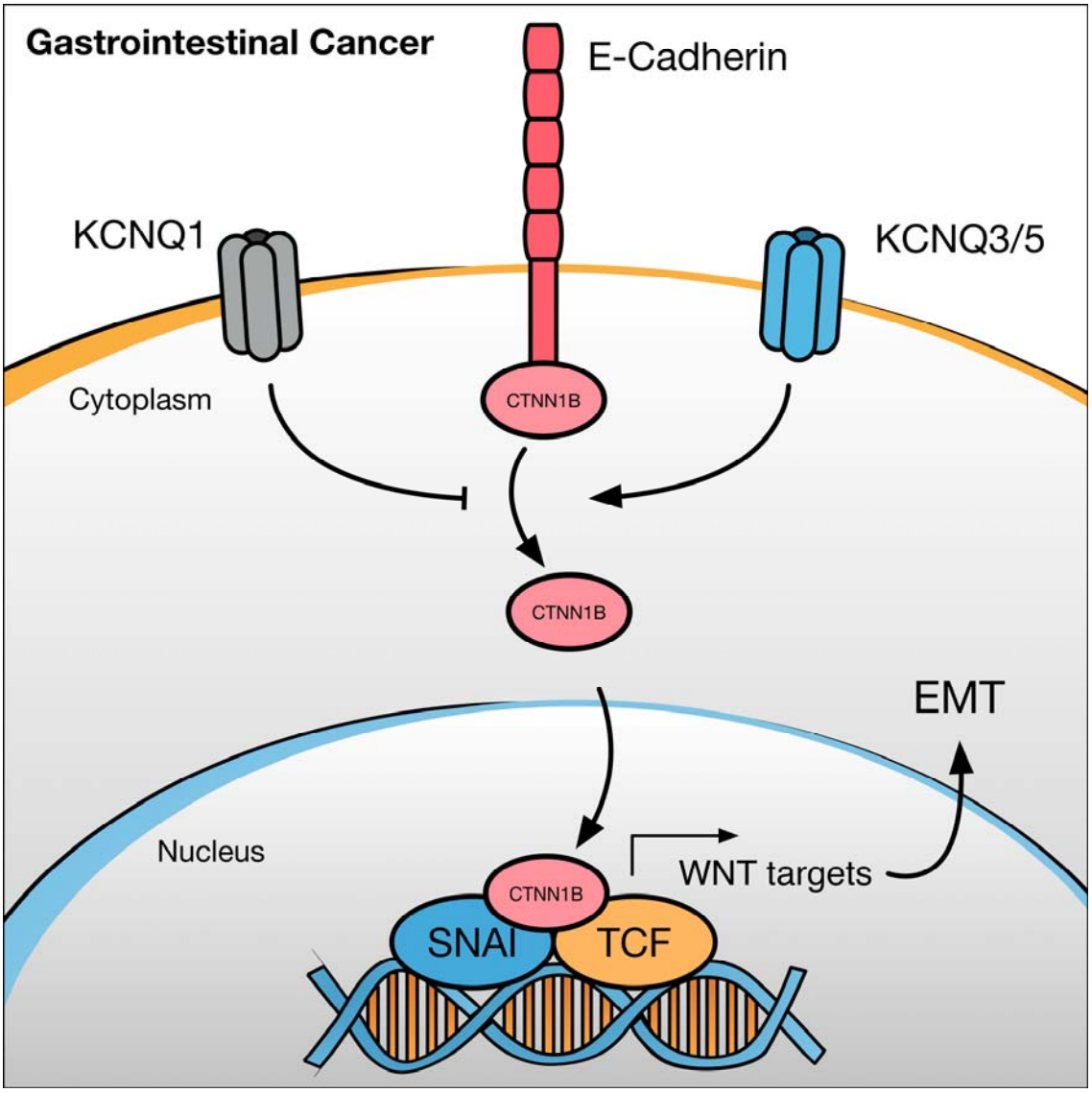
(110 words)



## KEYWORDS

KCNQ, Ion Channels, Gastrointestinal Cancer, Epithelial to Mesenchymal Transformation (EMT), WNT signalling.

## GRAPHICAL ABSTRACT



## SIGNIFICANCE

Our results show that the ion channel family KCNQ contribute to progression of gastrointestinal cancer. KCNQ genes are highly mutated and detected mutations select for inactivation of KCNQ1, but increased activity of KCNQ3. Analysis of gene expression profiles uncovers KCNQ expression correlating with the WNT pathway and EMT. Moreover, clinical data correlates with expression data showing KCNQ1 has hallmarks of a tumor suppressor, whilst KCNQ3 and KCNQ5 are oncogenic. These findings implicate KCNQ genes in control/mediation of the WNT pathway in cancer, and potentially as prognostic markers or therapeutic targets.

(90 words)

## INTRODUCTION

The KCNQ family of ion channels are a set of evolutionarily related genes encoding for proteins involved in potassium transport (Abbott, 2014; Robbins, 2001; Wang and Li, 2016). KCNQ channels typically repolarise the plasma membrane of a cell after depolarisation through other channels. KCNQ are therefore involved in wide ranging biological functions such as cardiac action potentials (Nerbonne and Kass, 2005; Robbins, 2001), hearing sensitivity (Kharkovets et al., 2002; Kubisch et al., 1999), neural excitability (Brown and Passmore, 2009; Wang and Li, 2016), and ionic homeostasis in the gastrointestinal tract (Ohya et al., 2015; Warth et al., 2002).

Diseases resulting from loss or gain-of-function (LOF and GOF respectively) mutations in the KCNQ family are wide ranging, and include epilepsy (Allen et al., 2014; Millichap et al., 2016; Rogawski, 2000), long and short QT syndrome (Morita et al., 2008), deafness (Kubisch et al., 1999), and more recently, Autism-like disorders (Sands et al., 2019). Neurological disorders caused by KCNQ mutations are an active area of research, and specific mutations in many individuals have been characterised using electrophysiological methods - many mutations therefore have prior functional characterisation as LOF of GOF (Miceli et al., 2008, 2015; Panaghie and Abbott, 2007; Sands et al., 2019). Whilst some of the mechanisms by which these diseases occur have been partially elucidated, there are complications in that the KCNQ family of proteins interact heavily with the KCNE family, to form complex heteromeric combinations of varying subunits that are not completely understood. In particular, KCNQ1 interacts with all known KCNE ancillary proteins in varying tissues, but is otherwise homotetrameric (Abbott, 2014). KCNQ2, 3, 4, and 5 however, can interact with each other and the KCNE family to theoretically form hundreds of combinations of channels, and the impact of specific mutations on any

one member will have an unknown influence on the greater function of the pool of channels (Howard et al., 2007; Yu et al., 2013).

Previous work has further highlighted the potential roles of individual ion channels across a broad range of cancers, including TRPA1 in breast cancer (Takahashi et al., 2018), HERG channels in leukaemia (Pillozzi et al., 2002), and osmotic regulatory machinery in several cancers (Haas and Sontheimer, 2010; Pedersen et al., 2013). The complex interplay between ionic gradients and cellular phenotypes has been studied at a high level before, including in cancer (Pardo and Stühmer, 2013; Pedersen and Stock, 2013; Shorthouse et al., 2018), but the functional contributions of the KCNQ family to the phenotype of non-excitable tissue is not fully known.

There is preliminary evidence to suggest that KCNQ1 plays a tumor suppressive role in the stomach and colon (Rapetti-Mauss et al., 2017; Than et al., 2014), and hepatocellular carcinoma (Fan et al., 2018), and that KCNQ3 potentially plays a role in oesophageal adenocarcinoma (Frankell et al., 2019). It is of particular interest to study the roles of ion channels in the gastrointestinal tract due to the critical nature of the ionic homeostasis that is required for their varied physiological functions. Cells in the gastrointestinal tract are often involved at the interface between the tissue and the luminal environment, and as such membrane transport is critical to their homeostatic function. The gradients electrochemical ion channels generate are used in protective mucus production (Tarran, 2004), acid secretion (Grahammer et al., 2001), and immune recognition (Feske et al., 2015), among other processes. Additionally, there is some evidence that activity of membrane transporters may play

roles in key cancer pathways, such as a reported interaction between KCNQ1 and beta catenin (Rapetti-Mauss et al., 2017).

With the wide availability of publicly available databases such as The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/) and the International Cancer Genome Consortium (ICGC) (https://icgc.org/) it is possible to analyse samples from large populations of cancer patients in terms of gene expression, mutations, copy number variation, and gene methylation. Whilst in many cases, mutations have been previously characterised as LOF or GOF, in cases where specific data is not available to categorise the mutation, structures of the protein become critically important in studying their potential effects. Where structures are not available for a particular gene, homology modelling is a useful method for generating consistently reliable predictive structures and allow the study and simulation of the spatial context of individual mutations (Šali and Blundell, 1993; Šali et al., 1995).

In this study, we aimed to investigate the role of the KCNQ family in gastrointestinal cancer through study of highly annotated clinical data sets with DNA and RNA sequencing data paired with structural modelling of KCNQ proteins.

# RESULTS

KCNQ genes are highly genetically altered in GI cancers

We studied the mutations in the KCNQ and related KCNE family across all cancers using the TCGA database. We calculated the missense mutational frequency of all KCNQ genes within each TCGA cohort, and grouped them into previously defined (Hoadley et al., 2018) subgroups (**Figure 1 A**). Ranking subgroups by percentage of patients that have a mutation in a KCNQ/E gene shows enrichment for the core GI subgroup when compared to other subgroups, a higher overall mutational burden is seen only in the melanoma cohort, which has a significantly higher baseline mutational burden (Chalmers et al., 2017). Comparison of the rates of synonymous to nonsynonymous mutations within the core GI cohort potentially indicates the presence of mutational selection, there are significantly more nonsynonymous mutations when compared to synonymous.

We further studied all genetic alterations in the KCNQ and related KCNE family across GI cancers using the publicly available TCGA and our own Oesophageal Adenocarcinoma data which has detailed clinical annotation (OCCAMS). The GI cancers studied were: Oesophageal Squamous Cell Carcinoma (ESCC, n = 103),; Oesophageal Adenocarcinoma in two cohorts: the TCGA (EAC-TCGA, n = 93)(Kim et al., 2017) and our own data (EAC-OCCAMS part of ICGC, n = 378); Stomach Adenocarcinoma (STAD-TCGA, n = 426) (Bass et al., 2014), and Colorectal Adenocarcinoma (COADREAD-TCGA, n = 594)(Muzny et al., 2012). We found that 31% of patients with GI cancers (n = 1594) had genetic alterations in the KCNQ/E families of genes. 30-40% of patients in all subtypes contain a mutation or copy number alteration (CAN), except for COADREAD-TCGA, in which 26% of patients are altered for KCNQ/E. The ratio of amplifications to mutation and deletion are also

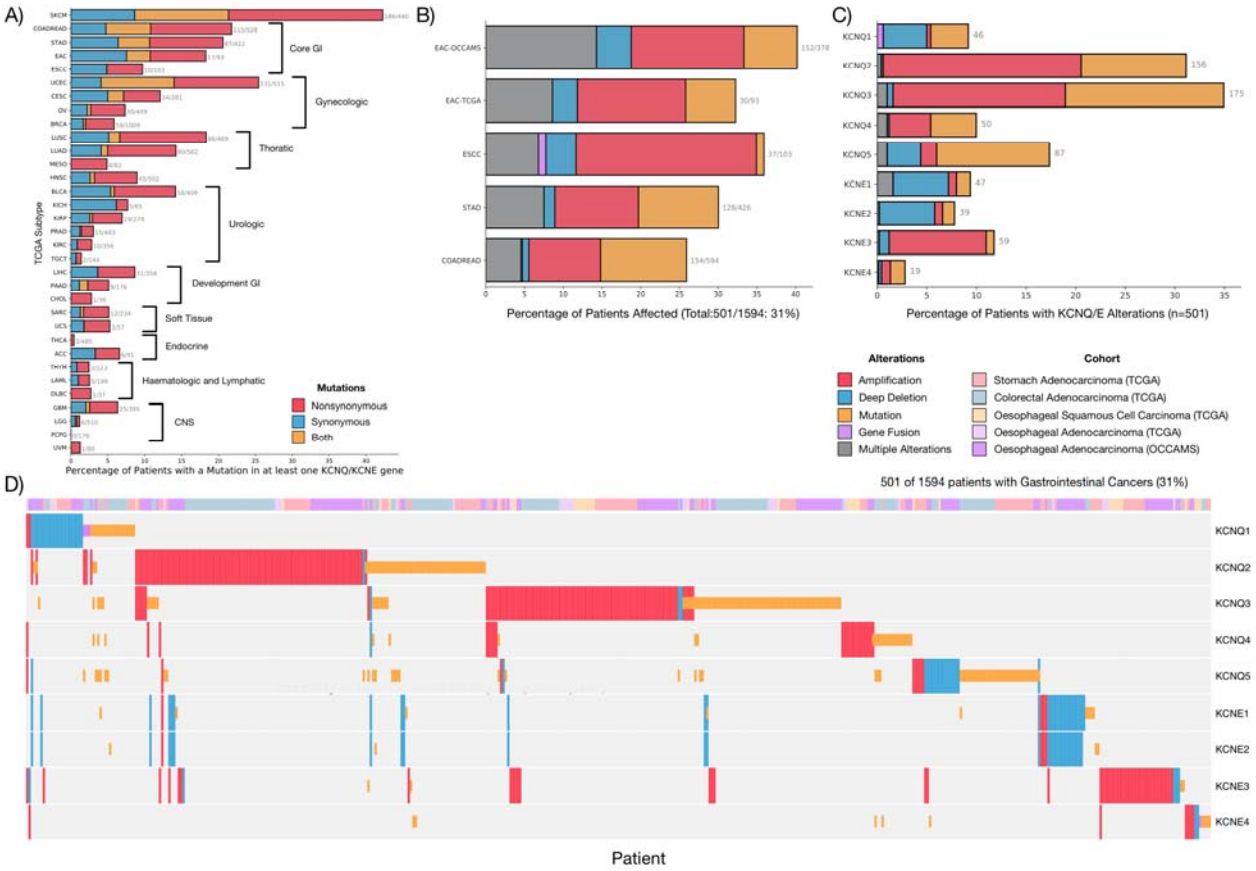
conserved between cohorts aside from ESCC-TCGA, which mostly contains amplification events (**Figure 1 B**). Comparing the copy number profiles, KCNQ1, KCNE1 and KCNE2 appear mostly to be deleted overall, whereas KCNQ2, KCNQ3, and KCNE3 are generally amplified. In terms of SNVs we find that KCNQ2, KCNQ3, and KCNQ5 are heavily mutated (**Figure 1 C**). Interestingly, 175 (11%) of all patients have a mutation/copy number change in KCNQ3, the most altered member of the family. The alterations are generally distributed across tissue types (**Figure 1 D**), and there is no observed correlation between mutations in the KCNQ/E family and cancer stage where annotated (**Figure S1**).

