

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property

Organization

International Bureau

(43) International Publication Date

08 August 2019 (08.08.2019)



(10) International Publication Number

WO 2019/149706 A1

(51) International Patent Classification:

C12N 5/07 (2010.01)

C12Q 1/6886 (2018.01)

G01N 33/50 (2006.01)

Published:

— with international search report (Art. 21(3))

— with sequence listing part of description (Rule 5.2(a))

(21) International Application Number:

PCT/EP2019/052149

(22) International Filing Date:

29 January 2019 (29.01.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

1801495.1

30 January 2018 (30.01.2018)

GB

(71) Applicant: **THE QUEEN'S UNIVERSITY OF BELFAST** [GB/GB]; University Road, Belfast Antrim BT7 INN (GB).

(72) Inventor: **CAMPBELL, Frederick Charles**; 12 Rosemary Park, Belfast Antrim BT96RF (GB).

(74) Agent: **FRKELLY**; 27 Clyde Road, Dublin D04 F838 (IE).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: A MODEL OF LOW- AND HIGH- GRADE CANCER

(57) Abstract: The present invention relates to the field of cancer research and in particular provides novel tools for cancer research and drug and biomarker discovery. Specifically, the present invention relates to models of low- and high- grade cancer, specifically adenocarcinoma. In particular, the present invention concerns Caco-2 cells, which have been modified to reproduce the events and cellular morphology associated with low- or high-grade colorectal adenocarcinoma in a controllable manner. In particular, the present invention concerns modified Caco-2 cells, which have been engineered to be defective in centrosome anchoring to the cell cortex, clustering of interphase centrosomes and having supernumerary centrosomes. The invention concerns modified Caco-2 cells, cancer model systems comprising modified Caco-2 cells and polynucleotides, expression vectors comprising said polynucleotides and methods of producing the cells and cancer model systems.



WO 2019/149706 A1

A model of low and high grade cancer

Field of the Invention

The present invention relates to models of low- and high- grade cancer. In particular, the present invention concerns Caco-2 cells, which have been modified to reproduce the events and cellular morphology associated with low- or high-grade colorectal cancer. In particular, the present invention concerns modified Caco-2 cells, which have been engineered to be defective in centrosome anchoring to the cell cortex, clustering of interphase centrosomes and having supernumerary centrosomes. The invention concerns modified Caco-2 cells, cell cultures comprising modified Caco-2 cells and methods of producing the cells and cell cultures.

Background to the Invention

Colorectal cancer (CRC) may represent the major cancer challenge of the 21st century because it is the third most lethal global malignancy and its incidence is expected to increase by 60% over the next two decades. Growth of CRC ranges from indolent to highly aggressive. Prognostic stratification is aided by histological grading but aggressive CRC is characterized by chromosome segregation error as well as high-grade morphology. Chromosome partitioning is coupled to multicellular morphology by mitotic apparatus and relevant interplay may be dissected in organotypic culture models.

Ordered cell division maintains the epithelial barrier in the healthy colon. In preparation for mitosis, cells copy their genome and remodel their internal architecture to enable mitotic spindle assembly. In Caco-2 CRC cells, redistribution of the cytoskeletal linker protein ezrin to form a cap-like accumulation at one pole of the cell cortex provides a cue for astral microtubule (MT) capture and stabilization of the interphase centrosome. Thus anchored, the centrosome normally replicates to generate one mother and one daughter. Centrosome anchoring to the cell cortex is necessary for separation of mother and daughter centrosomes, construction and orientation of the mitotic spindle, formation of cell shape and multicellular assembly. The ezrin cap stabilizes the normal centrosome, and it also anchors and “clusters” extra centrosomes during interphase. Centrosome amplification characterizes many human cancers. Effective clustering of extra centrosomes during interphase enables assembly of a bipolar mitotic spindle, error-free segregation of a diploid chromosome complement and normal multicellular pattern formation. Conversely, ineffective clustering of interphase centrosomes can activate failsafe processes that cluster extra centrosomes later in the cell cycle, during metaphase. However, these metaphase centrosome clustering processes invoke substantive segregation error.

Molecular controls of ezrin spatiotemporal dynamics are unclear. The polarity regulator protein kinase C zeta (PKCz) phosphorylates ezrin to initiate embryonic morphogenesis. PKCz also controls centrosome positioning, orientated mitosis, chromosome segregation and multicellular assembly.

