ONCOchannelome: A computational framework to investigate altered ion channels across tumor types

K. T. Shreya Parthasarathi[1,2], Soundharya Ramu[3], Mohit Kumar Jolly[4] and Jyoti Sharma[1,2]*

[1]Manipal Academy of Higher Education (MAHE), Manipal, Karnataka 576104, India

[2]Institute of Bioinformatics, International Technology Park, Bangalore 560066, India

[3] IISc Mathematics Initiative, Indian Institute of Science, Bangalore 560012, India

[4]Department of Bioengineering, Indian Institute of Science, Bangalore 560012, India

***Correspondence:**

Dr. Jyoti Sharma, Institute of Bioinformatics, International Technology Park, Bangalore 560066, India

jyoti@ibioinformatics.org

Short title: ONCOchannelome – Investigating altered ion channels in cancers

Abstract:

Conventional approaches in analyzing ion channels in cancer primarily focus on detection of membrane potential to record the polarization states of the cellular membrane. Ion channels are known to exhibit context dependent activity that often depends on cancer type and metabolic state of the tumor. Here, we developed a computational framework – ONCOchannelome that allows identification of potential ion channels using transcriptomic datasets. We collected publicly available RNA-Seq datasets corresponding to 2421 normal, 5442 primary tumors and 588 metastatic samples of 15 tumor types. Using the data we designed a computational strategy to merge the datasets, identifying differentially expressed ion channels in tumor and metastatic states aiding in determination of ion channels in the different tumor states. Subsequently, we developed strategies to determine the correlation of altered ion channels with epithelial to mesenchymal transition and to identify differentially co-expressed ion channel networks along with their possible transcription factors. The ONCOchannelome can be utilized to compare the alterations in ion channels in histological subtypes of lung cancer. The ONCOchannelome supports visualizing the altered ion channels in tumor and metastatic states in addition to providing the changes resulting in biological processes and molecular functions as the tumor progresses. The ONCOchannelome can be used to identify altered ion channels significant in various tumors that may aid in understanding tumor behaviour in turn facilitating exploration of their role in hallmarks of cancers.

Author Summary

Studying ion channels using current available methods often rely on monitoring membrane voltage using patch clamp technique or computationally simulating channel behaviour using models. Using these methods in the initial stages are considered to be time-consuming,

difficult to scale and dependent on models that may or may not reflect actual tumor environment. However, gene expression patterns of the ion channels could be studied to obtain significant insights on their role and importance in a particular tumor type. In view of this, we developed ONCOchannelome, a computational framework to explore alterations in ion channels across tumors by utilizing publicly available transcriptome profiles of patients with different tumor types. ONCOchannelome would enhance our understanding of the dependence of ion channels in the different tumors in addition to providing their observed alterations on progressing from primary tumor state to metastatic state.

1. Introduction:

Ion channels, the sequentially and structurally distinct membrane proteins expressed in all cell types, are known for selective permeability of ion species within and across cell membranes. Furini and Domene reported that the human genome encodes over 400 distinct ion channels, accounting for approximately 2% of all identified genes [1]. Traditional methods of studying these proteins lie within the scope of biophysicists and physiologists primarily assessing physiological processes including membrane excitability and epithelial transport. However, in recent years growing evidence suggests ion channels to be part of processes far beyond traditional physiological functions. It is now known that understanding the chemistry, electric potential and mechanics of ion channels could aid in regulating epigenetic and gene expression related activities in a cell [2], [3], [4]. Additionally, ion channels were observed to function as transcription factors as well as participate in cellular signaling through molecular interactions with macromolecular complexes [5], [6].

Aberrant expression of ion channels has been reported in multiple tumors [7], [8], [9], [10], [11]. Essentially, the altered ion channel landscape or "channelome" in tumor cells appear to play an active role in processes such as neoplastic transformation, tumor progression, adaptation to tumor microenvironment, metastasis, and therapy resistance [12], [13]. Unlike oncogenes, ion channels are generally not considered to trigger the onset of cancer, but they are known to drive oncogenic processes by modulating factors including calcium signaling, membrane voltage, sodium concentration, and cellular pH [13]. These observations along with experimental interferences demonstrating associations of altered ion channels and tumor progression resulted in the formation of the term oncochannels [13]. In parallel, several pharmacological modulators of ion channels are already in clinical use or under clinical trials [14]. Overexpression of some oncochannels could qualify them as tumor-associated antigens suitable for vaccine development [15]. Additionally, their presence at the cell surface would

also allow for selective targeting. Specific antibodies with various strategies are already being used to exploit ion channels demonstrating success in preclinical models and clinical trials [16], [17], [18].

Despite recent progress, characterizing ion channels continues to present theoretical and computational challenges. Involvement of ion channels in multiple physiological processes along with their role in the development of cancer increases the challenges in assigning specific mechanisms to ion channels in cancer hallmarks [19]. Additionally, the small size, dynamic nature and diversity, present difficulties in achieving high resolution structural information for understanding complex interactions with neighboring molecules [20], [21]. Moreover, molecular dynamics simulations of ion channels face limitations due to computational constraints including the accuracy of force fields that lack the ability to capture biologically relevant timescales [22]. The lack of annotations for many ion channels in cancer can be attributed to these limitations. To facilitate the detection of potential ion channels that could be targeted in various tumors, we developed the ONCOchannelome – a computational framework to explore deregulated ion channels across tumors. Omics technologies such as transcriptomics, aid in understanding various diseases including stroke, diabetes, and cancers [23].

2. Results:

2.1 ONCOchannelome modules

ONCOchannelome is comprised of four modules: Browse by tumor type, ion channels retrieval, EMT correlation and ion channels co-expression network (Figure 1).