Co-occurrence analysis with common GI cancer drivers (CCND1, CDKN2A, CTNNB1, ERBB2, KRAS, PTEN, and TP53) using DISCOVER (Canisius et al., 2016) show that many members of the KCNQ family are mutually exclusive with common drivers (**Table S1**), in particular we find that KCNQ2 and KCNQ5 are statistically significantly ($q <= 0.1$) mutually exclusive with CCND1 and CDKN2A, and KCNQ2 and KCNQ3 are mutually exclusive with both KRAS and TP53, indicating that these mutations tend to occur separately, potentially revealing that co-occurrence of these mutations is unfavourable to overall cell survival (Campbell, 2017).

RNA expression comparison between cancer and normal samples reveals that KCNQ3 is generally upregulated in GI cancers compared to normal tissue, and KCNQ1 shows a range of altered expression, but generally has its expression reduced (**Figure S2**).

Analysis of methylation data also indicates that the CPG promoters of the KCNQ family are often differently methylated when compared with their normal counterparts

(**Figure S3**), KCNQ5 appears hypermethylated, consistent with a decrease in gene expression. We also find that KCNQ1 and KCNQ3 show methylation patterns opposite to their general change in gene expression, with a hypermethylation of CPG promoter of KCNQ3 observed, despite an increase in RNA expression, and a general decrease in the methylation of CPG promoter in KCNQ1. This indicates the gene expression is likely altered through means other than methylation in the case of KCNQ1 and KCNQ3 such as mutation of the promoter region.



Mutations in the KCNQ family show evidence of selection

In order to study mutations in the KCNQ family across gastrointestinal cancers we chose to look in depth at the COSMIC database of somatic mutations in cancers, selecting for mutations occurring in any tissue along the gastrointestinal tract. The COSMIC database aggregates mutations from multiple sources, including TCGA and ICGC, and so is a comprehensive database for mutations across human cancers.

We performed a MUSCLE alignment of protein sequences for members of the KCNQ family to allow alignment of sequences into conserved domains, primarily transmembrane helices.

We find that there are regions of increased mutational frequency in different genes within the KCNQ family (**Figure 2 A**). In particular, KCNQ1 shows increased frequency of mutations in the S2-S3 helical region, and the S5-pore-S6 region of the protein. KCNQ3 shows a clear region of increased mutational burden within the S4 voltage sensing helix, and a second region in the Calmodulin-binding HB-HC helices. There are regions of high mutational burden also present within the S6 helix for KCNQ2, and the S4 helix for KCNQ5 potentially indicating mutational selection.

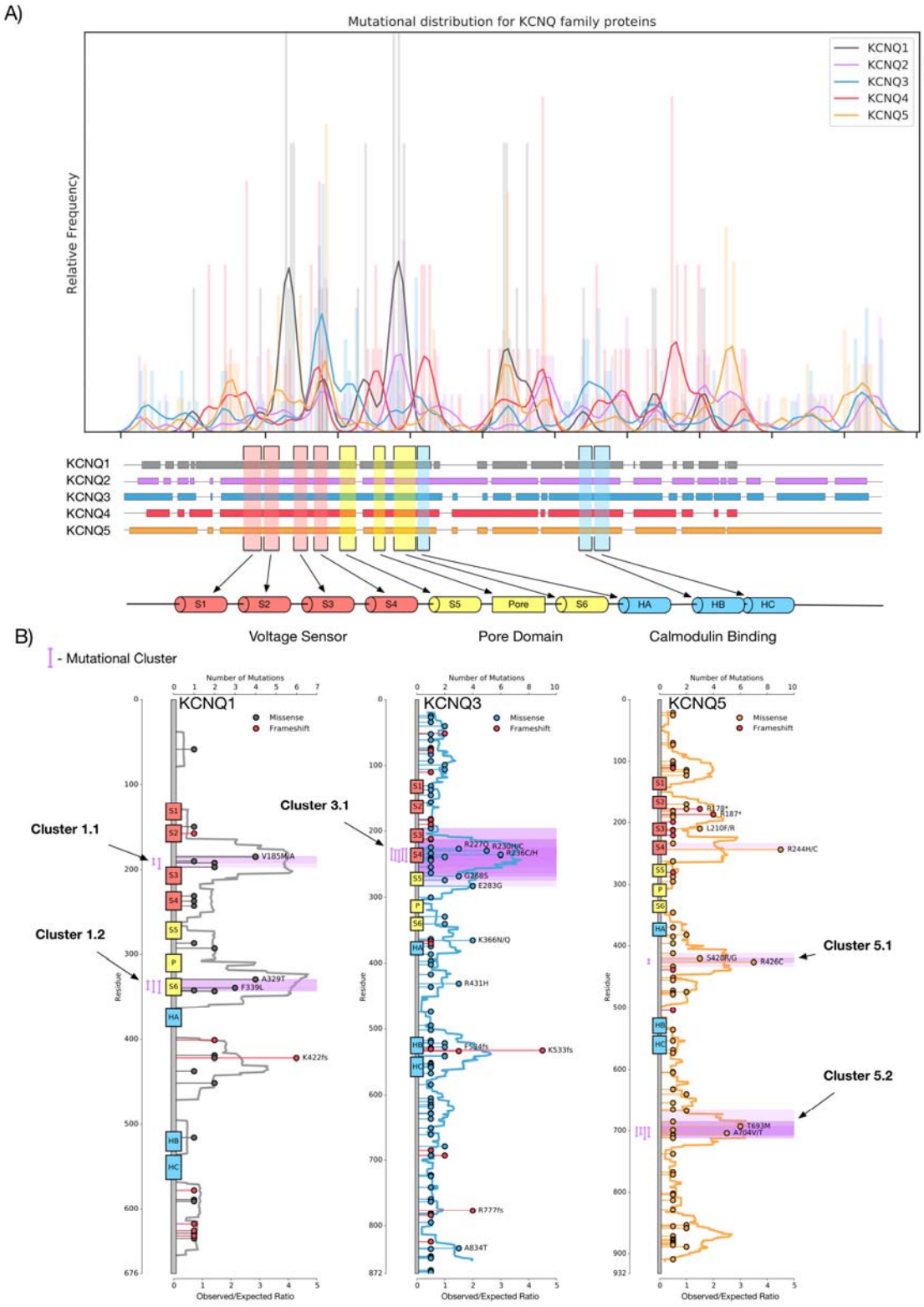
Fathmm scores calculated for KCNQ1-5 by COSMIC for mutations observed across GI cancers show that 65% of all mutations observed are predicted with high confidence to be pathogenic (**Table S2**). We additionally calculated the cancer-associated CScape scores (Rogers et al., 2017) for all possible mis-sense mutations in KCNQ1 and KCNQ3, the two genes showing highest relative mutational clustering. We find evidence that particular structural regions, with good overlap with those regions showing an increased mutational burden in GI cancer, have a high proportion of predicted pathogenic mutations. (**Figure S4**). Additionally there is a

significantly larger number of missense vs nonsense mutations within the domains of all KCNQ genes (**Figure S5**).

To statistically explore regions of mutational clustering within the sequence of each member of the KCNQ family, we applied the NMC method (Ye et al., 2010) to look for potential clusters of mutations occurring more often than expected by chance within the 1D sequence (**Figure 2 B**). We additionally overlay a mutational signature-based observed vs expected mutation ratio applied across a sliding window of the protein sequence, similar to the dn/ds method previously applied to mutational selection in cancer (Martincorena et al., 2017). We find that for KCNQ1, there is a clear significant cluster of mutations within the S2-S3 linker region (cluster 1.1), and within the S6 helix (cluster 1.2). KCNQ3 showed a significant region of mutational selection within the S4 voltage sensing helix, particularly arginine residues 227, 230, and 236 (cluster 3.1) KCNQ5 shows a major region of mutational clustering within the S4 helix, and a structurally and functionally undefined region at residue 680-700,. For KCNQ2, we identified a region of increased mutational frequency around the S6 helix, and no regions of mutational selection within KCNQ4 (**Figure S6**).

Finally, we performed a literature search for functional characterisation of observed mutations. Many of the mutations observed in the COSMIC database are functionally characterised mutations in other KCNQ-related disorders (**Table 1**).

Overall we find mutations in KCNQ1 appear to cluster primarily in the S2/S3 and S6 regions and mutations to KCNQ3 and 5 show similar clusters within the S4 helix.


A)
Mutational distribution for KCNQ family proteins
KCNQ1
KCNQ2
KCNQ3
KCNQ4
KCNQ5
Relative Frequency
S1
S2
S3
S4
S5
Pore
S6
HA
HB
HC
Voltage Sensor
Pore Domain
Calmodulin Binding
B)
- Mutational Cluster
Number of Mutations
KCNQ1
KCNQ3
KCNQ5
Missense
Frameshift
Cluster 1.1
Cluster 1.2
Cluster 3.1
Cluster 5.1
Cluster 5.2
V185M/A
A329T
F339L
K422fs
R227Q
R230H/C
R236C/H
G268S
E283G
K366N/Q
R431H
F584fs
K533fs
R777fs
A834T
R178*
R187*
L210F/R
R244H/C
S420R/G
R426C
T693M
A704V/T
Residue
Observed/Expected Ratio

255

256

Mutations to KCNQ genes are clustered in space

The structural assembly of the KCNQ/KCNE family is complex – KCNQ1 is known to assemble into homotetramers, which are modulated by members of the KCNE gene family, whereas KCNQ2, 3, 4, and 5 have varying stoichiometries, and are also modulated by KCNE gene products (**Figure 3 A**). We performed a bayesian inference analysis to study the gene expression correlations between members of the KCNQ/KCNE family in different gasto-intestinal tissues. In particular for the COADREAD RNAseq dataset **(Figure 3 B**) we find two clusters of genes that positively correlate with each other, this effect is observed to a lesser extent in other tissue specific analysis (**Figure S7**). The two clusters that show negative correlation to each other contain KCNQ1, KCNE2, and KCNE3 (Set 1), and KCNQ3, KCNQ4, KCNQ5, KCNE1, and KCNE4 (Set 2). KCNQ1 and KCNE2/KCNE3 complexes have been observed in vivo wild-type gastrointestinal tissue, and are essential for the homeostatic control of proton and potassium gradients (Grahammer et al., 2001; Heitzmann et al., 2004). This analysis potentially indicates an antagonistic relationship between the two clusters.

In order to further demonstrate functional effects of mutations on the structure of KCNQ genes, we chose to structurally model members of the KCNQ family. KCNQ proteins contain 6 transmembrane helices (shown schematically in **Figure 3 C**, shown structurally in **Figure 3 D,** full structure shown in **Figure S8**). Helices S1, S2, S3, and S4 (red) make up a voltage sensor domain, whose conformation changes in response to voltage shifts across the membrane control the gating of channel through positively charged arginine residues in the S4 helix. The S5, pore, and S6

domain (yellow) contain the gating components of the channel, additionally there are three helices HA, HB, and HC (blue), which make up a structure to which calmodulin binds and modulated the channel.

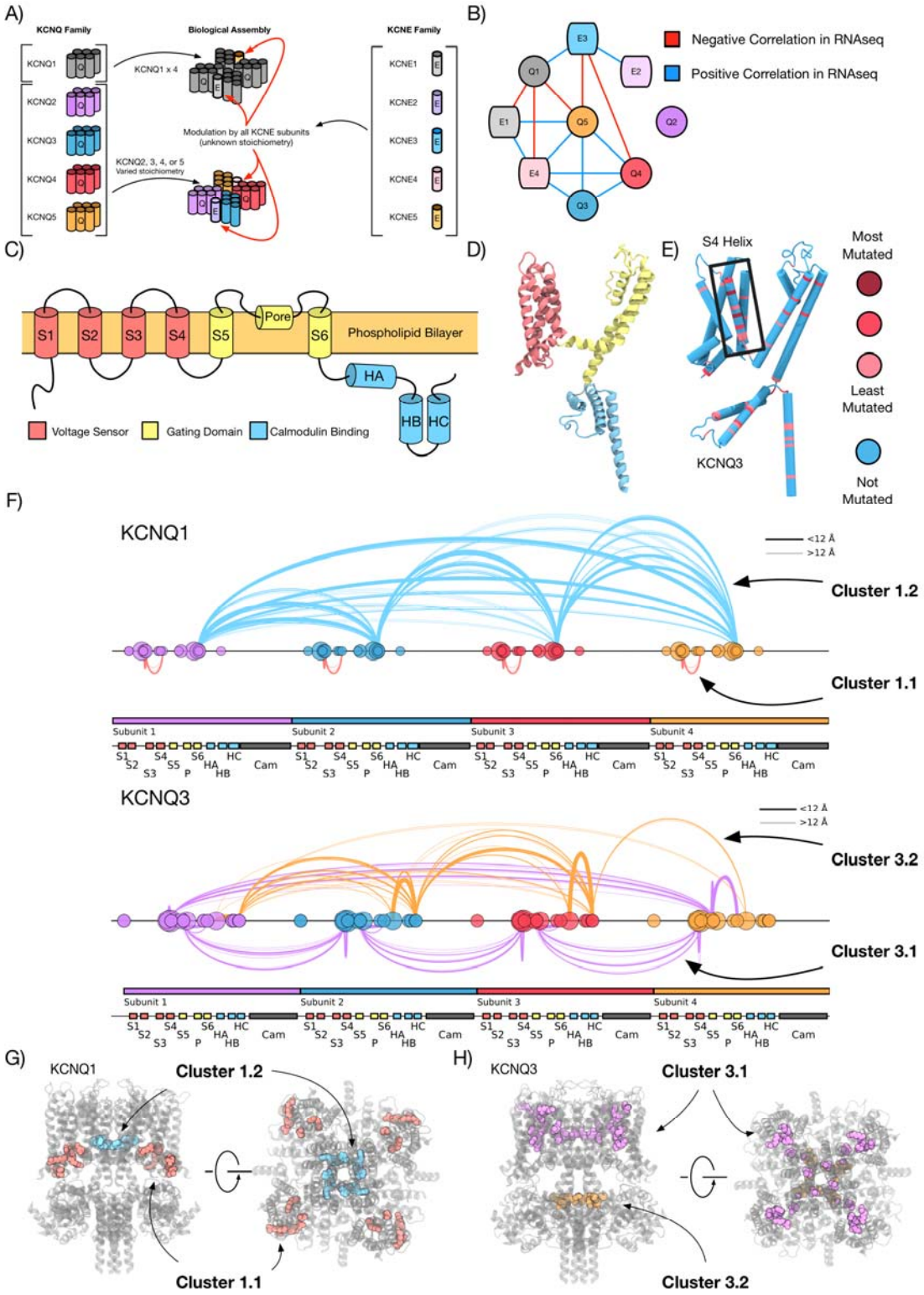
Homology models of the human proteins of each member were generated from the cryo-em structure of *Xenopus laevis* KCNQ1 5VMS. Models exhibit high nativity scores (**Table S3**), and were validated as viable by 100ns of equilibrium atomistic level molecular dynamics (**Figure S8**).