5 In this study, the inventors have dissected PKCz regulation of ezrin interactions with its known binding partner NHERF1 (also known as ezrin binding protein 50 [EBP50] or Solute carrier family 9, Sodium/Hydrogen exchanger, isoform 3, regulator 1 [SLC9A3R1]) that is important for maintenance of ezrin at the cell cortex. The inventors have also investigated PKCz regulation of merlin that is known to be involved in ezrin cap formation. The inventors
10 have shown that perturbation of ezrin cap formation alone or in combination with centrosome amplification drives evolution of phenotypes evocative of low- or high-grade CRC, in 3D organotypic culture models.

PKCz controls mitotic spindle dynamics, chromosome segregation and multicellular patterns but its role in CRC phenotype evolution remains unclear.

15

Until now, a lack of understanding as to the key molecular regulators of centrosome number and anchoring in the cell cycle has prevented the generation of cell-based models of low- and high-grade cancer which faithfully reproduce the genomic and multicellular attributes of low- or high-grade cancer respectively, in particular in relation to colorectal cancer. As a
20 consequence, despite the existence of a pressing need for such models as research tools for conducting mechanistic studies and drug screening programmes, until now they have remained elusive.

Summary of the Invention

The invention relates to the development of low- and high-grade cancer models. In
25 particular, the invention relates to the development of low- and high-grade colorectal cancer models. The invention relates to the development of low- and high-grade colorectal adenocarcinoma models.

The invention is based on the discovery of the molecular and mechanistic basis of defective anchoring of interphase centrosomes in low- or high-grade colorectal cancer types.

30 In particular the inventors have now shown that PKCz couples genome segregation to multicellular morphology through control of interphase centrosome anchoring. In this study the inventors have shown that PKCz regulates interdependent processes that control centrosome positioning. Among these, interaction between the cytoskeletal linker protein ezrin and its binding partner NHERF1 promotes formation of a localized cue for anchoring

interphase centrosomes to the cell cortex. In this study the inventors have shown that perturbation of these phenomena induced different outcomes in cells with single or extra centrosomes. Defective anchoring of a single centrosome promoted bipolar spindle misorientation, multi-lumen formation and aberrant epithelial stratification. Collectively, these disturbances induce cribriform multicellular morphology that is typical of some categories of low-grade CRC. By contrast, defective anchoring of extra centrosomes promoted multipolar spindle formation, chromosomal instability (CIN), disruption of glandular morphology and cell outgrowth across the extracellular matrix interface characteristic of aggressive, high-grade CRC. The inventors have shown that apical NHERF1 IHC intensity inversely associated with multipolar spindle frequency and high-grade morphology in formalin fixed human CRC samples. The discovery that defective PKC ζ control of interphase centrosome anchoring underlies distinct categories of mitotic slippage that shape development of low- or high-grade CRC phenotypes has now allowed the development of a cell-based model for LG and HG CRC phenotypes.

These discoveries lay the technical foundations for the development of two types of cells which reproduce the molecular and morphological attributes of low- or high-grade cancer respectively. This has never been achieved previously.

The inventors have also demonstrated that usefully, the modified cells can form the basis of multi-cellular models which reproduce the combined genomic and multicellular attributes of low- or high-grade cancer, in particular low-grade adenocarcinoma or high-grade adenocarcinoma. The cells are particularly useful in the study of colorectal adenocarcinoma and in research for diagnostic or prognostic markers and treatments thereof. In particular, the inventors have developed a low-grade system which shows combined features of cribriform morphology with normal chromosome segregation. Common features of low-grade cancer include preservation of glandular architecture. This low-grade model is characterised by some level of preservation of glandular architecture and may include multiple lumens, stratified epithelium and cribriform morphology.

In contrast, features of high-grade cancer include breakdown or loss of glandular architecture, nuclear and cellular pleomorphism, atypical mitotic figures and extension of cells into extracellular matrix. Although not an essential feature of high-grade cancer, chromosomal instability typically occurs in high-grade cancer. Here the inventors have additionally developed a high-grade system, which reproduces all of these attributes, for the first time. The inventors have therefore developed a high-grade system which for the first time combines many features of human high-grade cancer e.g. chromosomal instability, aberrant mitotic figures, nuclear and cellular pleomorphism and cell penetration into extracellular matrix. This high-grade model is characterised by severe perturbation of

glandular architecture, severe cellular and nuclear pleomorphism, malignant cell extension into extracellular matrix and chromosomal instability.