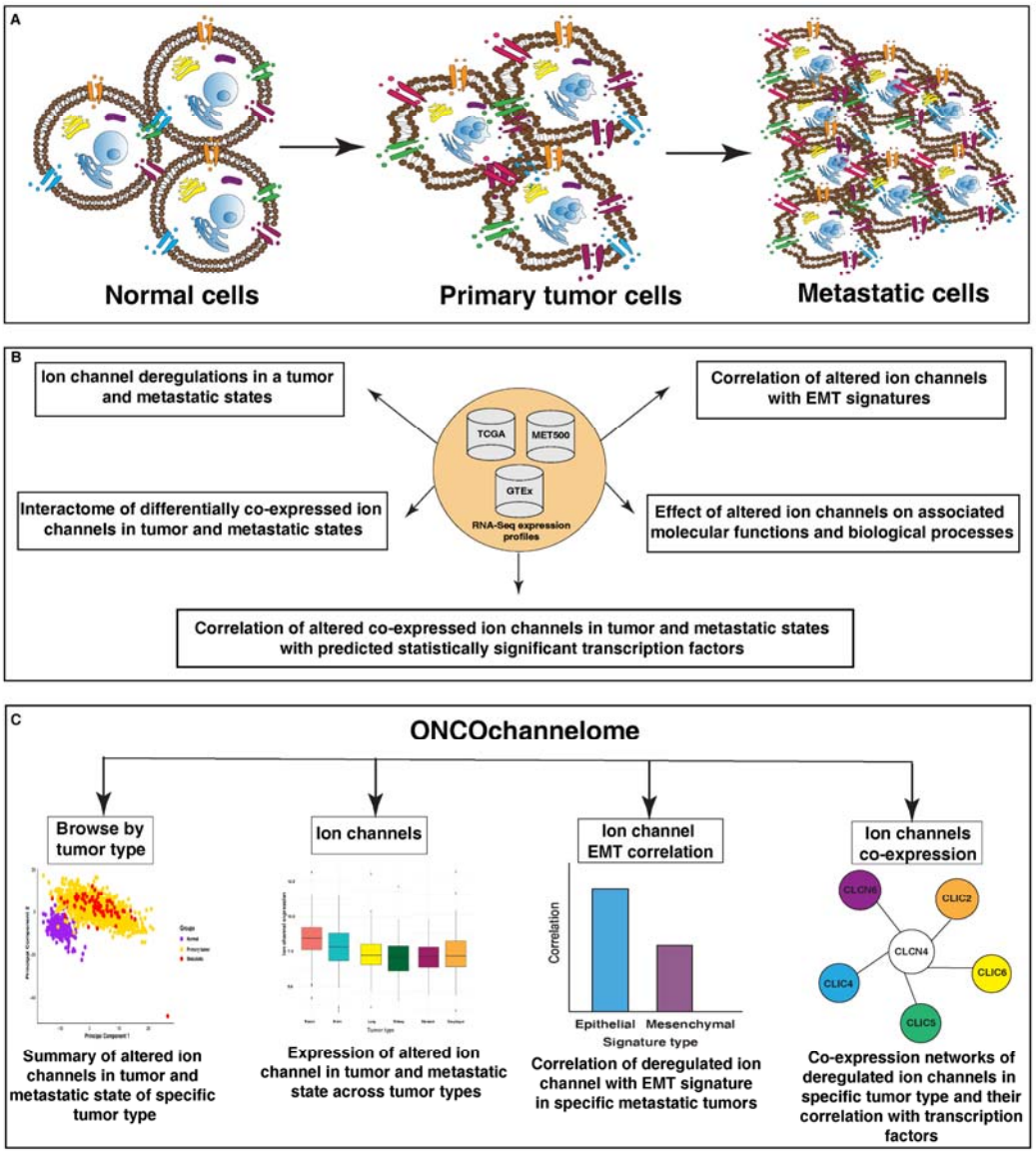

Figure 1: Schematic overview of the ONCOchannelome: A) The transformation between normal to tumor and tumor to metastatic states during tumor progression involves the cells to undergo molecular alterations resulting in significant changes in the cell morphology that reflect the traits required for metastatic invasion. B) The ONCOchannelome consists of modules that facilitates determination of differentially expressed (DE) ion channels in tumor and metastatic states of 15 tumor types, correlates the DE ion channels with epithelial and

mesenchymal gene signatures, constructs ion channel co-expression networks, provides details on biological processes and molecular functions of DE ion channels and provides correlations with potential transcription factors. C) Schematic of the ONCOchannelome. The computational framework allows user to explore ion channels using four modules: Browse by tumor type, Ion channels, Ion channel EMT correlation and Ion channels co-expression

The browse by tumor type module allows the users to explore the tumor types in terms of the overall deregulations observed in the ion channels. The ion channel retrieval module allows searches of expression and transcriptomic changes of a specific ion channel across tumor types. The ion channel EMT correlation module can be used to determine the correlations between the altered ion channel in a specific tumor type at a specific tumor state with epithelial and mesenchymal gene signatures. The fourth module, ion channel co-expression, can be used to determine the ion channels that are found to be co-expressed with the ion channel of interest in a specific tumor type at a tumor state. The significant associations of the ion channel co-expression network with the transcription factors can also be identified through this module.

2.2 Browse by tumor type

We have included 5442 primary tumor samples; 588 metastatic samples and 2421 normal samples distributed across 15 tumor types from 14 primary sites - breast invasive carcinoma (BRCA), colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), head and neck squamous cell carcinoma (HNSC), pancreatic adenocarcinoma (PAAD), prostate adenocarcinoma (PRAD), skin cutaneous melanoma (SKCM), thyroid carcinoma (THCA), bladder urothelial carcinoma (BLCA), cholangiocarcinoma (CHOL), glioblastoma multiforme (GBM), kidney chromophobe (KICH), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), and stomach adenocarcinoma (STAD) (Table 1).

Tumor	GTEx-Normal	TCGA-Normal	Total-Normal	TCGA-Tumor	TCGA-Metastatic	MET500-Metastatic	Total-Metastatic
BRCA	178	113	291	1092	7	67	74
COAD	241	41	282	287	1	10	11
ESCA	168	13	181	181	1	9	10
HNSC	0	44	44	518	2	21	23
PAAD	167	4	171	178	1	12	13
PRAD	99	52	151	494	1	69	70
SKCM	101	1	102	102	366	12	378
THCA	279	59	338	504	8	1	9
BLCA	9	19	28	407	0	0	0
CHOL	0	9	9	36	0	0	0
GBM	147	5	152	153	0	0	0
KICH	0	65	65	66	0	0	0
LUAD	288	109	397	513	0	0	0
LUSC	288	109	397	498	0	0	0
STAD	174	36	210	413	0	0	0

Table 1: Number of samples available across tumor types across normal, primary and metastatic states

Previously, we have reported the DE ion channels in tumor and metastatic state of breast cancer [9]. Ion channels were identified as potential candidates for the determination of prognosis in patients with breast cancer. Given the non-uniform expression pattern of ion channels in a specific sample state, the ion channels overexpressed in a particular cancer type may or may not overlap with the ion channels overexpressed in other cancer types or sample states. Furthermore, the details available through literature are scattered and not easily accessible. In view of this, we collected the list of ion channels and overlayed them to represent the unique patterns of altered expression observed in each tumor type and tumor state (Table 2).