Mapping COSMIC mutations onto the resultant structures and colouring by frequency reveals structural regions with a high frequency of mutations. KCNQ3 in particular shows a high number of mutations clustered in the S4 helix as expected (**Figure 3 E**). Other members show some mutational clustering (**Figure S9**).

To explore the spatial clustering of mutations quantitatively within the predicted structures we calculated clusters of colocalized mutations within each structure. Mutations within a 12 Å cutoff we grouped together, we find that each structure contains mutational clusters that give more functional information than 1D analysis alone. For KCNQ1 we find two distinctive spatial clusters, those in the S2/S3 linker domain, and a large interlinked cluster of residues around the S6 helix (**Figure 3 F**). Mutations in cluster 1.1 are clearly part of the previously defined phosphatidylinositol (PIP) binding site, essential to channel function (Zaydman and Cui, 2014). Mutations in cluster 1.2 converge in the centre of the protein and appear to cluster around the pore restriction (**Figure 3 G**). For KCNQ3, we find that the previously defined

mutational cluster in the S4 helix, denoted cluster 3.1, makes up a large region of spatially clustered mutations within the S4/S5 helices of the assembled tetramer (**Figure 3 H**). We additionally identify a further region of mutational clustering within the HA/HB helices of the protein, which are spatially close within the full tetrameric protein, but not within the 1D structure (cluster 3.2).

Overall, clustered mutations in KCNQ1 appear to mainly influence cofactor binding (cluster 1.1) and the pore of the protein (cluster 1.2), whereas KCNQ3 mutational clusters are focussed in the activation domain (cluster 3.1).


A)
KCNQ Family
Biological Assembly
KCNE Family
KCNQ1
KCNQ2
KCNQ3
KCNQ4
KCNQ5
KCNQ1 x 4
KCNQ2, 3, 4, or 5
Varied stoichiometry
Modulation by all KCNE subunits
(unknown stoichiometry)
KCNE1
KCNE2
KCNE3
KCNE4
KCNE5
B)
Negative Correlation in RNAseq
Positive Correlation in RNAseq
C)
S1
S2
S3
S4
S5
Pore
S6
Phospholipid Bilayer
HA
HB
HC
Voltage Sensor
Gating Domain
Calmodulin Binding
D)
E)
S4 Helix
Most Mutated
Least Mutated
Not Mutated
KCNQ3
F)
KCNQ1
<12 Å
>12 Å
Cluster 1.2
Cluster 1.1
Subunit 1
Subunit 2
Subunit 3
Subunit 4
Cam
KCNQ3
Cluster 3.2
Cluster 3.1
G)
KCNQ1
Cluster 1.2
Cluster 1.1
H)
KCNQ3
Cluster 3.1
Cluster 3.2

313

314

Mutations to KCNQ genes in GI cancers are functionally relevant

Due to the spatial clustering of mutations in the S6 gating helix and the S4 voltage sensing helix in KCNQ1 and KCNQ3 respectively we sought to investigate the functional effects of these mutations on the activity of the assembled channels.

For mutations in the KCNQ1 S6 helix, we generated additional homology models for each of the mutations found to be clustered in that region (Cluster 1.2 - F339L, L342F, P343L, and P343S), and an additional nearby mutation that is the most frequently observed KCNQ1 mutation across GI cancers (A329T). Analysis of the pore domain of the protein (**Figure 4 A**), shows that all mutations except F339L are predicted to occlude the pore, reducing or eliminating its ability to gate potassium ions, even when a single subunit is mutated, as may be the case when a patient has a KCNQ1 mutation in only a single allele. We conclude that mutations in cluster 1.2 for KCNQ1 are likely loss-of-function.

The S4 domain of KCNQ channels is particularly important due to its involvement in channel gating. It is of special interest because it is a “mutational hotspot” in KCNQ2 mutation induced epilepsies, and strong prior evidence exists for mutations in S4 controlling channel activity (DeMarco, 2012; Miceli et al., 2008, 2015; Sands et al., 2019). The positively charged arginines within S4 are the principal drivers of the proteins response to membrane polarity. We find a significant number of mutations in KCNQ2, KCNQ3, and KCNQ5 are within arginines in S4, with KCNQ3 containing mutations in every one of the 5 arginines (**Figure 4 C**). The S4 helix of all KCNQ channels includes 4 (KCNQ1) or 5 (KCNQ2-5) positively charged arginine residues essential for gating and activation of a potassium current, similarly to other

potassium channels (Aggarwal and MacKinnon, 1996; Jentsch, 2000). We refer to these arginines by convention in ascending order from N to C terminus as R1, R2, R4, R5, and R6. The position generally referred to as R3 is replaced with a glutamine in the KCNQ family and is therefore not considered. We find that R1, R2, and R4 are highly mutated in KCNQ3, and R4 and R5 are mutated in KCNQ2, whilst KCNQ5 shows 9 total mutations in R5 only.

We chose to analyse individual mutations in the helix using a previously developed tool utilizing molecular dynamics, Sidekick (Hall et al., 2014) (**Figure S10**), that simulates the effects of single mutations on the properties of an alpha helix. We surmised that single helical positioning in the membrane can act as a proxy for the forces the helix exerts on the full protein upon mutation, without costly simulations of the entire mutant protein. We chose to first replicate an experimentally determined alanine scanning experiment on the KCNQ1 S4 helix (Panaghie and Abbott, 2007), and a tryptophan scanning experiment of the KCNQ2 S4 helix, before simulating mutations in the KCNQ S4 helices from both cancer and epilepsy, some of which have previously reported experimental characterization.

Alanine scanning experiments (Panaghie and Abbott, 2007) revealed that mutations of the R2 and R4 positions of KCNQ1 to alanine resulted in constitutively active, or more readily activated currents, whereas mutations to alanine at R1, R6, or D242 (referred to as D6) impaired the activation of KCNQ1. Calculating helical tilt within a DPPC membrane reveals mutations to alanine at R2 and R4 decrease the tilt angle of S4, and has no effect (R1) or an increased angle for R6 and D6 mutations. A decrease in tilt angle within Sidekick analysis therefore correlates with increased activity of the channel (**Figure 4 D**). We additionally replicated a glutamine scanning experiment performed on the KCNQ2 S4 helix (Miceli et al., 2008) (**Figure S11**). We

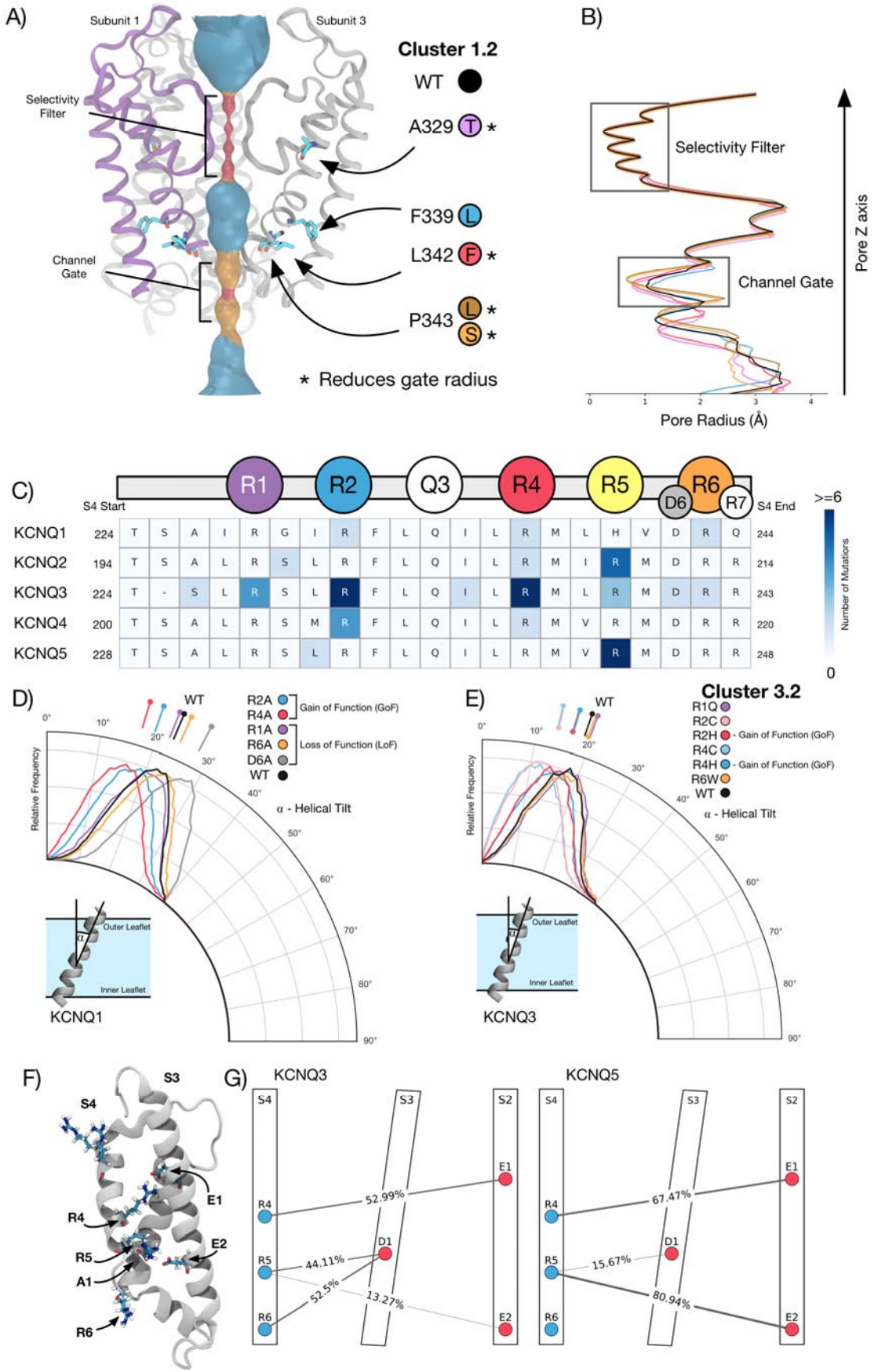
find that sidekick correctly identifies a decrease in tilt angle for constitutively active mutant R2Q but not for R1Q, which has a reduced activation threshold. We also find no change in tilt angle for mutants with a reduced activation threshold (R4Q, R5Q, R6Q, R7Q).

We simulated all arginine mutations in KCNQ3 including 3 that occur within our dataset, and additional mutations that are recently validated as gain of function causes of autism related disorders (Sands et al., 2019). Simulations agree with previous data, indicating a decrease in helical tilt angles for R2H and R2C, known gain-of-function mutations, and predicting that similar properties occur with unvalidated mutations R236C, and R236H (**Figure 4 E**). Overall we find that 14/17 (82%) of all mutations within the KCNQ3 S4 helix, and by extension involved in mutational cluster 3.1 are either previously validated gain of function, or predicted by our simulations to be gain of function (**Table 2**).

Finally we sought to examine the large number of mutations in the R5 residue of KCNQ2 and KCNQ5, simulations do not show any change in helical tilt upon mutation to H or C (**Table 2, Figure S11**), but looking at the arginines in the context of the full protein in molecular dynamics simulations shows that hydrogen bond occupancy for residues involved in stabilizing the active form of the protein are different. Residues thought essential for stabilizing the active form of the channel (Sands et al., 2019; Soldovieri et al., 2019) are R4, R5, R6, an aspartic acid in helix S2 (D1), and two glutamic acids within helix S3 (E1, and E2) (**Figure 4 F**). We find the KCNQ5 bonding network differs in occupancy compared to KCNQ2-4. R5 is principally involved in a hydrogen bonding interaction with D1 in KCNQ2-4, with occupancies ranging from 40% to 62% (**Figure S12**), but KCNQ5 R5 shows only 16% occupancy in bonding with D1, and instead is found to have an 81% occupancy

with E2 (**Figure 4 G**). Disruption of this hydrogen bonding network by mutation is likely to have functional consequences on the activity of the protein, particularly as the primary mutation to KCNQ5 R5 is to a cysteine, which could potentially form a disulphide bond with the unpaired cysteine R203 situated ~7Å from R5 in our model.

Overall we find evidence that many mutations observed in the KCNQ in GI cancer are of functional consequence. In particular, we find that mutations in Cluster 1.2 of KCNQ1 are mostly predicted to be loss-of-function, whilst those in Cluster 3.1 of KCNQ3 are mostly gain-of-function. We additionally find predicted functional consequences for many other mutations observed in the KCNQ family. This reinforces prior data showing that KCNQ1 appears tumor suppressive, and can be mutationally inactivated in cancers, whilst KCNQ3 appears to primarily accumulate gain-of-function mutations.