Of particular note is the development of a Caco-2 model of high-grade cancer model in which it is possible to induce micronuclei formation, and specifically to bring about errors in genome partitioning including CIN and chromosome missegregation into micronuclei in cells with extra centrosomes.

An understanding of the key molecular regulators of centrosome number and anchoring in the cell cycle has enabled the generation of cell-based models which faithfully recapitulate the events associated with low- and high grade colorectal cancer. This allows both grades of colorectal cancer to be interrogated (and compared) at the molecular level in a way which was not previously possible and thus allows the underlying causes to be studied in more detail.

The modified cells described herein and the models based thereon fill an important gap in scientific methods and for the first time make rapid mechanistic assessment of treatment effects on low or high-grade cancer outcomes feasible without the need for recourse to expensive, non-mechanistic patient-derived xenograft studies or large scale clinical trials.

Consequently, the cells usefully find application as tools to more efficiently and accurately conduct mechanistic studies of both low- and high-grade cancer, in particular colorectal cancer, at the individual cell level and when provided as part of 2- (2D) or 3-dimensional (3D) models and will form the basis of much needed drug screening programmes for the treatment of low- or high-grade cancer, especially colorectal cancer (CRC), without the practical limitations associated with currently available technologies. The models are cell-based and therefore permit drug screening programmes to be high-throughput. Additionally, the cells which form the basis of the models may be studied individually or as part of a cell culture, depending on the chosen application. Cell cultures can either be 2D or 3D, which allows the characteristic aspects of the cells and the effects of various interventions thereon to be studied in an environment which more closely approximates that of cells *in vivo*.

The low- and high-grade colorectal cancer models of the present invention are based on human caucasian colon adenocarcinoma cells, i.e. Caco-2 cells, which have been modified so as to reproduce the events associated with low- or high-grade colorectal cancer respectively. Conveniently, the invention makes use of commercially available Caco-2 cells as a starting point. These cells are widely available and the required culture conditions are well characterized and have been used previously by many other researchers. The starting cells used in the examples have been validated to confirm that they were unmodified Caco-2 cells, without any genetic or cellular modification whatsoever, prior to their modification using

methods, polynucleotides and expression vectors described herein in order to produce modified Caco-2 cells of the invention.

Accordingly, the present invention provides a modified Caco-2 cell, having defective centrosome anchoring relative to an equivalent, unmodified Caco-2 cell.

- 5 The present invention provides a modified Caco-2 cell, which is defective in or incapable of centrosome anchoring to the cell cortex when compared to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having ineffective clustering of interphase centrosomes relative to an equivalent, unmodified Caco-2 cell.

- 10 The present invention provides a modified Caco-2 cell, having defective centrosome anchoring relative to an equivalent, unmodified Caco-2 cell and having supernumerary centrosomes (at a given point in the cell cycle e.g. interphase) relative to an equivalent, unmodified Caco-2 cell. The modified Caco-2 cells have extra centrosomes that are not anchored to one region of the circumference of the cell (i.e. the cell cortex) during
15 interphase. Extra centrosomes are recognized by the presence of more than 2 centrosomes in the cell at any time. This situation contrasts with the normal situation, in which a cell has only one centrosome during early interphase and the centrosome duplicates during interphase to form a total of 2 centrosomes (one mother and one daughter centrosome). Defective anchoring of extra centrosomes is recognisable by dispersal of the centrosomes
20 (rather than confinement of the centrosomes to one location) and subsequent assembly of a multipolar mitotic spindle (i.e. the spindle has more than 2 poles). This last feature contrasts with the normal (i.e. non-cancerous) situation where the spindle has only 2 poles (i.e. is bipolar).

- The inventors have surprisingly discovered that reduction in the levels or activity of Protein
25 Kinase C-zeta (PKCz) and/or phosphatase and tensin homologue deleted on chromosome 10 (PTEN) and/or Na⁺/H⁺ exchanger regulatory factor-1 (NHERF1), for example in a Caco-2 cell background, results in changes consistent with low-grade cancer, particularly low-grade adenocarcinoma. The resulting cells can therefore serve as a model of low-grade cancer, particularly low-grade adenocarcinoma and more particularly, low-grade colorectal
30 adenocarcinoma.