Table 2: Number of upregulated and downregulated ion channels across tumor types across primary and metastatic states

Tumor	TN		MT		MN	
	Up	Down	Up	Down	Up	Down

BRCA	88	139	56	71	107	148
COAD	67	206	79	131	83	220
ESCA	81	151	122	65	117	120
HNSC	85	123	89	149	87	172
PAAD	98	91	126	96	156	118
PRAD	46	169	80	180	76	226
SKCM	131	110	59	36	140	90
THCA	109	72	4	1	120	62
BLCA	76	138	-	-	-	-
CHOL	200	59	-	-	-	-
GBM	95	229	-	-	-	-
KICH	93	201	-	-	-	-
LUAD	83	163	-	-	-	-
LUSC	100	188	-	-	-	-
STAD	101	133	-	-	-	-

The impact on the molecular functions, biological processes and cellular compartments due to alterations observed could also be established from the datasets analysed.

2.3 Ion channel retrieval

The module includes details on 493 ion channels. For each ion channel, information retrieved from public data resources like, FireBrowse, TumorPortal, GTEx, depmap, PubMed, Harmonizome, cancerscem and MET500 were linked with ONCOchannelome. Thereafter, the expression levels post performing batch correction to combine the datasets from the different data sources were plotted to depict the expression profiles for each individual ion channel across tumors. ONCOchannelome enabled altered ion channel visualisation will help users to identify the extent of deregulation of a particular ion channel across tumors at different tumor state.

2.4 Ion channel EMT correlation

Previously we have reported the correlation of altered ion channels with EMT gene signature scores indicating mesenchymal phenotypes in patients with breast cancer and esophageal squamous cell carcinoma [9], [8]. EMT-signature scores are calculated using the previously

published signature gene-sets. ssGSEA enrichment scores for epithelial and mesenchymal signature gene sets were computed across transcriptome profiles of the available metastatic tumor samples. Normalized enrichment scores revealed distinct patterns of epithelial and mesenchymal signatures across tumor types. Thereafter, the correlation of the DE ion channels in the specific tumor type with the calculated EMT scores was computed. Through ONCOchannelome ion channel EMT correlation panel, users can select an ion channel of interest and the tumor type and submit a query to evaluate the relationship between epithelial and mesenchymal gene signatures with the ion channel in normal, tumor and metastatic states of the particular tumor type.

2.5 Ion channel co-expression

The association between the altered ion channels with strongly co-expressed ion channels along with its extent of deregulation can facilitate uncover the possible relationship between the ion channels based on their expression patterns across different samples. ONCOchannelome allows user to explore this association of altered ion channel expression with expression of co-expressed ion channels. Furthermore, it allows user to identify the transcription factors that may regulate the expression of the co-expressed network. This might provide insights into the regulatory mechanisms underlying the observed ion channel expression patterns in a particular tumor type.

2.6 Case study

2.6.1 Comparative analysis of ion channel modulation in histological subtypes of lung cancer

ONCOchannelome can be utilized to compare the alterations in ion channels observed in adenocarcinoma versus squamous cell carcinoma. We performed a comprehensive comparison study between LUAD and LUSC. ONCOchannelome consisted of 397 normal

samples in both LUAD and LUSC whereas it consisted of 513 primary tumor samples in LUAD and 498 primary tumor samples in LUSC. 83 ion channels were observed to be upregulated in LUAD and 100 ion channels were upregulated in LUSC. Similarly, there were about 163 and 188 ion channels observed to be downregulated in the two tumor types, respectively. Overlaying the altered ion channels in the two subtypes indicated an overlap of 208 ion channels. Alternatively, 38 and 80 ion channels were identified as subtype exclusive ion channels in LUAD and LUSC respectively. The ion channels *GJB7*, *KCTD15*, *SLC10A6*, *CACNA2D3*, *KCNK5*, *CRACR2B* and *SLC41A2* were found to show divergent expression patterns in LUAD and LUSC subtypes. The variability observed in terms of alterations was reflected in the enriched biological processes and molecular functions as well. Two biological processes namely, glycerol transmembrane transport and inhibitory synapse assembly were identified as exclusively activated in LUSC subtype. Similarly, the biological processes: inorganic cation transmembrane transport, plasma membrane organization and plasma membrane phospholipid scrambling were activated only in LUAD subtype. The molecular functions: carbohydrate transmembrane transporter activity, polyol transmembrane transporter activity and glycerol transmembrane transporter activity were exclusively activated and symporter activity, active monoatomic ion transmembrane transporter activity, antiporter activity, organic acid transmembrane transporter activity and carboxylic acid transmembrane transporter activity were suppressed exclusively in LUSC subtype. Likewise, intracellular calcium activated chloride channel activity, GABA-gated chloride ion channel activity, high voltage gated calcium channel activity and GABA receptor activity were activated whereas monoatomic cation transmembrane transporter activity, metal ion transmembrane transporter activity, inorganic cation transmembrane transporter activity, calcium channel activity and calcium ion transmembrane transporter activity were observed to be suppressed exclusively in LUAD subtype. The cellular compartments observed to be