A)
Subunit 1
Subunit 3
Selectivity Filter
Channel Gate
Cluster 1.2
WT
A329 T *
F339 L
L342 F *
P343 L *
S *
* Reduces gate radius
B)
Selectivity Filter
Channel Gate
Pore Z axis
Pore Radius (Å)
C)
S4 Start
R1
R2
Q3
R4
R5
D6
R6
R7
S4 End
>=6
Number of Mutations
0
KCNQ1 224 T S A I R G I R F L Q I L R M L H V D R Q 244
KCNQ2 194 T S A L R S L R F L Q I L R M I R M D R R 214
KCNQ3 224 T - S L R S L R F L Q I L R M L R M D R R 243
KCNQ4 200 T S A L R S M R F L Q I L R M V R M D R R 220
KCNQ5 228 T S A L R S L R F L Q I L R M V R M D R R 248
D)
WT
R2A
R4A
Gain of Function (GoF)
R1A
R6A
D6A
Loss of Function (LoF)
WT
Relative Frequency
α - Helical Tilt
Outer Leaflet
Inner Leaflet
KCNQ1
E)
Cluster 3.2
WT
R1Q
R2C
R2H - Gain of Function (GoF)
R4C
R4H - Gain of Function (GoF)
R6W
WT
Relative Frequency
α - Helical Tilt
Outer Leaflet
Inner Leaflet
KCNQ3
F)
S3
S4
E1
R4
E2
R5
A1
R6
G)
KCNQ3
S4
S3
S2
E1
R4
52.99%
D1
R5
44.11%
52.5%
13.27%
R6
E2
KCNQ5
S4
S3
S2
E1
R4
67.47%
D1
R5
15.67%
80.94%
R6
E2

KCNQ gene expression correlates with WNT signalling and EMT in GI cancer

To interrogate the functional and signalling properties of KCNQ genes in cancer, and in particular to identify potential interactions with known cancer-associated pathways, we looked at the RNAseq data associated with both TCGA and the OCCAMS cohorts. We chose to study KCNQ1, KCNQ3, and KCNQ5, as they show the strongest mutational clustering. Firstly, genes from each of 10 major cancer-associated pathways were defined based on previously published analysis of the TCGA (Sanchez-Vega et al., 2018). (**Table S4**). Given KCNQ genes are altered in expression and copy number within GI cancer, we infer that patients with high and low expression of each KCNQ gene may show differences in the activation of cancer-associated pathways they interact with.

We trained a machine learning model using Gradient Boosted Decision Forests (GBDF) to learn the difference between high and low expressors of each KCNQ gene when given only the genes in each pathway. A more accurate model (better discrimination between high and low expressors) indicates greater association of a particular pathway to the query gene.

We tested our method on the canonical cancer pathway genes CTNNB1, PTEN, KRAS, and NOTCH1 by generating 20 classification models, and calculating the average scores for each pathway. All classifications result in the correct pathway (WNT for CTNNB1, PI3K for PTEN, RTK-RAS for KRAS, and NOTCH for NOTCH1) being classified in the top 2 pathways, as expected (**Figure S13**), for this analysis we excluded the query gene from the pathway definition so as not to pre-bias the model.

Running the model with KCNQ1, KCNQ3 and KCNQ5 as the query genes across all GI cancers (**Figure 5 A**), we find that the WNT pathway scores the highest for

average Matthews Correlation Coefficient (MCC). We also find that accuracy scores are generally higher for KCNQ1 and KCNQ3, potentially indicating that they are better classifiers and thus relatively more associated with their respective pathways. This result reinforces previous evidence that associates the expression of KCNQ1 in colorectal cell lines with a direct physical interaction with beta catenin (Rapetti-Mauss et al., 2017), and potentially expands the interaction to involve other KCNQ genes.

One advantage of using GBDFs for classification is the ability to interpret models and analyse which features contribute most to a score, in our case indicating how much influence each gene in each pathway had on the MCC score. We generated average Shaply Additive Explanation (SHAP) scores for the WNT pathway for (Lundberg and Lee, 2017; Lundberg et al., 2018) for each model trained on high and low KCNQ expressors (**Table S5**). Genes with a high SHAP score (>0.1) are extremely influential in the classification and therefore most associated with the query gene. The highest ranked WNT genes that correlate with KCNQ1 expression are transcription factors (TCF7, TLE3, RNF43) (**Figure 5 B**), interestingly TCF7 has been shown to be responsive to potassium in a tumor setting in immune cell populations (Vodnala et al., 2019). KCNQ3 classifications were mostly driven by membrane bound members of the WNT pathway and ligands (FZDs and WNTs). Overall however, all regions of the WNT pathway scored highly for each gene, indicating involvement of the entire pathway, rather than a small subset of individual members.

To further explore the previous suggestions of an association of the KCNQ family with the WNT pathway, we generated clustered heatmaps of all WNT genes for the top and bottom 25 expressers of KCNQ1, KCNQ3, and KCNQ5 (**Figure 5 C**).

Heatmaps show patient-type clustering as predicted, and genes associate strongly into co-regulated groups that correlate with KCNQ expression, showing that KCNQ expression is strongly linked to activity of genes within the WNT pathway. We further performed GSEA analysis of the top 25 and bottom 25 overall expressers for each KCNQ and identify WNT signalling as a hallmark pathway significantly associated with KCNQ3 and KCNQ5 (FDR q value <= 0.1)(**Figure S14**).

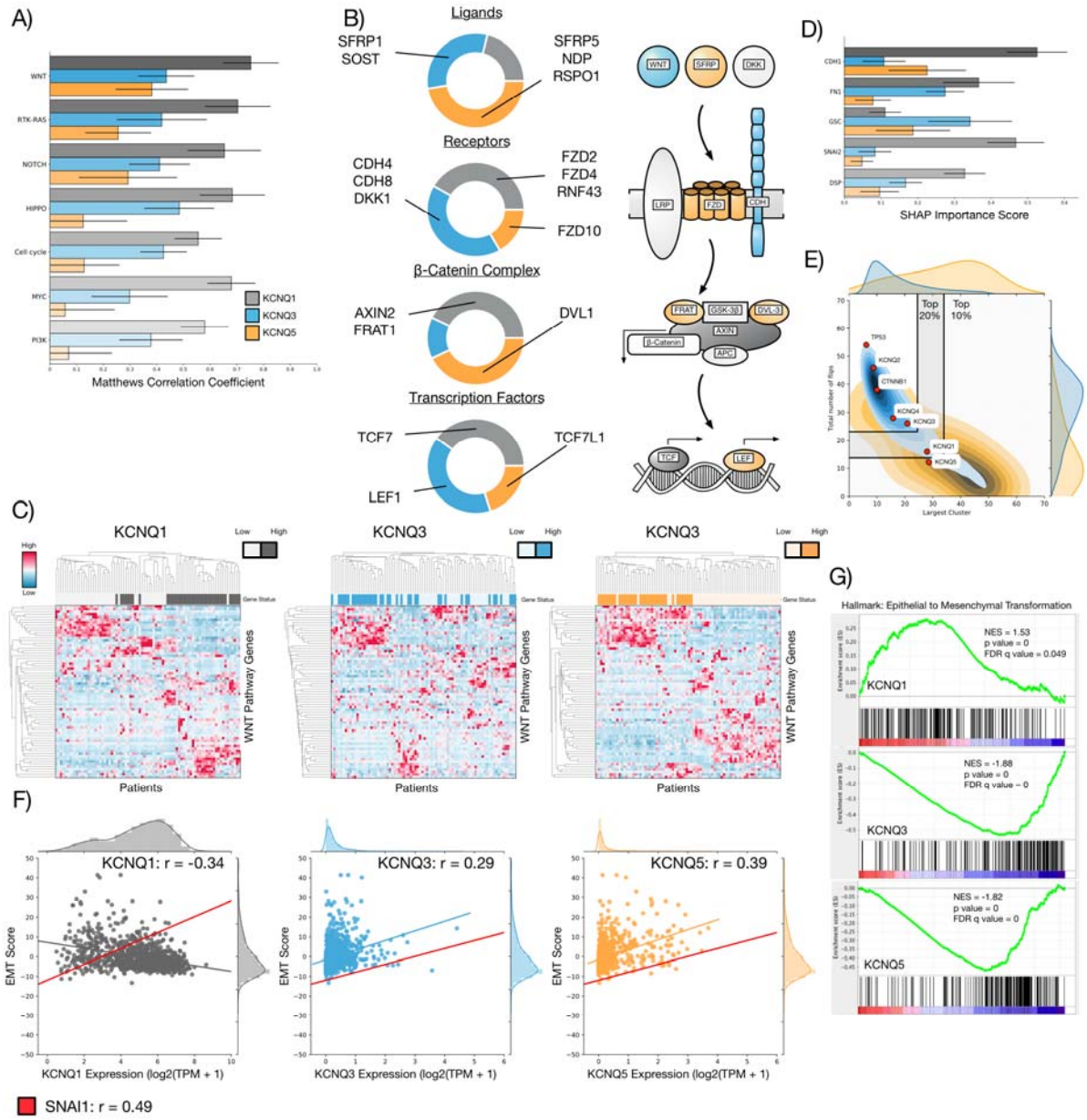
Because of the high association of KCNQ with WNT through our analysis, and previous work linking KCNQ1 to beta-catenin (Rapetti-Mauss et al., 2017), we chose to study the link between KCNQ genes and epithelial-to-mesenchymal transition (EMT). We trained a further GBDF on a set of genes previously implicated in RNAseq analysis of EMT (Gibbons and Creighton, 2018). Feature extraction of the models shows a strong correlation between KCNQ1 and E-cadherin and SNAI2 (**Figure 5 D**), and of KCNQ3 with fibronectin (FN1). Additionally, we identify KCNQ1 and KCNQ5 as being within the top ~20% and ~10% respectively of all genes for correlation with an EMT signature in the bulk RNAseq data (**Figure 5 E**).

To further unpick the relationship between KCNQ1/3/5, and EMT we looked at the association of a previously defined EMT-score generated from RNAseq (Creighton and Gibbons, 2013) with the expression of members of the KCNQ family. The EMT score is significantly correlated with the RNA expression levels for KCNQ1, KCNQ3, and KCNQ5 (**Figure 5 F**). KCNQ1 shows a negative correlation with EMT score, indicating that higher KCNQ1 expression is associated with a more epithelial phenotype (pearson rho = -0.34, p <0.0001). Conversely, KCNQ3 and KCNQ5 show

a positive correlation, indicating that high expression is associated with a more mesenchymal phenotype (pearson rho KCNQ3 = 0.29, KCNQ5 = 0.39, $p < 0.0001$), for comparison, the pearson rho for the classical EMT transcription factor SNAI1 is 0.49 ($p < 0.0001$). This reinforces the previous mutational data implicating KCNQ1 with a tumor suppressive role, and KCNQ3/5 with an oncogenic-like role. GSEA analysis finds a high correlation between EMT signatures in KCNQ1 and KCNQ3/KCNQ5 high and low expressors (FDR q value <= 0.05) (**Figure 5 G**).

As a validation of the RNAseq data we find significant disruption in the expression of WNT proteins from analysis of The Cancer Protein Atlas (TCPA) data (Li et al., 2017) when protein expression data is matched with TCGA RNAseq for the same patients (**Figure S15**), in particular we find levels of beta-catenin and alpha-catenin are commonly significantly altered in patients that are high vs low for KCNQ RNA (corrected $p <= 0.1$). We also find a statistically significant difference in the expression of EMT proteins between high and low RNA expressors of KCNQ genes (**Figure S16**), in particular, E-cadherin, claudin-7, are generally higher in high expressors of KCNQ1 and low expressors of KCNQ3/5, and fibronectin is generally found lower in high KCNQ1 patients and low KCNQ3/5 patients. Interestingly we also find a significant difference in both RNA and protein levels of BRCA2 for expressors of KCNQ3, potentially indicating a difference in DNA damage response.

Additional validation was performed by analysis of oesophageal adenocarcinoma organoid systems (Li et al., 2018). Whilst RNA levels of KCNQ3 and 5 in each system are extremely low, KCNQ1 is expressed at different levels in the different organoid systems. We find a strong correlation between the expression level of KCNQ1 and activation of the WNT pathway, similar to that seen in the patient RNAseq data (**Figure S17**).


A)
WNT
RTK-RAS
NOTCH
HIPPO
Cell cycle
MYC
PI3K
KCNQ1
KCNQ3
KCNQ5
Matthews Correlation Coefficient
B)
Ligands
SFRP1
SOST
SFRP5
NDP
RSPO1
Receptors
CDH4
CDH8
DKK1
FZD2
FZD4
RNF43
FZD10
β-Catenin Complex
AXIN2
FRAT1
DVL1
Transcription Factors
TCF7
TCF7L1
LEF1
WNT
SFRP
DKK
LRP
FZD
CDH
FRAT
GSK-3β
DVL-3
AXIN
β-Catenin
APC
TCF
LEF
D)
CDH1
FN1
GSC
SNAI2
DSP
SHAP Importance Score
E)
Top 20%
Top 10%
TP53
KCNQ2
CTNNB1
KCNQ4
KCNQ3
KCNQ1
KCNQ5
Total number of flips
Largest Cluster
C)
KCNQ1
KCNQ3
KCNQ3
Low
High
Gene Status
WNT Pathway Genes
Patients
G)
Hallmark: Epithelial to Mesenchymal Transformation
NES = 1.53
p value = 0
FDR q value = 0.049
KCNQ1
NES = -1.88
p value = 0
FDR q value = 0
KCNQ3
NES = -1.82
p value = 0
FDR q value = 0
KCNQ5
F)
KCNQ1: r = -0.34
KCNQ3: r = 0.29
KCNQ5: r = 0.39
EMT Score
KCNQ1 Expression (log2(TPM + 1)
KCNQ3 Expression (log2(TPM + 1)
KCNQ5 Expression (log2(TPM + 1)
SNAI1: r = 0.49

KCNQ gene expression correlates strongly with clinical prognosis

Kaplan Meier analysis of overall survival (**Figure 6 A**) and disease-free survival (**Figure S18**) for TCGA datasets shows that many members of the KCNQ family have a significant correlation with prognosis and recurrence. Across all GI cancers KCNQ1 expression positively correlates with survival ($p<0.0001$), confirming previous work suggesting that it has tumor suppressive properties. When broken down into constitutive cancers however, we find statistical significance for KCNQ1 only in the STAD dataset ($p<0.001$) (**Figure S19**), potentially indicating a stomach specificity. Though previous work has shown tumor suppressive properties for KCNQ1 in colorectal tissue (Rapetti-Mauss et al., 2017), we do not find a statistically significant correlation (**Figure S19**). We additionally find a statistically significant correlation with poorer overall survival and the high expression of KCNQ3, and KCNQ5 over all GI patients, indicating an oncogene-like effect associated with their expression. Tissue breakdowns suggest that KCNQ3 expression is negatively correlated with overall survival in all subtypes except ESCC, where there is no correlation, though statistical significance ($p < 0.05$) is not reached in all cases (**Figure S19**). We also studied the impact of KCNQ3 on survival in EAC in the combined TCGA-EAC/OCCAMS dataset, and find a statistically significant correlation with survival ($p=0.017$).