Accordingly, the present invention provides a genetically modified Caco-2 cell model, the modified Caco-2 cell comprising a heterologous polynucleotide capable of hybridising to a polynucleotide encoding PKCz or PTEN or NHERF1 such that expression of said heterologous polynucleotide results in a reduction in, or elimination of, expression or level or

activity of native PKCz or PTEN or NHERF1 polypeptide in the modified Caco-2 cell compared to an equivalent, unmodified cell.

Polynucleotides encoding PKCz, PTEN, NHERF1, PLK4 and Aurora A may be DNA or RNA. In particular, cDNA reference sequences are provided which encode PKCz (SEQ ID NO: 1),
 5 PTEN (SEQ ID NO: 2), NHERF1 (SEQ ID NO: 3), PLK4 (SEQ ID NO: 4) and Aurora A (SEQ ID NO: 5). It will be understood that where a polynucleotide is required that is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NOs: 1-5 (for example in order to reduce transcript levels, downregulate levels or activity of the encoded protein, or as a probe to measure expression at the mRNA level), then the mRNA
 10 sequence may be generated by converting the relevant reference cDNA sequence, using a conversion tool that substitutes uracils for thiamines, which is widely available.

The present invention also provides a method of generating a cancer cell, comprising the steps of:

- a. providing a Caco-2 cell; and
- 15 b. reducing the expression or activity of PKCz and/or PTEN and/or NHERF1 polypeptide in said Caco-2 cell, relative to that of an equivalent, unmodified Caco-2 cell.

The present invention also provides a method of generating a low-grade cancer cell,
 20 comprising the steps of:

- a. providing a Caco-2 cell; and
- b. reducing the expression or activity of PKCz and/or PTEN and/or NHERF1 polypeptide in said Caco-2 cell, relative to that of an equivalent, unmodified Caco-2 cell.

25 Preferably the low-grade cancer cell is a low-grade adenocarcinoma cell. More preferably the low-grade cancer cell is a low-grade colorectal adenocarcinoma cell.

The present invention also provides a method of generating a cancer cell, preferably an adenocarcinoma cell, more preferably a colorectal adenocarcinoma cell, even more
 30 preferably a low-grade adenocarcinoma cell, still more preferably a low-grade colorectal adenocarcinoma cell, wherein a Caco-2 cell is transformed with a heterologous polynucleotide capable of hybridising to a polynucleotide encoding PKCz or PTEN or NHERF1 such that there is a reduction in, or elimination of, expression or level or activity of

native PKCz or PTEN or NHERF1 polypeptide in said Caco-2 cell compared to an untransformed control Caco-2 cell. Optionally, the Caco-2 cell may be transformed with a polynucleotide capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1 and/or a polynucleotide capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2 and/or a polynucleotide capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3.

The present invention also provides a method of generating a cancer cell, which method comprises; reducing the levels or activity of an endogenous PKCz and/or PTEN and/or NHERF1 protein within a Caco-2 cell such that the resulting modified Caco-2 cell, has reduced levels or activity of PKCz and/or PTEN and/or NHERF1 relative to an equivalent, unmodified Caco-2 cell. The cancer cell is preferably an adenocarcinoma cell, in particular a low-grade adenocarcinoma cell. More preferably the cancer cell is a colorectal adenocarcinoma cell, in particular a low-grade adenocarcinoma cell. Still more preferably the cancer cell is a low-grade colorectal adenocarcinoma cell.

The present invention also provides a genetically modified Caco-2 cell, the modified Caco-2 cell comprising a polynucleotide expression construct, wherein the construct comprises a polynucleotide sequence which is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1, and/or a polynucleotide sequence which is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2, and/or polynucleotide sequence which is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3, and which upon expression in a host Caco-2 cell results in a reduction in, or elimination of, the expression level or activity of Protein Kinase C-zeta (PKCz) and/or phosphatase and tensin homologue deleted on chromosome 10 (PTEN) and/or Na⁺/H⁺ exchanger regulatory factor-1 (NHERF1) polypeptide respectively in the modified Caco-2 cell, relative to an equivalent, unmodified Caco-2 cell.

Preferably, the expression of the polynucleotide sequence which is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1, and/or that which is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2, and/or that which is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3; is under the control of an inducer molecule.

The present invention provides a modified Caco-2 cell, having reduced levels or activity of Protein Kinase C-zeta (PKCz) and/or phosphatase and tensin homologue deleted on chromosome 10 (PTEN) and/or Na⁺/H⁺ exchanger regulatory factor-1 (NHERF1) relative to an equivalent, unmodified Caco-2 cell.