activated in LUAD involved the regions specific to neuronal transport indicating its potential in brain metastasis whereas the ion channels involved in neuronal mechanisms were suppressed in LUSC subtype. Co-expression networks of the ion channels showing divergent expression patterns provided further insights on their importance in terms of tumorigenicity. For example, *GJB7* was identified as elevated in LUSC and suppressed in LUAD. The co-expression network of ion channels in LUSC showed *GJB7* to be highly significant within the network (0.96, p-value: $1.20e^{-52}$) while the network was negatively correlated with normal samples (-0.89, p-value: $2.00e^{-300}$) and positively correlated with tumor samples (0.89, p-value: $2.00e^{-300}$). All the ion channels in the network were downregulated. The transcription factor *EZH2* was identified as highly positively correlated with the co-expression network (0.87, p-value: $2.28e^{-238}$) and *E2F1* was moderately positively correlated with the network (0.67, p-value: $6.44e^{-122}$). In LUAD the co-expression network indicated *GJB7* to be highly significant with the network (0.87, p-value: $1.30e^{-51}$) with the network being positively correlated to normal samples (0.8, p-value: $2.00e^{-207}$) and negatively correlated to tumor samples (-0.8, p-value: $2.00e^{-207}$). All the ion channels excluding *AQP6*, a water transporter were identified as downregulated in the network. *JUN* was identified as the transcription factor (p-value: 0.03) regulating *GJB7* downregulation along with the ion channels *SLC10A4*, *GABRB3*, *SLC10A6*, *SLC30A8*, *KCNIP1*, *KCNJ12*, *KCNC4*, *SLC16A12*, *KCNAB1*, *KCNA5*¸ *GJC1*, *GRIN2A*, *GJA5*, *KCNMB1*, *CACNG4* and *KCNJ3* (Figure 2).

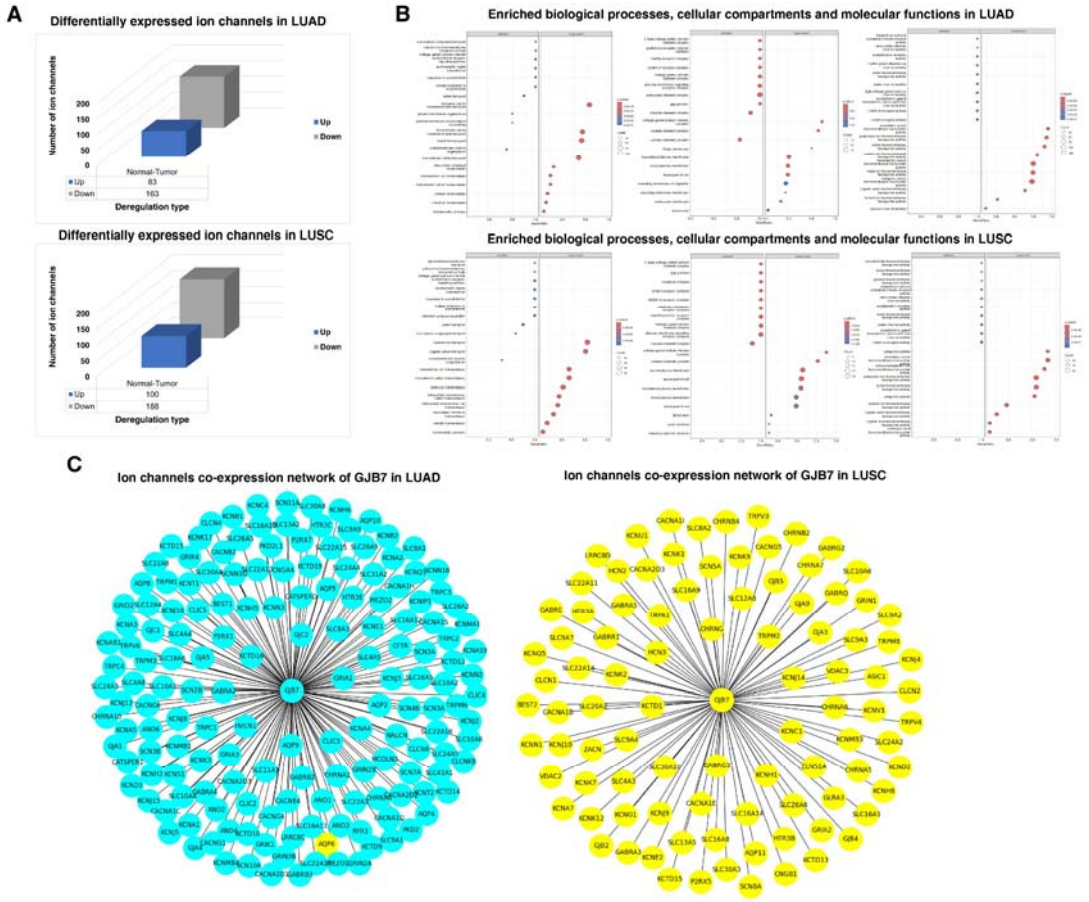

Figure 2: Altered ion channel in histological subtypes of lung cancer: A) Number of differentially expressed ion channels in LUAD and LUSC samples. B) Dot plot of activated and suppressed biological processes, molecular functions and cellular compartments corresponding to differentially expressed ion channels in LUAD and LUSC. C) Ion channel co-expression network of differentially co-expressed ion channels with GJB7 in LUAD and LUSC

2.6.2 Alterations in ion channels in tumor and metastatic states of skin cutaneous melanoma

The ONCOchannelome has eight tumors with samples corresponding to primary tumor and metastatic states. We report a comprehensive overview of the ion channels found to be altered in tumor and metastatic states of SKCM. ONCOchannelome consisted of 102 normal and

primary tumor samples corresponding to SKCM and 378 metastatic samples. The differential gene expression analysis carried out on comparing normal with tumor samples (TN), tumor with metastatic samples (MT) and normal with metastatic samples (MN) indicated 43 DE ion channels to overlap across the states. 27, 5 and 31 ion channels were exclusively DE in TN, MT and MN respectively. Of the 43 ion channels, 10 exhibited uniform expression patterns: either consistently upregulated or downregulated across all SKCM states. The remaining 33 ion channels demonstrated consistent deregulation when comparing normal tissue to both tumor and metastatic samples. However, when tumor samples were compared to metastatic ones, these ion channels showed a reversal in expression direction. The gene set enrichment analysis obtained from ONCOchannelome shows unique set of molecular functions to be enriched in each comparison state with only the function 'identical protein binding' to be common in comparison of normal samples with tumor and metastatic samples. Similarly, the molecular functions adenyl nucleotide binding, adenyl ribonucleotide binding and neurotransmitter receptor activity overlapped in comparison of metastatic samples with tumor and normal samples. Of the ten ion channels, *GJB2*, *GLRB*, *GRIA3*, *GRID2*, *HTR3A*, *KCND2*, *KCTD16*, *PIEZO2* and *SLC26A4* showed positive correlation with mesenchymal gene signatures. Of the ion channels identified as exclusively altered in tumor to metastatic state comparison, *CACNA1E*, *SCN7A*, *CACNA1G* and *KCNK1* showed an increase in positive correlation with mesenchymal gene signatures as compared to epithelial gene signatures indicating their significance in tumor progression (Figure 3).