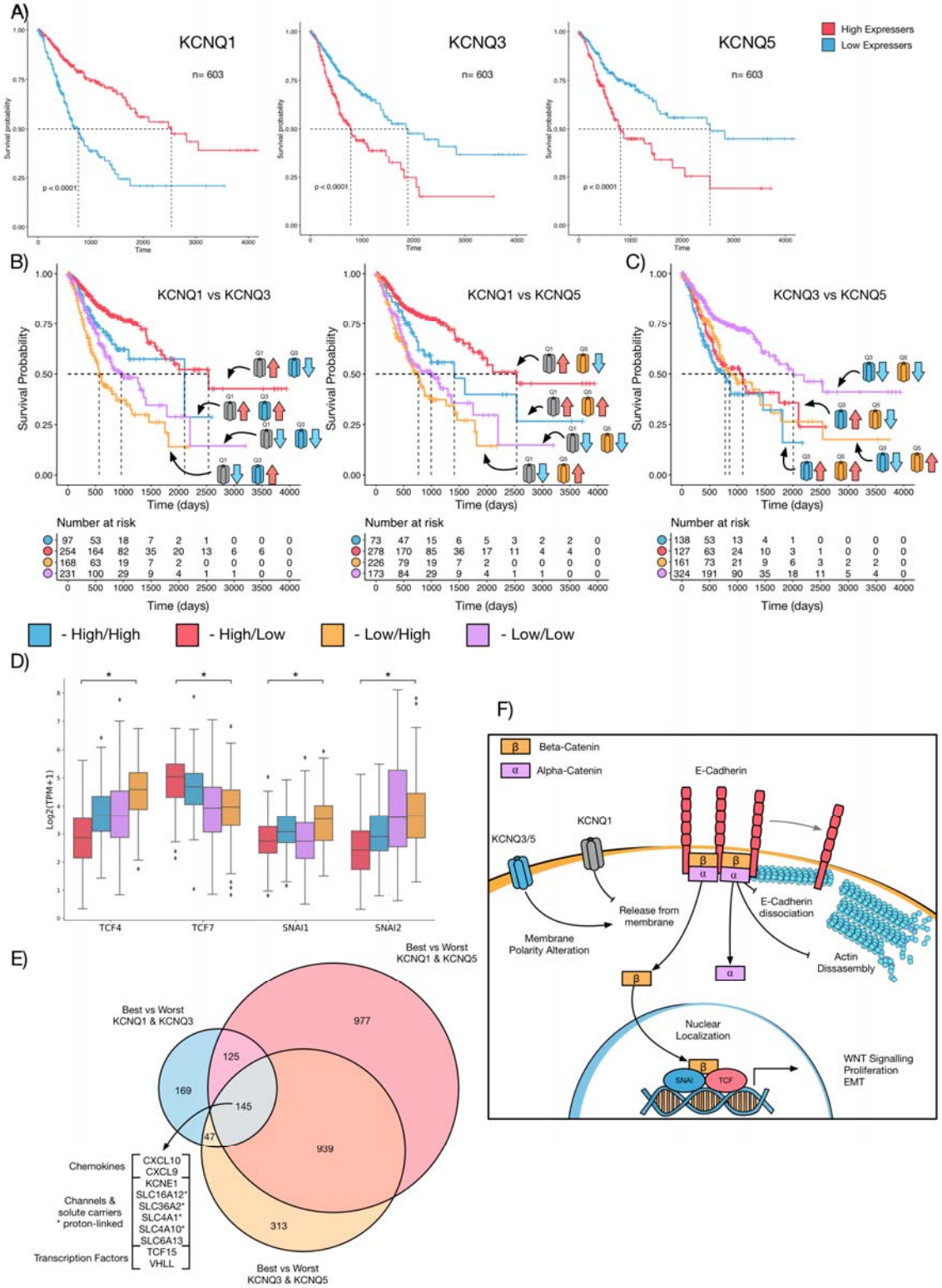
To study the inter-gene relationship between KCNQ1, KCNQ3, KCNQ5, and survival we separated patients into groups that were high and low for every pair of combinations of the three genes. Comparing KCNQ1 and KCNQ3 (**Figure 6 B**) we find the patients high for KCNQ1, and low for KCNQ3 have the best prognosis. The

worst prognosis is found for patients with low KCNQ1 and high KCNQ3, whilst patients high or low for both have an intermediate survival. The graded change in survival for these 4 sets of patients indicates that KCNQ1 and KCNQ3 may be independent of each other, as alterations in each individual gene combines with the effect of the other one survival. The same is also true when KCNQ1 and KCNQ5 are compared. The same is not true for comparison of patients with differing expression of KCNQ3 and KCNQ5 (**Figure 6 C**). Patients who are low expressers of both KCNQ3 and KCNQ5 have the best prognosis, whilst those with an increased expression of either, or both genes have a similarly worse prognosis. This indicates that KCNQ3 and KCNQ5 are mutually redundant, in that an alteration in a single one of them is enough to cause a significant prognostic change, and alteration of the second has no further impact on the phenotype. Taken together, this provides evidence that KCNQ1 can be treated as independently influencing the cell from KCNQ3 and KCNQ5.

We performed differential expression analysis between the best and worse prognosis patients in each case presented in Figure 6C for oesophageal adenocarcinoma. Classifying genes as differently expressed with a corrected $p > 0.1$, and log2 fold change > ±1.5 (**Figure 6 D**). We find that some commonly deregulated genes are chemokines, proton-coupled membrane transporters, and some transcription factors, including VHLL (**Table S6**). Additionally, we find a clear alteration in the expression levels of WNT transcription factors TCF4, TCF7, SNAI1, and SNAI2, shown for KCNQ1/KCNQ3 best and worst prognoses in (**Figure 6 E)**.

Finally we combine all of our findings into a single descriptive model. We generated a discrete, executable systems biology model of KCNQ activity in GI cancers using the BioModelAnalyzer. (**Figure S20**). The model describes the effects of mutations and expression changes in the KCNQ family based on the previously presented data, and predicts the impact of alteration on 5 cellular phenotypes, Proliferation, Cell Cycle Arrest, Apoptosis, Cell-Cell Adhesion, and EMT.

Overall our results point to a model where WNT signalling is mediated negatively by KCNQ1, and positively by KCNQ3/KCNQ5 (**Figure 6 F**).


A)
KCNQ1
n= 603
KCNQ3
n= 603
KCNQ5
n= 603
High Expressers
Low Expressers
p < 0.0001
Survival probability
Time
B)
KCNQ1 vs KCNQ3
KCNQ1 vs KCNQ5
C)
KCNQ3 vs KCNQ5
Survival Probability
Time (days)
Number at risk
- High/High
- High/Low
- Low/High
- Low/Low
D)
Log2(TPM+1)
TCF4
TCF7
SNAI1
SNAI2
E)
Best vs Worst KCNQ1 & KCNQ5
Best vs Worst KCNQ1 & KCNQ3
Best vs Worst KCNQ3 & KCNQ5
977
125
169
145
47
939
313
Chemokines
CXCL10
CXCL9
Channels & solute carriers
* proton-linked
KCNE1
SLC16A12*
SLC36A2*
SLC4A1*
SLC4A10*
SLC6A13
Transcription Factors
TCF15
VHLL
F)
Beta-Catenin
Alpha-Catenin
E-Cadherin
KCNQ1
KCNQ3/5
Release from membrane
Membrane Polarity Alteration
E-Cadherin dissociation
Actin Dissasembly
Nuclear Localization
SNAI
TCF
WNT Signalling
Proliferation
EMT

584

585

586

## DISCUSSION

There is emerging evidence that ion channels play a role in many, and potentially all cancers. Therapeutics against voltage gated potassium channels have been shown in vivo to improve prognosis for glioblastoma and breast cancer (Arcangeli and Becchetti, 2017; Pointer et al., 2017; Wang et al., 2015), and there are studies implicating specific sodium (Brisson et al., 2011; Roger et al., 2015) and calcium channels (Phan et al., 2017; Qu et al., 2014) in cancers. We present strong evidence that the KCNQ family of genes also play a significant and functional role in human gastrointestinal cancers. Through integration of data at the patient, cell, and protein structural levels we show that KCNQ genes and protein products contribute to cancer phenotype, and provide a selective advantage to cancers. We show that a large number of patients with gastrointestinal cancers show genetic alterations in a member of the KCNQ family, and that the expression levels of these genes correlate strongly with survival. We find that mutations in the KCNQ family of genes are clustered within KCNQ1, KCNQ3, and KCNQ5, and that there is evidence of mutational selection. Furthermore, we prove that many of these mutations have functional effects on the protein products, through alteration of the gating properties of the channel, or through occlusion of the pore region. We go on to analyse RNAseq and RPPA level data, using novel machine learning techniques, to show that KCNQ gene expression highly correlates with the WNT pathway, and with EMT, indicating a potential role for the gene family in cancer signalling and transformation. Finally, we generate a systems biology model to sum our findings into a descriptive model, and show that patients with different expression levels of KCNQ genes have varying prognosis.

By studying data from varying levels simultaneously (patient level, cellular level, and protein level) we find consistently that KCNQ1 exhibits properties of a tumor suppressor – it is often deleted or lost in cancer, mutations are generally inactivating, and it is negatively correlated with WNT signalling and EMT. Opposingly, we find that other KCNQ genes (mainly KCNQ3 and KCNQ5) show hallmarks of onocogenic properties, they are often amplified in cancers, and their mutations are synonymous with gain-of-function mutations. Additionally, KCNQ3 and KCNQ5 are positively correlated with WNT signalling and the EMT pathway, thus genes within the same family, with very similar chemical activity have opposing influences on cancer phenotype. The correlation with survival is of particular interest for potential patient stratification, and may additionally represent a clinical window for reduction of certain cancer phenotypes, especially given the recent developments of highly specific and potent KCNQ subtype specific drugs (Manville and Abbott, 2019). Caution must be taken when considering therapeutic applications of KCNQ involvement in cancer, due to the extreme importance of KCNQ1 in cardiac activity, nevertheless compounds specific to KCNQ3 or KCNQ5 may have a therapeutic window, as is thought to be the base with hERG inhibitors (Fukushiro-Lopes et al., 2018).

That every level of data (genetic alteration/clinical, mutational/structural, gene expression/pathway) consistently agrees that KCNQ1 appears tumor suppressive, whilst KCNQ3 and 5 have oncogenic properties is a strong indicator that the results are clinically relevant.

Whilst we confirm previous work that KCNQ1 may be involved in WNT signalling (Rapetti-Mauss et al., 2017), and expand to show that KCNQ3 and KCNQ5 play an opposing role, further work is necessary to elucidate exactly how ion channels can drive these pathways, through either membrane polarity, physical protein-protein interactions, or another currently undiscovered mechanism.

A further result of this finding is the implication that patients with congenital mutations in KCNQ genes may be more at risk or protected against particular GI cancers, this is especially poignant in the light of recent results suggesting that KCNQ channels may be acted upon by particular dietary vitamins such as vitamin B6 (Reid et al., 2015).

In summary, we identify a strong role for KCNQ1, KCNQ3, and KCNQ5 in the progression of human gastrointestinal cancers. We find that KCNQ1 is tumor suppressive, whereas KCNQ2, KCNQ3, and KCNQ5 have oncogenic properties. We confirm our findings across multiple levels of data, and evidence is consistent at the patient, cellular, and protein/gene levels. We also find a role for KCNQ genes in the WNT pathway and EMT, expanding upon previous work that found an interaction between KCNQ1 and beta-catenin. This work highlights the emerging functional role of the KCNQ family in human gastrointestinal cancers.

## ACKNOWLEDGEMENTS

We acknowledge the help and support of Benjamin Marshal on the Bayesian Inference analysis. We are grateful to the OCCAMS consortium for providing the data for the oesophageal analysis linked with clinical outcome data, and Dr Xiaodun Li for advice with the organoid data. We acknowledge the services of Cancer Research UK Cambridge Institute Genomics Core for performing RNAseq. This work has been supported by the Royal Society (URF to BH grant no. UF130039, studentship to VK), Medical Research Council (Grant-in-Aid to the MRC Cancer unit grant number MC_UU_12022/9 and NIRG to BH grant number MR/S000216/1), and the Harrison Watson Fund at Clare College, Cambridge (M.W.J.H.).

## AUTHOR CONTRIBUTIONS

DS and BH conceived the study, DS curated the data and performed all analysis. MH wrote code for DN/DS style mutation frequency, ER and CK performed cell culture and prepared RNA for sequencing, GD supported dataset processing and analysis. RCF provided the oesophageal data and expertise on human GI cancer. DS, BH, and RCF cowrote the manuscript. All authors were responsible for editing of the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## FIGURE LEGENDS

Figure 1:

KCNQ genes are highly altered in GI cancer.

A) Mutations in any member of the KCNQ/KCNE family across all cancers in The Cancer Genome Atlas (TCGA)
B) Percentage of patients with an alteration in a KCNQ gene for 5 cohorts: Oesophageal Adenocarcinoma (EAC – OCCAMS and TCGA), Oesophageal Squamous Cell Carcinoma (ESCC), Stomach Adenocarcinoma (STAD), and Colorectal Adenocarcinoma (COADREAD).
C) Percentage of patients across all cohorts with an alteration in each KCNQ/KCNE gene.
D) Oncoprint showing alterations in KCNQ/KCNE genes per patient.

Figure 2:

Mutations in KCNQ genes cluster in cancer

A) Mutational Frequency for all KCNQ genes after sequence alignment
B) Mutational clustering for KCNQ1, line represents Observed vs Expected mutation ratio

C) Mutational clustering for KCNQ3, line represents Observed vs Expected mutation ratio

D) Mutational clustering for KCNQ5, line represents Observed vs Expected mutation ratio

Figure 3:

Mutations in KCNQ genes cluster in space

A) Schematic of the assembly of KCNQ/KCNE genes.

B) Bayesian inference analysis of the RNA expression levels of KCNQ/KCNE genes in the colorectal (COADREAD) dataset.

C) Schematic of the KCNQ canonical structure.

D) Render of the KCNQ1 structure.

E) Render of the KCNQ3 structure with mutational frequency overlaid, darker red regions represent a higher mutational frequency.