- 5 The present invention provides a modified Caco-2 cell, having reduced levels or activity of PKCz relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having reduced levels or activity of PKCz and PTEN relative to an equivalent, unmodified Caco-2 cell.

- 10 The present invention provides a modified Caco-2 cell, having reduced levels or activity of PKCz and NHERF1 relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having reduced levels or activity of PTEN and NHERF1 relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having reduced levels or activity of PKCz and PTEN and NHERF1 relative to an equivalent, unmodified Caco-2 cell.

15

- The inventors have also discovered that a reduction in the levels or activity of Protein Kinase C-zeta (PKCz) and/or phosphatase and tensin homologue deleted on chromosome 10 (PTEN) and/or Na⁺/H⁺ exchanger regulatory factor-1 (NHERF1), for example in a Caco-2 cell background, results in low-grade cancer. Independently, an increase in the levels or activity of Polo Like Kinase 4 (PLK4) and/or Aurora A, for example in a Caco-2 cell background can affect centrosome replication. Surprisingly, however, a reduction in the levels or activity of Protein Kinase C-zeta (PKCz) and/or phosphatase and tensin homologue deleted on chromosome 10 (PTEN) and/or Na⁺/H⁺ exchanger regulatory factor-1 (NHERF1), *in combination* with increased levels or activity of Polo Like Kinase 4 (PLK4) and/or Aurora A for example in a Caco-2 cell background, compared with the levels or activities of the corresponding polypeptides in an unmodified Caco-2 cell, results in high-grade cancer. Advantageously, the resulting cells (i.e. those having reduced levels or activities of PKCz and/or PTEN and/or NHERF1, in conjunction with increased levels or activities of PLK4 and/or Aurora A; compared with the levels or activities of their counterparts in equivalent, unmodified Caco-2 cells) can therefore serve as a model of high-grade cancer, in particular high-grade adenocarcinoma. The resulting cells can therefore serve as a model of high-grade colorectal adenocarcinoma.
- 20
- 25
- 30

The present invention therefore provides a modified Caco-2 cell, having reduced levels or activity of PKC ζ and/or PTEN and/or NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having increased levels or activity of PLK4 and/or Aurora A relative to an equivalent, unmodified Caco-2 cell.

5

The following represent preferred combinations:

The present invention provides a modified Caco-2 cell, having a reduced level or activity of PKC ζ relative to an equivalent, unmodified Caco-2 cell; and having an increased level or activity of PLK4 relative to an equivalent, unmodified Caco-2 cell.

- 10 The present invention provides a modified Caco-2 cell, having a reduced level or activity of PTEN relative to an equivalent, unmodified Caco-2 cell; and having an increased level or activity of PLK4 relative to an equivalent, unmodified Caco-2 cell.

- The present invention provides a modified Caco-2 cell, having a reduced level or activity of NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having an increased level or
15 activity of PLK4 relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having reduced levels or activities of PKC ζ and PTEN relative to an equivalent, unmodified Caco-2 cell; and having an increased level or activity of PLK4 relative to an equivalent, unmodified Caco-2 cell.

- The present invention provides a modified Caco-2 cell, having reduced levels or activities of
20 PKC ζ and NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having an increased level or activity of PLK4 relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having reduced levels or activities of PTEN and NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having an increased level or activity of PLK4 relative to an equivalent, unmodified Caco-2 cell.

- 25 The present invention provides a modified caco-2 cell, having reduced levels or activities of PKC ζ and PTEN and NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having an increased level or activity of PLK4 relative to an equivalent, unmodified Caco-2 cell.

- The present invention provides a modified Caco-2 cell, having a reduced level or activity of PKC ζ relative to an equivalent, unmodified Caco-2 cell; and having an increased level or
30 activity of Aurora A relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having a reduced level or activity of PTEN relative to an equivalent, unmodified Caco-2 cell; and having an increased level or activity of Aurora A relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having a reduced level or activity of NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having an increased level or activity of Aurora A relative to an equivalent, unmodified Caco-2 cell.

5 The present invention provides a modified Caco-2 cell, having reduced levels or activities of PKCz and PTEN relative to an equivalent, unmodified Caco-2 cell; and having an increased level or activity of Aurora A relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having reduced levels or activities of PKCz and NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having an increased level or activity of Aurora A relative to an equivalent, unmodified Caco-2 cell.