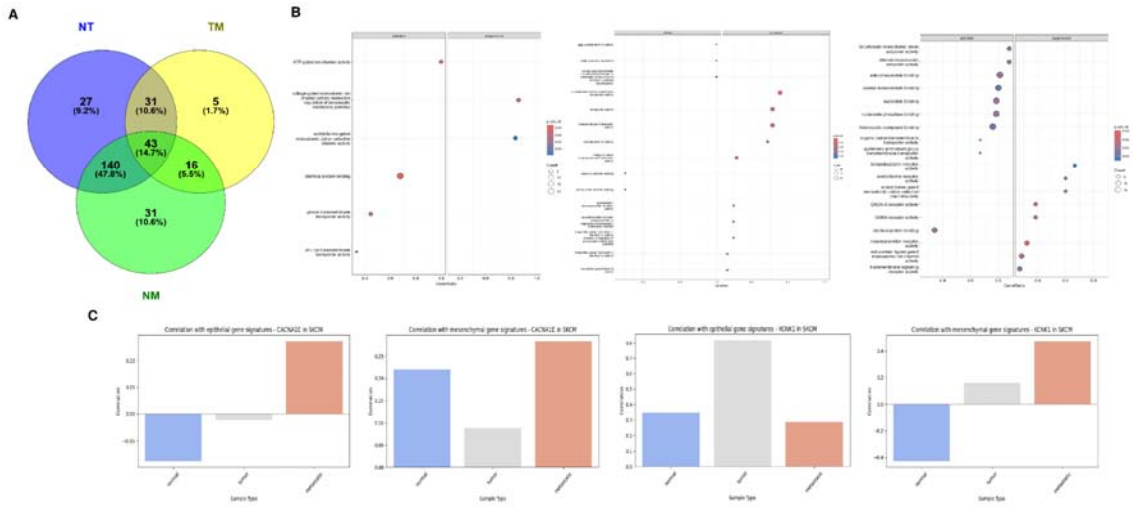

Figure 3: Ion channels altered in tumor and metastatic states of skin cutaneous melanoma: A) Overlaps in differentially expressed (DE) ion channels in comparison of tumor samples with normal samples (TN), metastatic samples with tumor samples (MT) and metastatic samples with normal samples (MN). B) Dot plot of activated and suppressed molecular functions corresponding to DE ion channels in TN, MT and MN states. C) Correlation of epithelial and mesenchymal gene signatures with CACNA1E and KCNK1 in normal, tumor and metastatic expression profiles

2.7 User interface and information retrieval

ONCOchannelome has a simple user interface for data exploration of the altered ion channels across tumor types. The interface enables user to identify the alterations in ion channels and retrieve their significance in various biological processes and molecular functions. Hyperlinks are provided to various projects that allow user to correlate the altered ion channels with other studies that have identified the mutations and expression patterns in ion channels in various tumors using different datasets. Correlation with EMT gene signatures can be obtained for individual ion channels altered in specific tumors. Co-expression of ion channels with other significantly altered ion channels would allow the user to look for gene

clusters that may function together resulting in downstream processes. The potential transcription factors can also be identified in order to hypothesize the regulation of biological processes being deregulated within the tumor that may provide insights into regulatory mechanisms underlying observed gene expression patterns.

3. Discussion:

The ONCOchannelome computational framework was designed to devise a computational method to identify oncochannels across tumor types and to enhance our understanding of ion channels in tumor progression. It represents a computational platform to provide exploration of ion channels being deregulated along with provision of details in terms of their correlation with EMT. ONCOchannelome goes beyond currently available resources, IUPHAR-DB [24], ChanFAD [25] and Channelpedia [26] with a focus on ion channels and their involvement in cancer. ONCOchannelome can be used for: (i) Exploration of alterations in ion channels in tumor and metastatic states across tumor types, (ii) Correlation of specific ion channels identified to be altered in particular tumors with epithelial and mesenchymal gene signatures (iii) Retrieval of biological processes, molecular functions and cellular compartments with a possibility to be activated or suppressed on deregulation of specific ion channels (iv) Prediction of potential transcription factors with a potential to regulate the altered ion channels (v) Determination of ion channel interactome based on statistical significance from transcriptome gene co-expression analysis and visualization of results.

Currently, ONCOchannelome incorporates processed data utilizing the developed computational pipeline using TCGA, MET500 and GTEx dataset consisting of 5442 primary tumor samples, 588 metastatic samples and 2421 normal samples with information on 493 ion channels. Also included are hyperlinks towards high throughput data analysis projects: FireBrowse, TumorPortal, GTEx, PubMed, Harmonizome, cancerscem and MET500 that

allows broad spectrum comparison of ion channels across datasets other than the ones included in the computational framework. Notably, the alterations of specific ion channels across tumors provides insights on tumor specific ion channel alteration detection. Venn diagrams generated are provided to look for overlaps across altered ion channels in tumor and metastatic states of specific tumor types. The distribution of RNA abundance in normal, tumor and metastatic states across tumors for each ion channel are provided in the form of boxplots. The correlation between EMT gene signatures and DE ion channels were assessed across normal, tumor, and metastatic states across tumor types. This correlation of altered ion channels in each tumor type with epithelial and mesenchymal gene signatures can be visualised in the form of bar plots. The plots could be utilised to observe associations that may reflect the general EMT-related transcription shifts across heterogeneous samples of a specific tumor. Although the correlation analysis did not take into account for alternative variables such as histological subtype, tumor stage or demographics, the observed correlations were consistent with findings from previous studies. For instance, mRNA expression of transient receptor potential cation channel subfamily C member 1 (*TRPC1*) was reported to have a strong association with mesenchymal phenotype in endometrial carcinoma [27]. Through our methodology, expression of *TRPC1* in BRCA, COAD, HNSC, PRAD and SKCM had a higher correlation with mesenchymal gene signatures as compared to epithelial gene signatures that may indicate aggressive tumor phenotype in these tumors. Similarly, potassium voltage-gated channel subfamily Q member 1 (*KCNQ1*) was identified to have potential significance in epithelial cancers wherein low expression of *KCNQ1* was reported to promote colon adenocarcinoma [28]. The expression of *KCNQ1* was identified to be associated with epithelial cell plasticity indicating it to promote epithelial features [28]. Through our methodology, expression of *KCNQ1* in multiple tumors including COAD, ESCA, PAAD, PRAD and THCA had higher correlation with epithelial gene signatures as

compared to mesenchymal gene signatures. Similarly, gap junction *GJB7* has been previously reported to harbour somatic mutations in gastric and colorectal cancers [29]. Our study reports transcriptional alterations in *GJB7* in lung carcinoma. Further, gene co-expression networks along with indication of upregulation and downregulation of each ion channel in the network is provided. The correlation values of the network with binary traits - normal, tumor and metastatic further allows to establish statistical relevance of the network. The computational framework also includes transcription factor identification based on the gene co-expression networks generated and their correlation with the complete network. Further, we have provided two case studies to help users explore the framework for their research.