F) Spatial clustering plot for KCNQ1 (top) and KCNQ3 (bottom), arcs represent residues that are within 12Å of each other, size of each circle represents the number of mutations occurring at that location in the structure. Labelled are mutational clusters identified in figure 2 B.

G) Structure of KCNQ1 with mutational clusters highlighted in red (cluster 1.1), and blue (cluster 1.2).

H) Structure of KCNQ3 with mutational clusters highlighted in pink (cluster 3.1), and orange (cluster 3.2).

Figure 4:

Mutations in KCNQ1 and KCNQ3 in GI cancers impact protein function.

A) Render of the pore region of KCNQ1. Residues mutated in cluster 1.2 are rendered in blue. Mutations that constrict the pore region when modelled are starred.
B) HOLE analysis of the pore region of KCNQ1 wt (black) WT mutations in cluster 1.2 (coloured according to Figure 4, A).
C) Mutational frequency from the COSMIC GI cancer dataset for the S4 region of the KCNQ family. Residues involved in gating (R1, R2, Q3, R4, R5, R6) are highlighted (top).
D) Helical tilt angle for alanine scanning mutagenesis for the KCNQ1 S4 helix.
E) Helical tilt angle for mutational analysis of KCNQ3 S4 helix.
F) Residues involved in stabilizing the active conformation of KCNQ2 through hydrogen bonding.
G) Hydrogen bonding networks generated from molecular dynamics trajectories for KCNQ3 (left) and KCNQ5 (right). Percentages represent total bond occupancy over 100 ns equilibrium molecular dynamics.

Figure 5:

KCNQ1, KCNQ3, and KCNQ5 gene expression correlates with WNT signalling and EMT

A) Gradient Boosted Decision Forest (GBDF) based pathway analysis results for comparing patients that are high and low for KCNQ1 (grey), KCNQ3 (blue) and KCNQ5 (orange). Pathways are ranked by average Matthews Correlation Coefficient.

B) Feature extraction schematic for WNT pathway classification. Genes with a high Shapely Additive Explanation (SHAP score > 0.1) are counted for each model

C) WNT pathway RNA expression clustering between patients high (dark) and low (light) for KCNQ1 (left), KCNQ3 (middle), and KCNQ5 (right).

D) Feature extraction for GBDF model of the EMT pathway when classifying patients high and low for KCNQ1 (grey), KCNQ3 (blue), and KCNQ5 (orange).

E) Ranking metric of KCNQ genes for the EMT pathway. KCNQ genes are ranked against all genes for ability to cluster EMT pathway genes between high and low expressing patients.

F) EMT score for KCNQ1 (left), KCNQ3 (middle), KCNQ5 (right). Red line represents EMT score for EMT associated transcription factor SNAI1 ($r = 0.49$, $p < 0.01$).

G) GSEA hallmark: EMT analysis for high vs low expressors of KCNQ1 (top), KCNQ3 (middle), and KCNQ5 (bottom).

Figure 6:

KCNQ expression correlates with clinical prognosis:

A) Survival analysis for KCNQ1 (left), KCNQ3 (middle), and KCNQ5 (right) high vs low expressers.

B) Survival analysis for patients with different co-expressions of KCNQ1 vs KCNQ3 (left), and KCNQ1 vs KCNQ5 (right). Schematics indicate up and downregulation of KCNQ1 (grey), KCNQ3 (blue), and KCNQ5 (orange).

C) Survival analysis for patients with different co-expressions of KCNQ3 vs KCNQ5. Schematics indicate up and downregulation of KCNQ3 (blue), and KCNQ5 (orange).

D) Venn diagram of overlapping genes from differential gene expression analysis of patients in 6 cohorts divided into 3 sets. Shown are differentially expressed (corrected $p < 0.05$, $\log_2$ fold change > 1.5) genes for comparison of KCNQ1 high/KCNQ3 low patients vs KCNQ1 low/KCNQ3 high patients (blue), KCNQ1 high/KCNQ5 low patients vs KCNQ1 high/KCNQ5 low patients (pink), and KCNQ3 low/KCNQ5 low patients vs KCNQ3 high/KCNQ5 high patients (blue). Grey region represents genes differentially expressed in all cases.

E) RNA expression of WNT associated transcription factors between patients with high KCNQ1/low KCNQ3 (red), high KCNQ1/high KCNQ3 (blue), low KCNQ1/low KCNQ3 (pink), and low KCNQ1/high KCNQ3 (orange).

F) Schematic of systems biology model representing influence of KCNQ1 and KCNQ3/5 on cellular phenotype.

# TABLES

| Gene | Mutation | No. Obs. | Specific Effect | Functional Effect | Clinical Impact | References |
|---|---|---|---|---|---|---|
| KCNQ1 | K422fs | 6 | Does not traffic | LOF | Long QT Syndrome | Harmer et al 2014 |
| KCNQ1 | F339L | 3 | Reduced Currents | LOF | Romano-Ward Syndrome | Thomas et al (2005) |
| KCNQ1 | K422R | 2 | Reduced Currents | LOF | Long QT Syndrome | Veerman et al (2013) |
| KCNQ1 | P343S | 2 | Reduced Currents | LOF | Romano-Ward Syndrome | Zehelein et al (2004) |
| KCNQ1 | R293H | 2 | Loss of Predicted salt bridge to KCNE1 | - | - | Xu et al (2013) |
| KCNQ1 | R192C | 2 | Loss of PIP2 binding site | LOF | Long QT Syndrome | Zaydman et al (2013), Eckey et al (2014) |
| KCNQ1 | R192H | 2 | Loss of PIP2 binding site | LOF | Long QT Syndrome | Zaydman et al (2013), Eckey et al (2014) |
| KCNQ2 | R210H | 4 | Increased Activation threshold | LOF | Benign Familial Neonatal Convulsions | Neverisky et al (2017) |
| KCNQ2 | A294V | 4 | Reduced currents | LOF | Epilepsy of Infancy with Migrating Focal Seizures | Duan et al (2018) |
| KCNQ2 | E119D | 2 | Negative Shift in Activation Threshold | LOF | Benign Familial Neonatal Convulsions | Wuttke et al (2008) |
| KCNQ2 | G301S | 2 | Loss of Retigabine Binding Site | - | - | Kalappa et al (2015) |
| KCNQ2 | R581Q | 2 | - | - | Benign Familial Neonatal Convulsions | Singh et al (2003) |
| KCNQ3 | K533fs | 9 | Loss of PIP2 binding site | - | - | Choveau et al (2018) |
| KCNQ3 | R230C | 5 | Increased current amplitude, increased activation | GOF | Neurodevelopmental Disability | Sands et al (2019) |
| KCNQ3 | R230H | 5 | Increased current amplitude, increased activation | GOF | Neurodevelopmental Disability | Sands et al (2019) |
| KCNQ3 | K366N | 4 | Loss of PIP2 binding site | - | - | Choveau et al (2018) |
| KCNQ3 | R227Q | 3 | Reduced Activation Threshold | GOF | Neurodevelopmental Disability | Sands et al (2019) |

Table 1: KCNQ mutations in GI Cosmic patients. Shown are mutations with known functional or clinical consequences

| Gene | Mutation | Equivalent Name | No. Obs. | Specific Effect | Functional Effect | Sidekick Result | Reference |
|---|---|---|---|---|---|---|---|
| KCNQ1 | R228A | R1A | - | Impaired activation | LOF | No change | Panaghie et al (2007) |
| KCNQ1 | R231A | R2A | - | Constitutively open | GOF | Decrease tilt | Panaghie et al (2007) |
| KCNQ1 | R237A | R4A | - | Accumulates in the open state | GOF | Decrease tilt | Panaghie et al (2007) |
| KCNQ1 | D242A | D6A | - | More difficult to activate | LOF | Increase tilt | Panaghie et al (2007) |
| KCNQ1 | R243A | R6A | - | Impaired activation | LOF | Increase tilt | Panaghie et al (2007) |
| KCNQ2 | R201H | R2H | - | Reduced Activation Threshold | GOF | Decrease tilt | Miceli et al (2015) |
| KCNQ2 | R201C | R2C | - | Constitutively Active | GOF | Decrease tilt | Miceli et al (2015) |
| KCNQ2 | I205V | - | - | Loss of function | LOF | No change | Niday et al (2017) |
| KCNQ2 | R207W | R4W | - | Increased Activation Threshold | LOF | Decrease tilt | Dedek et al (2001) |
| KCNQ2 | R210H | R5H | 4 | - | - | No change | Rikee database[1] |
| KCNQ2 | R213W | R6W | - | Increased Activation Threshold | LOF | Increase tilt | Miceli et al (2013) |
| KCNQ2 | R198Q | R1Q | - | Reduced Activation Threshold | GOF | No change | Miceli et al (2008) |
| KCNQ2 | R201Q | R2Q | - | Constitutively Active | GOF | Decrease tilt | Miceli et al (2008) |
| KCNQ2 | R207Q | R4Q | 1 | Increased Activation Threshold | LOF | No change | Miceli et al (2008) |
| KCNQ2 | R210Q | R5Q | - | Increased Activation Threshold | LOF | No change | Miceli et al (2008) |
| KCNQ2 | R213Q | R6Q | - | Increased Activation Threshold | LOF | No change | Miceli et al (2008) |
| KCNQ2 | R214Q | R7Q | - | Increased Activation Threshold | LOF | No change | Miceli et al (2008) |
| KCNQ3 | R227Q | R1Q | 3 | Reduced Activation Threshold | GOF | No change | Sands et al (2019) |
| KCNQ3 | R230H | R2H | 4 | Reduced Activation Threshold | GOF | Decrease tilt | Sands et al (2019) |
| KCNQ3 | R230C | R2C | 1 | Reduced Activation Threshold | GOF | Decrease tilt | Sands et al (2019) |
| KCNQ3 | R236C | R4C | 2 | - | - | Decrease tilt | - |
| KCNQ3 | R236H | R4H | 4 | - | - | Decrease tilt | - |
| KCNQ3 | R239W | R5W | 2 | - | - | No change | - |
| KCNQ3 | R242W | R6W | 1 | - | - | No change | - |
| KCNQ4 | R207H | R2H | 2 | - | - | Decrease tilt | - |
| KCNQ4 | R207C | R2C | 1 | - | - | Decrease tilt | - |
| KCNQ4 | R213C | R4C | 1 | - | - | Decrease tilt | - |

[1] Rikee database at https://rikee.org

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| KCNQ4 | R216H | R5H | - | Reduced Activation Threshold | GOF | Decrease tilt | Panaghie et al (2007) |
| KCNQ5 | R244H | R5H | 3 | - | - | No change | - |
| KCNQ5 | R244C | R5C | 6 | - | - | No change | - |

Table 2: Results of Molecular Dynamics Simulations on single S4 helices from the KCNQ family.

# METHODS

## TCGA Data

TCGA level 3 data was downloaded using Firebrowse (RNAseq), cBioportal (CNAs, mutation and clinical data)(Gao et al., 2013), or TCGA wanderer (Methylation data)(Díez-Villanueva et al., 2015).

## COSMIC Data

COSMIC data(Forbes et al., 2017; Tate et al., 2019) was downloaded from cancer.sanger.ac.uk. We subset mutations into those only found in gastrointestinal tissue, defined as those where the primary site is in one of the following categories: "large_intestine", "small_intestine", "gastrointestinal_tract_(site_indeterminate)", "oesophagus", "stomach".

## Oncoprints

Oncoprints were produced using the R library Oncoprint. Copy number alterations were determined as follows – relative copy number for each gene was defined as:

$$relative\ copy\ number = \ log_2(\frac{Total\ Copy\ Number\ of\ Gene}{Total\ ploidy\ of\ sample})$$

Genes were defined as deleted if total copy number == 0 OR relative copy number < -1, genes were defined as amplified if relative copy number > 1.

## Mutual Exclusivity/Cooperativity

Mutual exclusivity and cooperativity for KCNQ genes was calculated using the DISCOVER library in R(Canisius et al., 2016).

## Sequence alignments

Sequence alignment was performed using the MUSCLE algorithm(Edgar, 2004), implemented in python through Biopython(Cock et al., 2009), using default settings.

FASTA files for KCNQ sequences were obtained from uniprot using codes P51787 (KCNQ1), O43526 (KCNQ2), O43525 (KCNQ3), P56696 (KCNQ4), and Q9NR82 (KCNQ5).

## Expected vs Observed mutational distribution

For calculating the expected vs observed mutational distribution, exon data was downloaded from the Ensembl Biomart (ensembl.org/biomart). Ensembl 96, hg38.p12 was selected, and data downloaded for chromosomes 1-22, X, and Y. We used bedtools(Quinlan and Hall, 2010) to sort data and overlapping exons were merged. To sort data we used the following command:

```
tail -n +2 human_exon_bed_1-Y.txt | cut -f 1,2,3 | bedtools sort -i stdin > human_exon_bed_file_1-Y_sorted.bed
```

Merging was performed using:

```
bedtools merge -i human_exon_bed_file_1-Y_sorted.bed > human_exon_bed_file_1-Y_merged.bed
```

The R library Deconstructsigs(Rosenthal et al., 2016) was used to generate the mutational signature for all COSMIC mutations in KCNQ genes within GI tissue for these exons. The mutational spectrum was normalised using the mutational signature to generate the expected relative mutation rate for each possible missense mutation. This was then multiplied by the total number of mutations in each gene to get the expected distribution of events along the gene. The observed and expected mutational frequency ratio was average over a sliding window of 50 bases.

## 1D Mutational Clustering

Mutational clustering was calculated using the NMC clustering method from the R library IPAC(Ye et al., 2010) applied to sequence alone. All mutations to each KCNQ gene in the COSMIC database for GI cancers were considered. The top 5 mutational clusters ranked by adjusted P-value were plotted.

## Fathmm/CScape score calculation

To generate fathmm(Shihab et al., 2013) and CScape(Rogers et al., 2017) scores for each mutation we downloaded the fathmm and cscape databases (fathmm.biocompute.org.uk, cscape.biocompute.org.uk). We constructed a query for each database of every possible mutation in each KCNQ gene of interest using the python scripts available with each database and pyensembl using GRCH37 and the canonical transcript.

Bayesian inference analysis

RSEM normalised RNAseq data was used to calculate the association of KCNQ/E genes with each other. The covariance matrix for KCNQ genes was calculated and converted to correlation matrix. Permutation tests (N = 5000) were performed to calculate a p-value testing the null hypothesis that the corresponding pair of variables has 0 correlation. P-values of less than 0.05 were deemed a significant co-association.