10 The present invention provides a modified Caco-2 cell, having reduced levels or activities of PTEN and NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having an increased level or activity of Aurora A relative to an equivalent, unmodified caco-2 cell.

The present invention provides a modified Caco-2 cell, having reduced levels or activities of PKCz and PTEN and NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having
15 an increased level or activity of Aurora A relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having a reduced level or activity of PKCz relative to an equivalent, unmodified Caco-2 cell; and having increased levels or activities of PLK4 and Aurora A relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having a reduced level or activity of
20 PTEN relative to an equivalent, unmodified Caco-2 cell; and having increased levels or activities of PLK4 and Aurora A relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having a reduced level or activity of NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having increased levels or activities of PLK4 and Aurora A relative to an equivalent, unmodified Caco-2 cell.

25 The present invention provides a modified Caco-2 cell, having reduced levels or activities of PKCz and PTEN relative to an equivalent, unmodified caco-2 cell; and having increased levels or activities of PLK4 and Aurora A relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having reduced levels or activities of PKCz and NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having increased
30 levels or activities of PLK4 and Aurora A relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having reduced levels or activities of PTEN and NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having increased levels or activities of PLK4 and Aurora A relative to an equivalent, unmodified Caco-2 cell.

The present invention provides a modified Caco-2 cell, having reduced levels or activities of PKCz and PTEN and NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having increased levels or activities of PLK4 and Aurora A relative to an equivalent, unmodified Caco-2 cell.

- 5 The present invention provides a modified Caco-2 cell, having reduced levels or activities of PKCz and PTEN and NHERF1 relative to an equivalent, unmodified Caco-2 cell; and having increased levels or activities of PLK4 and Aurora A relative to an equivalent, unmodified Caco-2 cell.

10 The present invention provides a modified Caco-2 cell, having a reduced activity of PKCz polypeptide relative to an equivalent, unmodified Caco-2 cell; and increased expression levels of PLK4 polypeptide compared to an unmodified Caco-2 control cell.

The present invention also provides a genetically modified Caco-2 cell, the modified Caco-2 cell comprising a polynucleotide expression construct, wherein the construct comprises a polynucleotide sequence which is complementary to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1, and/or a polynucleotide sequence which is complementary to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2, and/or a polynucleotide sequence which is complementary to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3, and which upon expression results in a reduction in the expression level or activity of Protein Kinase C-zeta (PKCz) and/or phosphatase and tensin homologue deleted on chromosome 10 (PTEN) and/or Na⁺/H⁺ exchanger regulatory factor-1 (NHERF1) polypeptide respectively in the modified Caco-2 cell, relative to an equivalent, unmodified Caco-2 cell; and wherein the modified Caco-2 cell further comprises a polynucleotide expression construct, wherein the construct comprises a polynucleotide sequence of SEQ ID NO: 4 or a sequence of at least 90% identity thereto, and/or SEQ ID NO: 5 or a sequence of at least 90% identity thereto, and which upon expression results in an increase in the expression level or activity of Polo Like Kinase 4 (PLK4) and/or Aurora A polypeptide respectively, in the modified Caco-2 cell, relative to an equivalent, unmodified Caco-2 cell.

30 Preferably, the expression of the polynucleotide sequence comprising SEQ ID NO: 4 or a sequence of at least 90% identity thereto, and/or SEQ ID NO: 5 or a sequence of at least 90% identity thereto, is under the control of an inducer molecule.

The present invention also provides an expression vector comprising;

- 35 a. a polynucleotide sequence which is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1 or a fragment thereof; and/or a

polynucleotide sequence which is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2 or a fragment thereof; and/or a polynucleotide sequence which is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3 or a fragment thereof; and

- 5 b. a polynucleotide sequence having a sequence of SEQ ID NO: 4 or a sequence of at least 90% thereto; and/or a sequence of SEQ ID NO: 5 or a sequence of at least 90% thereto;

wherein said expression vector is capable of expression in a Caco-2 cell, and wherein upon expression in a Caco-2 cell results in a decrease in the expression of PKCz and/or PTEN
10 and/or NHERF1 and an increase in the expression of PLK4 and/or Aurora A in said caco-2 cell relative to that of an unmodified control Caco-2 cell.

Preferably the present invention provides an expression vector comprising;

- a. a polynucleotide sequence which is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1 or a fragment thereof; and
15 b. a polynucleotide sequence having a sequence of SEQ ID NO: 4 or a sequence of at least 90% thereto;

wherein said expression vector is capable of expression in a Caco-2 cell, and wherein upon expression in a Caco-2 cell results in a decrease in the expression of PKCz and an increase in the expression of PLK4 in said Caco-2 cell relative to that of an unmodified control Caco-2
20 cell.