Limitations:

The utility of the computational framework – ONCOchannelome is best realized during the discovery phase, as it facilitates generation of hypothesis by identifying potential ion channels altered in specific tumor types. Translational interpretation of the computationally established ion channels would require further experimental validation. Furthermore, the imbalance in the sample distribution across normal, tumor and metastatic states may constrain statistical power for detecting subtle alterations specific to progression. Nonetheless, the utilized dataset provides a robust and valuable foundation for comparative analysis and the devised framework could be used for larger datasets on availability. ONCOchannelome provides a broader view of EMT-association with DE ion channels. Incorporation of clinical metadata may be required to enhance and validate context-specific EMT dynamics. Transcription factors predicted for DE ion channels derived from multiple curated databases are computational in nature. They offer valuable insights and serve as basis for further investigations.

In summary, ONCOchannelome serves as a comprehensive resource for identifying tumor and metastatic related ion channels in 15 tumor types, exploring EMT-related correlations for each altered ion channels, and for analysis of the relevance of deregulation of the ion channel in biological processes and molecular functions. The interface and the structure of the framework allow for inclusion of additional cancer types on availability of datasets. Currently, ONCOchannelome lacks inclusion of other omics layers in terms of ion channel alterations data, but it could be explored with availability of new datasets.

4. Design and implementation

4.1 Databases

For the functionality of ONCOchannelome described herein, curated ion channel lists from literature and gene symbols established in Human Genome Organization Gene Nomenclature Committee (HGNC) database were used.

4.2 Input data

The analysis described in this manuscript to develop the computational framework utilized publicly available transcriptome profiling data accessible from The Cancer Genome Atlas (TGCA), Genotype Tissue Expression (GTEx) and Integrative Clinical Genomics of Metastatic Cancer (MET500) data resources. The cancer cohorts included were: BRCA, COAD, ESCA, HNSC, PAAD, PRAD, SKCM, THCA, BLCA, CHOL, GBM, KICH, LUAD, LUSC and STAD (Figure 4).

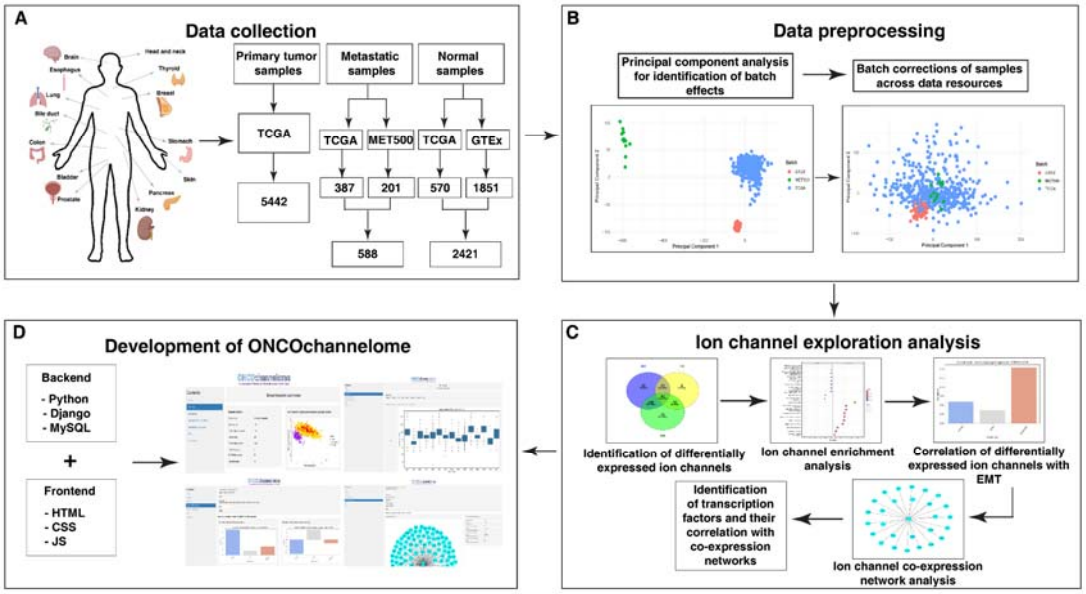

Figure 4: The depiction of overall workflow to develop the computational framework, 'ONCOchannelome' to investigate altered ion channels across tumor types

4.3 Preprocessing

Normal individual data from GTEx were selected based on age, gender and tumor sample ratio normalization. Principle component analysis (PCA) was performed across samples belonging to a specific tumor type and batch effects were identified. R Bioconductor limma (v.3.60.6) [30], afffy (v.1.82.0) [31] and sva (v.3.52.0) [32] package were utilized to perform batch corrections based on sample group as well as sample type specifically using removeBatchEffect function from limma package (S1 Fig and S2 Fig).

4.4 Detection of alterations in ion channels

R Bioconductor limma package was used to estimate the differentially expressed genes (DEGs) based on Bayes method. DEGs were identified for 3 subgroups – tumor vs normal (TN), metastatic vs. tumor (MT) and metastatic vs. normal (MN). Any gene with an adjusted p-value < 0. 05 calculated using Benjamini-Hochberg method of false discovery rate and

log2FC |0.6| (corresponding to FC > 1.5 or < 0.66) between the conditions were considered as DEGs. The number of altered ion channels in each subgroup was given as a subset of the total DEGs obtained as described by the following.

$$IC_D \subseteq D_G \subseteq G \text{ and } |IC_D| \leq |IC| \leq |G| \quad (1)$$

Where G is the total number of transcripts, D_G is the total number of DEGs, IC is the total number of curated ion channels and IC_D is the total number of DE ion channels.