Homology modelling

Homology modelling was performed using the template structure 5VMS from the protein data bank(Sun and MacKinnon, 2017). Modeller was used to generate 5 models, and the model with the highest GA341 (nativity) scores, and lowest DOPE (total energy of the structure) scores was selected. Alignments used to generate the models are available in the SI. Single point mutations were induced in the models using the mutate_model protocol in the modeller tool as described in Feyfant et al (Feyfant et al., 2007), and using the mutate_model.py script available on the modeller website (https://salilab.org/modeller/wiki/Mutate%20model).

Pore analysis

Pore analysis was performed using the algorithm HOLE (Smart et al., 1996). Pore profile was visualised using Visual Molecular Dynamics (VMD) (Humphrey et al., 1996).

Molecular dynamics simulations

Molecular dynamics was performed using Gromacs version 2018.1(Abraham et al., 2015).

For simulations of homology models in AT we used the charmm36 forcefield(Lee et al., 2014). In each case the protein was placed in a 15 x 15 x 15nm box with roughly 650 DPPC lipid molecules. Setup was performed in the same manner as systems in the memprotMD pipeline (Stansfeld et al., 2015). The system was converted to MARTINI coarse-grained structures (CG-MD) with an elastic network in the martiniv2 forcefield(Monticelli et al., 2008) and self-assembled by running a 1000ns molecular dynamics simulation at 323k to allow the formation of the bilayer around the protein. The final frame of the CG-MD simulation was converted back to atomistic detail using the CG2AT method (Stansfeld and Sansom, 2011). The AT system was neutralised with counterions, and additional ions added up to a total NACL concentration of 0.05 mol/litre. The system was minimized using the steepest descents algorithm until maximum force Fmax of the system converged. Equilibration was performed using NVT followed by NPT ensembles for 100 ps each with the protein backbone restrained. We used the Verlet cutoff scheme with PME electrostatics, and treated the box as periodic in the X, Y, and Z planes. Simulations were run for 200ns of unrestrained molecular dynamics. Root mean square deviation (RMSD) was calculated for structures using the g_rmsdist command in GROMACS.

CG simulations of single helices were performed as described previously (Hall et al., 2014). Models of single helices were generated and converted to MARTINI coarse grained structures. Helices were then inserted into POPC bilayers and simulated for short (100ns) simulations for 100 repeats of each sequence.

H-bond analysis

H-bond occupancy for molecular dynamics simulations was calculated using the VMD HBonds plugin.

Gradient Boosted Decision Forests

Gradient boosted decision forests were implemented using Xgboost in the python package scikit-learn(Chen and Guestrin, 2016). RNAseq data from TCGA and OCCAMS was combined. For training the models, a target gene is selected, and the top and bottom 50 expressors selected. A model was then trained to classify the 100 patients into high and low expressors based only on genes from a query pathways. Pathways and genes were generated from Sanchez-Vega et al (Sanchez-Vega et al., 2018). We trained 20 separate models for each pathway, in each case a random

20% (10) patients were removed from the dataset and used for testing. The model was run using the command:

```
Model = XGBClassifier(max_depth = 3, n_estimators = 100, learning_rate = 0.1, colsample_bytree = 0.3)
```

The model was fitted using the logloss evaluation metric and stopped using the early stopping method 40 epochs prior to a loss of accuracy. The method was tested on known genes from established pathways (KRAS in the KRAS pathway, CTNNB1 in the WNT pathway, and P53 in the P53 pathway). In these cases the query gene was removed from the respective pathway in order to unbias the model towards them. Models were evaluated and interpreted by generating the SHAP score using the SHAP implementation in python (Lundberg and Lee, 2017).

Gene ranking

For ranking individual genes against all other genes in the dataset for correlation with a gene set, a target gene is selected, and the top and bottom 50 expressors from the RNAseq dataset were selected. The RNAseq data is then subset by genes in a geneset (for EMT this subset consists of: CDH1, DSP, TJP1, VIM, CDH2, FOXC2, SNAI1, SNAI2, TWIST1, GSC, FN1, ITBG6, MMP2, MMP3, MMP9, and SOX10 – as defined by Chae et al (Chae et al., 2018), and the optimally ordered linkage is calculated for all patients using the hierarchical clustering method from scipy. The command used to generate the linkage is:

```
leave_order = scipy.cluster.hierarchy.linkage(gene_expression_matrix, optimal_ordering = True)
```

Where gene_expression_matrix is a matrix of gene expression values for a pathway and the top and bottom 50 patients for expression of a query gene. The order of leaves in the tree is calculated and converted to binary format (where 0 is low expressors and high expressors). We infer that genes highly associated with a geneset query will cluster well, and so the binary string will contain contiguous strings of 1s and 0s. To calculate the “orderedness” of the binary string we calculate two metrics:

1) The longest continuous string of 1s or 0s in the binary string
2) The number of times a 1 is flipped to a 0 or vice versa

We calculate these two scores for every gene in the dataset for the query pathway. We plot the resultant values against one another to get the degree to which a set of query genes correlates with every gene in the dataset.

Gene Set Enrichment Analysis

Gene set enrichment was performed using the GSEA desktop application (Subramanian et al., 2005). We subset patients into the 25 highest and lowest expressors for each of KCNQ1, KCNQ3, and KCNQ5. GSEA was run for 1000 permutations using the geneperm permuation method for the hallmarks and oncogenic signatures gene sets (h.all.v7.0 and c6.all.v7.0 respectively) and ranked by Ttest.

EMT score

EMT score was calculated following the protocol defined by Chae et al (Chae et al., 2018). RNAseq expression data for each sample in the TCGA and OCCAMS dataset was z-score normalised and genes corresponding to epithelial and mesenchymal phenotypes were extracted. We classed CDH1, DSP, and TJP1 as epithelial markers, and VIM, CDH2, FOXC2, SNAI1, SNAI2, TWIST1, GSC, FN1, ITBG6, MMP2, MMP3, MMP9, and SOX10 as mesenchymal. The EMT score for each patient sample was calculated by subtracting the mean z-score of epithelial markers from the mean for mesenchymal markers. For comparison we calculated the correlation for SNAI2, a known driver of EMT. For correlation of EMT scores against gene expression, we calculated the spearmans correlation for the TPM expression of a query gene against the EMT score.

TCPA analysis

For analysis of data from The Cancer Protein Atlas (TCPA) (Li et al., 2013)we downloaded RPPA data from the TCPA portal (tcpaportal.org). We matched samples from the TCPA with their respected TCGA ID, and took only patients where both RNAseq and protein level data was available. Patients were then split into the highest and lowest 25 expressors of a query gene in the RNA expression data, and protein level data compared. High and low expression cohorts were compared using a students T test, and multiple test corrected using the Bonferroni method.

Survival analysis

Survival analysis was performed using the R library Survival. We used the clinical data associated with the TCGA and OCCAMS, which includes overall survival and recurrence free survival data. Survival times were converted to days, and survival curves generated on the top and bottom 25% expressors of each query gene.

Differential expression analysis

Differential expression analysis was performed using the R library Deseq2 (Love et al., 2014).

Systems biology

Systems biology models were constructed with the BioModelAnalyzer tool (Chuang et al., 2015). Nodes were assigned to genes or interest, and relationships inferred through data analysis and through studying the literature (see **Table S7** for interactions in the models)

## REFERENCES

Abbott, G.W. (2014). Biology of the KCNQ1 Potassium Channel. New J. Sci.

Abraham, M.J., Murtola, T., Schulz, R., Páll, S., Smith, J.C., Hess, B., and Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX *1–2*, 19–25.

Aggarwal, S.K., and MacKinnon, R. (1996). Contribution of the S4 Segment to Gating Charge in the Shaker K+ Channel. Neuron *16*, 1169–1177.

Allen, N.M., Mannion, M., Conroy, J., Lynch, S.A., Shahwan, A., Lynch, B., and King, M.D. (2014). The variable phenotypes of KCNQ-related epilepsy. Epilepsia.

Arcangeli, A., and Becchetti, A. (2017). hERG Channels: From Antitargets to Novel

Targets for Cancer Therapy. Clin. Cancer Res. *23*, 3–5.

Bass, A.J., Thorsson, V., Shmulevich, I., Reynolds, S.M., Miller, M., Bernard, B., Hinoue, T., Laird, P.W., Curtis, C., Shen, H., et al. (2014). Comprehensive molecular characterization of gastric adenocarcinoma. Nature *513*, 202–209.

Brisson, L., Gillet, L., Calaghan, S., Besson, P., Le Guennec, J.-Y., Roger, S., and Gore, J. (2011). Na(V)1.5 enhances breast cancer cell invasiveness by increasing NHE1-dependent H(+) efflux in caveolae. Oncogene *30*, 2070–2076.

Brown, D.A., and Passmore, G.M. (2009). Neural KCNQ (Kv7) channels. Br. J. Pharmacol.

Campbell, P.J. (2017). Cliques and Schisms of Cancer Genes. Cancer Cell *32*, 129–130.

Canisius, S., Martens, J.W.M.M., and Wessels, L.F.A.A. (2016). A novel independence test for somatic alterations in cancer shows that biology drives mutual exclusivity but chance explains most co-occurrence. Genome Biol. *17*, 261.

Chae, Y.K., Chang, S., Ko, T., Anker, J., Agte, S., Iams, W., Choi, W.M., Lee, K., and Cruz, M. (2018). Epithelial-mesenchymal transition (EMT) signature is inversely associated with T-cell infiltration in non-small cell lung cancer (NSCLC). Sci. Rep.

Chalmers, Z.R., Connelly, C.F., Fabrizio, D., Gay, L., Ali, S.M., Ennis, R., Schrock, A., Campbell, B., Shlien, A., Chmielecki, J., et al. (2017). Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome Med. *9*, 34.

Chen, T., and Guestrin, C. (2016). XGBoost. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD

'16, (New York, New York, USA: ACM Press), pp. 785–794.

Chuang, R., Hall, B.A., Benque, D., Cook, B., Ishtiaq, S., Piterman, N., Taylor, A., Vardi, M., Koschmieder, S., Gottgens, B., et al. (2015). Drug Target Optimization in Chronic Myeloid Leukemia Using Innovative Computational Platform. Sci. Rep. *5*, 8190.

Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics *25*, 1422–1423.

Creighton, C., and Gibbons, D. (2013). Pan-cancer survey of epithelial–mesenchymal transition markers across The Cancer Genome Atlas. Dev. Dyn.

DeMarco, M.L. (2012). Three-Dimensional Structure of Glycolipids in Biological Membranes. Biochemistry *51*, 5725–5732.

Díez-Villanueva, A., Mallona, I., and Peinado, M.A. (2015). Wanderer, an interactive viewer to explore DNA methylation and gene expression data in human cancer. Epigenetics and Chromatin.

Edgar, R.C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res.

Fan, H., Zhang, M., and Liu, W. (2018). Hypermethylated KCNQ1 acts as a tumor suppressor in hepatocellular carcinoma. Biochem. Biophys. Res. Commun. *503*, 3100–3107.

Feske, S., Wulff, H., and Skolnik, E.Y. (2015). Ion channels in innate and adaptive immunity. Annu. Rev. Immunol. *33*, 291–353.

Feyfant, E., Sali, A., and Fiser, A. (2007). Modeling mutations in protein structures. Protein Sci.

Forbes, S.A., Beare, D., Boutselakis, H., Bamford, S., Bindal, N., Tate, J., Cole, C.G., Ward, S., Dawson, E., Ponting, L., et al. (2017). COSMIC: Somatic cancer genetics at high-resolution. Nucleic Acids Res.

Frankell, A.M., Jammula, S., Li, X., Contino, G., Killcoyne, S., Abbas, S., Perner, J., Bower, L., Devonshire, G., Ococks, E., et al. (2019). The landscape of selection in 551 esophageal adenocarcinomas defines genomic biomarkers for the clinic. Nat. Genet. *51*, 506–516.

Fukushiro-Lopes, D.F., Hegel, A.D., Rao, V., Wyatt, D., Baker, A., Breuer, E.-K., Osipo, C., Zartman, J.J., Burnette, M., Kaja, S., et al. (2018). Preclinical study of a Kv11.1 potassium channel activator as antineoplastic approach for breast cancer. Oncotarget *9*, 3321–3337.

Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci. Signal.

Gibbons, D.L., and Creighton, C.J. (2018). Pan-cancer survey of epithelial-mesenchymal transition markers across the Cancer Genome Atlas. Dev. Dyn. *247*, 555–564.

Grahammer, F., Herling, A.W., Lang, H.J., Schmitt-Gräff, A., Wittekindt, O.H., Nitschke, R., Bleich, M., Barhanin, J., and Warth, R. (2001). The cardiac K+ channel KCNQ1 is essential for gastric acid secretion. Gastroenterology *120*, 1363–1371.

Haas, B.R., and Sontheimer, H. (2010). Inhibition of the sodium-potassium-chloride

cotransporter isoform-1 reduces glioma invasion. Cancer Res.

Hall, B.A., Halim, K.B.A., Buyan, A., Emmanouil, B., and Sansom, M.S.P. (2014). Sidekick for Membrane Simulations: Automated Ensemble Molecular Dynamics Simulations of Transmembrane Helices. J. Chem. Theory Comput. *10*, 2165–2175.

Heitzmann, D., Grahammer, F., von Hahn, T., Schmitt-Gräff, A., Romeo, E., Nitschke, R., Gerlach, U., Lang, H.J., Verrey, F., Barhanin, J., et al. (2004). Heteromeric KCNE2/KCNQ1 potassium channels in the luminal membrane of gastric parietal cells. J. Physiol.

Hoadley, K.A., Yau, C., Hinoue, T., Wolf, D.M., Lazar, A.J., Drill, E., Shen, R., Taylor, A.M., Cherniack, A.D., Thorsson, V., et al. (2018). Cell-of-Origin Patterns Dominate the Molecular Classification of 10,000 Tumors from 33 Types of Cancer. Cell *173*, 291-304.e6.

Howard, R.J., Clark, K.A., Holton, J.M., Minor, D.L., and Jr. (2007). Structural insight into KCNQ (Kv7) channel assembly and channelopathy. Neuron *53*, 663–675.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: Visual molecular dynamics. J. Mol. Graph. *14*, 33–38, 27–28.

Jentsch, T.J. (2000). Neuronal KCNQ potassium channels: Physislogy and role in disease. Nat. Rev. Neurosci.

Kharkovets, T., Hardelin, J.-P., Safieddine, S., Schweizer, M., El-Amraoui, A., Petit, C., and Jentsch, T.J. (2002). KCNQ4, a K+ channel mutated in a form of dominant deafness, is expressed in the inner ear and the central auditory pathway. Proc. Natl. Acad. Sci.

Kim, J., Bowlby, R., Mungall, A.J., Robertson, A.G., Odze, R.D., Cherniack, A.D.,

Shih, J., Pedamallu, C.S., Cibulskis, C., Dunford, A., et al. (2017). Integrated genomic characterization of oesophageal carcinoma. Nature *541*, 169–174.

Kubisch, C., Schroeder, B.C., Friedrich, T., Lütjohann, B., El-Amraoui, A., Marlin, S., Petit, C., and Jentsch, T.J. (1999). KCNQ4, a novel potassium channel expressed in sensory outer hair cells, is mutated in dominant deafness. Cell.

Lee, S., Tran, A., Allsopp, M., Lim, J.B., Hénin, J., and Klauda, J.B. (2014). CHARMM36 united atom chain model for lipids and surfactants. J. Phys. Chem. B.

Li, J., Lu, Y., Akbani, R., Ju, Z., Roebuck, P.L., Liu, W., Yang, J.Y., Broom, B.M., Verhaak, R.G.W., Kane, D.W., et al. (2013). TCPA: A resource for cancer functional proteomics data. Nat. Methods.

Li, J., Akbani, R., Zhao, W., Lu, Y., Weinstein, J.N., Mills, G.B., and Liang, H. (2017). Explore, Visualize, and Analyze Functional Cancer Proteomic Data Using the Cancer Proteome Atlas. Cancer Res. *77*, e51–e54.

Li, X., Francies, H.E., Secrier, M., Perner, J., Miremadi, A., Galeano-Dalmau, N., Barendt, W.J., Letchford, L., Leyden, G.M., Goffin, E.K., et al. (2018). Organoid cultures recapitulate esophageal adenocarcinoma heterogeneity providing a model for clonality studies and precision therapeutics. Nat. Commun. *9*, 2983.

Love, M.I., Huber, W., and Anders, S. (2014). DESeq2.

Lundberg, S.M., and Lee, S.-I.I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems, pp. 4765–4774.

Lundberg, S.M., Erion, G.G., and Lee, S.-I. (2018). Consistent Individualized Feature Attribution for Tree Ensembles.

Manville, R.W., and Abbott, G.W. (2019). In silico re-engineering of a

neurotransmitter to activate KCNQ potassium channels in an isoform-specific manner. Commun. Biol. *2*, 401.

Martincorena, I., Raine, K.M., Gerstung, M., Dawson, K.J., Haase, K., Van Loo, P., Davies, H., Stratton, M.R., and Campbell, P.J. (2017). Universal Patterns of Selection in Cancer and Somatic Tissues. Cell *171*, 1029-1041.e21.

Miceli, F., Soldovieri, M.V., Hernandez, C.C., Shapiro, M.S., Annunziato, L., and Taglialatela, M. (2008). Gating consequences of charge neutralization of arginine residues in the S4 segment of Kv7.2, an epilepsy-linked K+ channel subunit. Biophys. J.

Miceli, F., Soldovieri, M.V., Ambrosino, P., De Maria, M., Migliore, M., Migliore, R., and Taglialatela, M. (2015). Early-onset epileptic encephalopathy caused by gain-of-function mutations in the voltage sensor of Kv7.2 and Kv7.3 potassium channel subunits. J. Neurosci. *35*, 3782–3793.

Millichap, J.J., Park, K.L., Tsuchida, T., Ben-Zeev, B., Carmant, L., Flamini, R., Joshi, N., Levisohn, P.M., Marsh, E., Nangia, S., et al. (2016). *KCNQ2* encephalopathy. Neurol. Genet. *2*, e96.

Monticelli, L., Kandasamy, S.K., Periole, X., Larson, R.G., Tieleman, D.P., and Marrink, S.-J. (2008). The MARTINI Coarse-Grained Force Field: Extension to Proteins. J. Chem. Theory Comput. *4*, 819–834.

Morita, H., Wu, J., and Zipes, D.P. (2008). The QT syndromes: long and short. Lancet *372*, 750–763.

Muzny, D.M., Bainbridge, M.N., Chang, K., Dinh, H.H., Drummond, J.A., Fowler, G., Kovar, C.L., Lewis, L.R., Morgan, M.B., Newsham, I.F., et al. (2012). Comprehensive

molecular characterization of human colon and rectal cancer. Nature *487*, 330–337.

Nerbonne, J.M., and Kass, R.S. (2005). Molecular Physiology of Cardiac Repolarization. Physiol. Rev.

Ohya, S., Asakura, K., Muraki, K., Watanabe, M., and Imaizumi, Y. (2015). Molecular and functional characterization of ERG, KCNQ, and KCNE subtypes in rat stomach smooth muscle. Am. J. Physiol. Liver Physiol.

Panaghie, G., and Abbott, G.W. (2007). The role of S4 charges in voltage-dependent and voltage-independent KCNQ1 potassium channel complexes. J. Gen. Physiol. *129*, 121–133.

Pardo, L.A., and Stühmer, W. (2013). The roles of K+ channels in cancer. Nat. Rev. Cancer *14*, 39–48.

Pedersen, S.F., and Stock, C. (2013). Ion channels and transporters in cancer: pathophysiology, regulation, and clinical potential. Cancer Res. *73*, 1658–1661.

Pedersen, S.F., Hoffmann, E.K., and Novak, I. (2013). Cell volume regulation in epithelial physiology and cancer. Front. Physiol. *4*, 233.

Phan, N.N., Wang, C.-Y., Chen, C.-F., Sun, Z., Lai, M.-D., and Lin, Y.-C. (2017). Voltage-gated calcium channels: Novel targets for cancer therapy. Oncol. Lett. *14*, 2059–2074.

Pillozzi, S., Brizzi, M., Balzi, M., Crociani, O., Cherubini, A., Guasti, L., Bartolozzi, B., Becchetti, A., Wanke, E., Bernabei, P., et al. (2002). HERG potassium channels are constitutively expressed in primary human acute myeloid leukemias and regulate cell proliferation of normal and leukemic hemopoietic progenitors. Leukemia *16*, 1791–1798.

Pointer, K.B., Clark, P.A., Eliceiri, K.W., Salamat, M.S., Robertson, G.A., and Kuo, J.S. (2017). Administration of Non-Torsadogenic human Ether-à-go-go-Related Gene Inhibitors Is Associated with Better Survival for High hERG–Expressing Glioblastoma Patients. Clin. Cancer Res. *23*, 73–80.

Qu, Z., Yao, W., Yao, R., Liu, X., Yu, K., and Hartzell, C. (2014). The Ca(2+) -activated Cl(-) channel, ANO1 (TMEM16A), is a double-edged sword in cell proliferation and tumorigenesis. Cancer Med. *3*, 453–461.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics.

Rapetti-Mauss, R., Bustos, V., Thomas, W., McBryan, J., Harvey, H., Lajczak, N., Madden, S.F., Pellissier, B., Borgese, F., Soriani, O., et al. (2017). Bidirectional KCNQ1:β-catenin interaction drives colorectal cancer cell differentiation. Proc. Natl. Acad. Sci. *114*, 4159–4164.

Reid, E.S., Williams, H., Stabej, P.L.Q., James, C., Ocaka, L., Bacchelli, C., Footitt, E.J., Boyd, S., Cleary, M.A., Mills, P.B., et al. (2015). Seizures Due to a KCNQ2 Mutation: Treatment with Vitamin B6. In JIMD Reports, (Wiley-Blackwell), pp. 79–84.

Robbins, J. (2001). KCNQ potassium channels: Physiology, pathophysiology, and pharmacology. Pharmacol. Ther.

Rogawski, M.A. (2000). KCNQ2/KCNQ3 K+channels and the molecular pathogenesis of epilepsy: Implications for therapy. Trends Neurosci.

Roger, S., Gillet, L., Le Guennec, J.-Y., and Besson, P. (2015). Voltage-gated sodium channels and cancer: is excitability their primary role? Front. Pharmacol. *6*, 152.

Rogers, M.F., Shihab, H.A., Gaunt, T.R., and Campbell, C. (2017). CScape: A tool for predicting oncogenic single-point mutations in the cancer genome. Sci. Rep.

Rosenthal, R., McGranahan, N., Herrero, J., Taylor, B.S., and Swanton, C. (2016). deconstructSigs: Delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. Genome Biol.

Šali, A., and Blundell, T.L. (1993). Comparative Protein Modelling by Satisfaction of Spatial Restraints. J. Mol. Biol. *234*, 779–815.

Šali, A., Potterton, L., Yuan, F., van Vlijmen, H., and Karplus, M. (1995). Evaluation of comparative protein modeling by MODELLER. Proteins Struct. Funct. Bioinforma.

Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W.K., Luna, A., La, K.C., Dimitriadoy, S., Liu, D.L., Kantheti, H.S., Saghafinia, S., et al. (2018). Oncogenic Signaling Pathways in The Cancer Genome Atlas. Cell.

Sands, T.T., Miceli, F., Lesca, G., Beck, A.E., Sadleir, L.G., Arrington, D.K., Schönewolf-Greulich, B., Moutton, S., Lauritano, A., Nappi, P., et al. (2019). Autism and developmental disability caused by *KCNQ3* gain-of-function variants. Ann. Neurol. *86*, 181–192.

Shihab, H.A., Gough, J., Cooper, D.N., Stenson, P.D., Barker, G.L.A., Edwards, K.J., Day, I.N.M., and Gaunt, T.R. (2013). Predicting the Functional, Molecular, and Phenotypic Consequences of Amino Acid Substitutions using Hidden Markov Models. Hum. Mutat.

Shorthouse, D., Riedel, A., Kerr, E., Pedro, L., Bihary, D., Samarajiwa, S., Martins, C.P.C.P., Shields, J., and Hall, B.A.B.A. (2018). Exploring the role of stromal osmoregulation in cancer and disease using executable modelling. Nat. Commun. *9*,

3011.

Smart, O.S., Neduvelil, J.G., Wang, X., Wallace, B.A., and Sansom, M.S.P. (1996). HOLE: A program for the analysis of the pore dimensions of ion channel structural models. J. Mol. Graph.

Soldovieri, M.V., Ambrosino, P., Mosca, I., Miceli, F., Franco, C., Canzoniero, L.M.T., Kline-Fath, B., Cooper, E.C., Venkatesan, C., and Taglialatela, M. (2019). Epileptic Encephalopathy In A Patient With A Novel Variant In The Kv7.2 S2 Transmembrane Segment: Clinical, Genetic, and Functional Features. Int. J. Mol. Sci. *20*.

Stansfeld, P.J., and Sansom, M.S.P.P. (2011). From Coarse Grained to Atomistic: A Serial Multiscale Approach to Membrane Protein Simulations. J. Chem. Theory Comput. *7*, 1157–1166.

Stansfeld, P.J., Goose, J.E., Caffrey, M., Carpenter, E.P., Parker, J.L., Newstead, S., and Sansom, M.S.P. (2015). MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. Structure.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U. S. A.

Sun, J., and MacKinnon, R. (2017). Cryo-EM Structure of a KCNQ1/CaM Complex Reveals Insights into Congenital Long QT Syndrome. Cell.

Takahashi, N., Chen, H.Y., Harris, I.S., Stover, D.G., Selfors, L.M., Bronson, R.T., Deraedt, T., Cichowski, K., Welm, A.L., Mori, Y., et al. (2018). Cancer Cells Co-opt the Neuronal Redox-Sensing Channel TRPA1 to Promote Oxidative-Stress

Tolerance. Cancer Cell.

Tarran, R. (2004). Regulation of Airway Surface Liquid Volume and Mucus Transport by Active Ion Transport. Proc. Am. Thorac. Soc. *1*, 42–46.

Tate, J.G., Bamford, S., Jubb, H.C., Sondka, Z., Beare, D.M., Bindal, N., Boutselakis, H., Cole, C.G., Creatore, C., Dawson, E., et al. (2019). COSMIC: The Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res.

Than, B.L.N., Goos, J.A.C.M., Sarver, A.L., O'Sullivan, M.G., Rod, A., Starr, T.K., Fijneman, R.J.A., Meijer, G.A., Zhao, L., Zhang, Y., et al. (2014). The role of KCNQ1 in mouse and human gastrointestinal cancers. Oncogene.

Vodnala, S.K., Eil, R., Kishton, R.J., Sukumar, M., Yamamoto, T.N., Ha, N.-H., Lee, P.-H., Shin, M., Patel, S.J., Yu, Z., et al. (2019). T cell stemness and dysfunction in tumors are triggered by a common mechanism. Science.

Wang, J.J., and Li, Y. (2016). KCNQ potassium channels in sensory system and neural circuits. Acta Pharmacol. Sin.

Wang, Y., Wang, L., Yin, C., An, B., Hao, Y., Wei, T., Li, L., and Song, G. (2015). Arsenic trioxide inhibits breast cancer cell growth via microRNA-328/hERG pathway in MCF-7 cells. Mol. Med. Rep. *12*, 1233–1238.

Warth, R., Garcia Alzamora, M., Kim, J., Zdebik, A., Nitschke, R., Bleich, M., Gerlach, U., Barhanin, J., and Kim, S. (2002). The role of KCNQ1/KCNE1 K+ channels in intestine and pancreas: lessons from the KCNE1 knockout mouse. Pflügers Arch. - Eur. J. Physiol. *443*, 822–828.

Ye, J., Pavlicek, A., Lunney, E.A., Rejto, P.A., and Teng, C.H. (2010). Statistical method on nonrandom clustering with application to somatic mutations in cancer.

BMC Bioinformatics.

Yu, H., Lin, Z., Mattmann, M.E., Zou, B., Terrenoire, C., Zhang, H., Wu, M., McManus, O.B., Kass, R.S., Lindsley, C.W., et al. (2013). Dynamic subunit stoichiometry confers a progressive continuum of pharmacological sensitivity by KCNQ potassium channels. Proc. Natl. Acad. Sci.

Zaydman, M.A., and Cui, J. (2014). PIP2 regulation of KCNQ channels: biophysical and molecular mechanisms for lipid modulation of voltage-dependent gating. Front. Physiol. *5*, 195.