Gene Ontology enrichment

We performed Gene Ontology Biological Process, Molecular Function and Cellular Compartment gene set enrichment analysis (GSEA) on the IC_D in each subgroup using R Bioconductor clusterProfiler library (v.4.2.2) [33]. The minimum gene size was set to 5 and maximum was set to 300, p-value cutoff was set to 0.05 whereas the number of permutations was set to 10000. We show the significantly enriched terms for the IC_D in each subgroup.

4.5 Construction of ion channel co-expression networks

To construct a state-specific ion channel network, we provided the transcriptome profiles corresponding to the IC_D to weighted gene co-expression network analysis (WGCNA) library (v.1.73) [34], an R Bioconductor package. The expression profiles were pre-processed using the goodSampleGenes function and the genes without any variance across samples were filtered out. A scale-free network model was generated followed by construction of adjacency correlation matrix in order to obtain an appropriate soft threshold power. The adjacency matrix was further transformed into a topological overlap matrix (TOM) further resulting into a dissimilarity matrix. Thereafter we performed hierarchical clustering using DynamicTreeCut algorithm and significant gene modules were selected. The expression profiles of each of the modules summarized in the form of module eigengenes were correlated with the binary trait values (normal, primary tumor, metastatic). Gene significance

(GS) and modules membership (MM) values were calculated and further used to generate GS vs MM plots to estimate the ion channel association with the binary trait. The co-expressed gene modules were identified as the ones that had a p-value <= 0.05. The soft threshold values interpreted using scale free network topology fit and the minimum module size defined based on the IC_D obtained for individual tumor type in each subgroup identified ion channels is mentioned in the S1 table.

4.6 Determination of correlation of ion channel co-expression networks with EMT

Single sample gene set enrichment analysis (ssGSEA) was performed on the epithelial and mesenchymal gene lists from the transcriptome profiles of tumor types that contained metastatic samples. The gene lists were obtained from a previously published study [35]. We quantified the enrichment of the gene signatures independently using the Gene Set Enrichment Analysis in Python (GSEAPY) library [36]. The enrichment scores were further normalized for the gene sets and used to calculate its correlation with the DE ion channels using Spearman correlation.

4.7 Computation of correlation of co-expressed network specific transcription factors with p-values

The co-expressed IC_D modules were analysed by Enrichr [37] and we determined the transcription factors that might control the ion channel co-expressed networks. The gene set libraries from Enrichr used were: ChIP enrichment analysis (ChEA), Enrichr for protein-protein interactions (PPIs), encyclopedia of DNA elements (ENCODE) transcription factors (TF) ChIP-seq 2015, TRANScription FACtor database (TRANSFAC) and JASPAR position weight matrix (PWMs), ENCODE and ChEA consensus TFs from ChIP-X, and transcriptional regulatory relationships unraveled by sentence-based text mining (TRRUST). The transcription factors with a p-value cut-off of 0.05 and with evidence of presence in

humans were shortlisted. The analysis was updated last on 29th March 2025. Thereafter, Pearson correlation along with t-distribution based p-value was calculated between the expression values of the transcription factors and the WGCNA obtained module eigen genes to associate the co-expressed gene modules with the determined transcription factors.

4.8 ONCOchannelome web implementation

The ONCOchannelome was formulated as a web interface developed using Python (v.3.10.11) scripts based on Django (v.5.0.3) framework. The data generated through the analysis were stored in MySQL relational database management system and can be accessed by the users through the web interface using the link (http://www.oncochannelome.co.in/). The front-end designing of the web interface was accomplished using Hyper-text markup language (HTML), Bootstrap, Cascading Style Sheets (CSS) and Java script (JS).

5. Author contributions

K.T.S.P.: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation. S.R.: Methodology, Formal analysis. M.K.J.: Formal analysis, Writing – review and editing. J.S.: Writing – review & editing, Methodology, Supervision, Resources, Project administration, Investigation, Conceptualization.

6. Acknowledgments

We would like to thank the authors of the manuscripts for making the datasets used in this study publicly available. We thank Harsh Pawar for helpful discussions and critical review. J.S. would like to thank Council of Scientific and Industrial Research (CSIR), Government of India [37WS(0114)/2023-24/EMR-II/ASPIRE] and Indian Council of Medical Research (ICMR), Government of India for the research support. K.T.S.P. was supported by ICMR,

Government of India [BMI/12(95)2021]. J.S. was a recipient of the Bio-CARe Women Scientists award from the Department of Biotechnology (DBT), Government of India.

7. References

1. Furini, S. and C. Domene, *Computational studies of transport in ion channels using metadynamics.* Biochim Biophys Acta, 2016. **1858**(7 Pt B): p. 1733-40.
2. Lobikin, M., et al., *Resting potential, oncogene-induced tumorigenesis, and metastasis: the bioelectric basis of cancer in vivo.* Phys Biol, 2012. **9**(6): p. 065002.
3. Tolon, R.M., et al., *Regulation of somatostatin gene expression by veratridine-induced depolarization in cultured fetal cerebrocortical cells.* Brain Res Mol Brain Res, 1996. **35**(1-2): p. 103-10.
4. Angrisani, A., et al., *The emerging role of the KCTD proteins in cancer.* Cell Commun Signal, 2021. **19**(1): p. 56.
5. Arcangeli, A., *Ion channels and transporters in cancer. 3. Ion channels in the tumor cell-microenvironment cross talk.* Am J Physiol Cell Physiol, 2011. **301**(4): p. C762-71.
6. Gomez-Ospina, N., et al., *The C terminus of the L-type voltage-gated calcium channel Ca(V)1.2 encodes a transcription factor.* Cell, 2006. **127**(3): p. 591-606.
7. Zhang, L., et al., *Pathophysiological role of ion channels and transporters in hepatocellular carcinoma.* Cancer Gene Ther, 2024. **31**(11): p. 1611-1618.
8. Parthasarathi, K.T.S., et al., *Aberrations in ion channels interacting with lipid metabolism and epithelial-mesenchymal transition in esophageal squamous cell carcinoma.* Front Mol Biosci, 2023. **10**: p. 1201459.
9. Parthasarathi, K.T.S., et al., *In Silico Analysis of Ion Channels and Their Correlation with Epithelial to Mesenchymal Transition in Breast Cancer.* Cancers (Basel), 2022. **14**(6).
10. Sharma, J., et al., *Somatic Mutations Profile of a Young Patient With Metastatic Urothelial Carcinoma Reveals Mutations in Genes Involved in Ion Channels.* Front Oncol, 2019. **9**: p. 435.
11. Cordier, C., N. Prevarskaya, and V. Lehen'kyi, *TRPM7 Ion Channel: Oncogenic Roles and Therapeutic Potential in Breast Cancer.* Cancers (Basel), 2021. **13**(24).
12. Jiang, L.H., E. Adinolfi, and S. Roger, *Editorial: Ion Channel Signalling in Cancer: From Molecular Mechanisms to Therapeutics*. Front Pharmacol, 2021. **12**: p. 711593.
13. Huber, S.M., *Oncochannels.* Cell Calcium, 2013. **53**(4): p. 241-55.
14. Wulff, H. and N.A. Castle, *Therapeutic potential of KCa3.1 blockers: recent advances and promising trends.* Expert Rev Clin Pharmacol, 2010. **3**(3): p. 385-96.
15. Fuessel, S., et al., *Vaccination of hormone-refractory prostate cancer patients with peptide cocktail-loaded dendritic cells: results of a phase I clinical trial.* Prostate, 2006. **66**(8): p. 811-21.
16. Hartung, F., W. Stuhmer, and L.A. Pardo, *Tumor cell-selective apoptosis induction through targeting of K(V)10.1 via bifunctional TRAIL antibody.* Mol Cancer, 2011. **10**: p. 109.
17. Mamelak, A.N. and D.B. Jacoby, *Targeted delivery of antitumoral therapy to glioma and other malignancies with synthetic chlorotoxin (TM-601).* Expert Opin Drug Deliv, 2007. **4**(2): p. 175-86.
18. Bharadwaj, R., et al., *Targeting Solute Carrier Transporters (SLCs) as a Therapeutic Target in Different Cancers.* Diseases, 2024. **12**(3).

19. Ko, J.H., et al., *Expression profiling of ion channel genes predicts clinical outcome in breast cancer.* Mol Cancer, 2013. **12**(1): p. 106.
20. DeMarco, K.R., S. Bekker, and I. Vorobyov, *Challenges and advances in atomistic simulations of potassium and sodium ion channel gating and permeation.* J Physiol, 2019. **597**(3): p. 679-698.
21. Carvalho-de-Souza, J.L., Saponaro, A., Bassetto, C. A. Z., Rauh, O., Schroeder, I., Franciolini, F., Moroni, A., *Experimental challenges in ion channel research: uncovering basic principles of permeation and gating in potassium channels.* Advances in Physics: X, 7(1), 2021.
22. Jardin, I., et al., *TRPC6 channel and its implications in breast cancer: an overview.* Biochim Biophys Acta Mol Cell Res, 2020. **1867**(12): p. 118828.
23. Sharma, J., et al., *Editorial: Multi-Omics Approaches to Study Signaling Pathways.* Front Bioeng Biotechnol, 2020. **8**: p. 829.
24. Harmar, A.J., et al., *IUPHAR-DB: the IUPHAR database of G protein-coupled receptors and ion channels.* Nucleic Acids Res, 2009. **37**(Database issue): p. D680-5.
25. Castro, E.V., et al., *ChanFAD: A Functional Annotation Database for Ion Channels.* Front Bioinform, 2022. **2**: p. 835805.
26. Ranjan, R., et al., *Channelpedia: an integrative and interactive database for ion channels*. Front Neuroinform, 2011. **5**: p. 36.
27. Van den Eynde, C., et al., *TRP channel expression correlates with the epithelial-mesenchymal transition and high-risk endometrial carcinoma.* Cell Mol Life Sci, 2021. **79**(1): p. 26.
28. Rapetti-Mauss, R., et al., *Interplay Between Ion Channels and the Wnt/beta-Catenin Signaling Pathway in Cancers.* Front Pharmacol, 2020. **11**: p. 525020.
29. Son, H.J., et al., *Tight Junction-Related CLDN5 and CLDN6 Genes, and Gap Junction-Related GJB6 and GJB7 Genes Are Somatically Mutated in Gastric and Colorectal Cancers.* Pathol Oncol Res, 2020. **26**(3): p. 1983-1987.
30. Ritchie, M.E., et al., *limma powers differential expression analyses for RNA-sequencing and microarray studies*. Nucleic Acids Res, 2015. **43**(7): p. e47.
31. Gautier, L., et al., *affy--analysis of Affymetrix GeneChip data at the probe level.* Bioinformatics, 2004. **20**(3): p. 307-15.
32. Leek, J.T., et al., *The sva package for removing batch effects and other unwanted variation in high-throughput experiments*. Bioinformatics, 2012. **28**(6): p. 882-3.
33. Yu, G., et al., *clusterProfiler: an R package for comparing biological themes among gene clusters*. OMICS, 2012. **16**(5): p. 284-7.
34. Langfelder, P. and S. Horvath, *WGCNA: an R package for weighted correlation network analysis.* BMC Bioinformatics, 2008. **9**: p. 559.
35. Tuan Zea Tan, Q.H.M., Yoshio Miki, Tetsuo Noda, Seiichi Mori, Ruby Yun□Ju Huang, and Jean Paul Thiery, *Epithelial□mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients.* EMBO Mol Med, 2014.
36. Fang, Z., X. Liu, and G. Peltz, *GSEApy: a comprehensive package for performing gene set enrichment analysis in Python.* Bioinformatics, 2023. **39**(1).
37. Kuleshov, M.V., et al., *Enrichr: a comprehensive gene set enrichment analysis web server 2016 update.* Nucleic Acids Res, 2016. **44**(W1): p. W90-7.

8. Supporting information captions

S1 table: Parameters used in WGCNA analysis of ion channels in TN, MT and MN across tumors

S1 Fig: PCA plot to identify batch effects in tumor types with normal, tumor and metastatic samples available

S2 Fig: PCA plot to identify batch effects in tumor types with normal and tumor samples available

9. Data availability statement

The ONCOchannelome is available at 'http://www.oncochannelome.co.in/'.