The present invention also provides a method of generating a high-grade cancer cell, comprising the steps of:

- a. providing a Caco-2 cell; and
25 b. reducing the expression or activity of PKCz and/or PTEN and/or NHERF1 polypeptide in said Caco-2 cell, relative to that of an equivalent, unmodified Caco-2 cell; and
c. increasing the expression or activity of PLK4 and/or Aurora A polypeptide in the Caco-2 cell, relative to that of an equivalent, unmodified Caco-2 cell.

30 The resulting cell is preferably a high-grade cancer cell, in particular a high-grade adenocarcinoma cell. More preferably, the cancer cell is a high-grade colorectal adenocarcinoma cell, i.e. having the characteristic morphology, intracellular architecture and/or gene expression characteristics of a high-grade colorectal adenocarcinoma cell.

The modification to induce ineffective clustering of interphase centrosomes involves 2 steps (b and c). Suitably, steps b and c can be performed consecutively or simultaneously. Where steps b and c are performed consecutively, they can be performed in any order. It will be appreciated, that, in some applications of the method, where, for example, a cell exhibiting the characteristic features of low-grade adenocarcinoma is required first of all, (for instance to examine the differences before and after a switch from normal to low-grade cancer and/or from low-grade to high-grade cancer) then the step(s) of reducing the expression or activity of PKCz and/or PTEN and/or NHERF1 polypeptide in said Caco-2 cell, relative to that of an equivalent, unmodified Caco-2 cell may be carried out first and then followed by step c. Alternatively, where it is not desired to make use of the ability to switch between low-grade and high-grade adenocarcinoma, for example, where a high-grade cell only is required, then steps b and c may be carried out simultaneously, or the order of steps b and c may be reversed; initially performing step c to generate extra centrosomes before inducing defective anchoring of the centrosomes by carrying out step b.

In preferred embodiment, where a high-grade cell is required, firstly, the starting cells (i.e. Caco-2 cells) are modified to have supernumerary centrosomes (e.g. by overexpression of PLK4) (corresponding to step c). Subsequently, anchoring of centrosomes is inhibited by inhibition or knockdown of PKCz (and/or NHERF1 and/or PTEN)(corresponding to step b). When these 2 steps have been carried out, the resulting cells have extra centrosomes that are not anchored to one region of the circumference of the cell (i.e. the cell cortex) during interphase. Defective anchoring of extra centrosomes is recognisable by dispersal of the centrosomes (rather than confinement of the centrosomes to one location) and subsequent assembly of a multipolar mitotic spindle (i.e. the spindle has more than 2 poles). This last feature contrasts with the normal (i.e. non-cancerous) situation where the spindle has only 2 poles (i.e. is bipolar).

The present invention also provides a method of generating a high-grade cancer cell, which method comprises; reducing the levels or activity of an endogenous PKCz and/or PTEN and/or NHERF1 protein within a Caco-2 cell such that the resulting modified Caco-2 cell has reduced levels or activity of PKCz and/or PTEN and/or NHERF1 relative to an equivalent, unmodified Caco-2 cell; wherein, the method further comprises transforming the Caco-2 cell with a recombinant polynucleotide capable of expressing PLK4 and/or a recombinant polynucleotide capable of expressing Aurora A; such that upon expression of the recombinant polynucleotide in the Caco-2 cell, the resulting modified Caco-2 cell has increased levels or activity of PLK4 and/or Aurora A relative to an equivalent, unmodified Caco-2 cell. The resulting cell is preferably a high-grade adenocarcinoma cell. More preferably, the cancer cell is a high-grade colorectal adenocarcinoma cell.

The present invention also provides a method of generating a high-grade cancer cell, preferably a high-grade adenocarcinoma cell, in particular a high-grade colorectal adenocarcinoma cell, wherein a Caco-2 cell is transformed with a heterologous polynucleotide capable of hybridising to a polynucleotide encoding PKC ζ or PTEN or NHERF1 such that there is a reduction in, or elimination of, expression or level or activity of native PKC ζ or PTEN or NHERF1 polypeptide in said Caco-2 cell compared to an untransformed control Caco-2 cell; and further comprising increasing the expression or activity of PLK4 and/or Aurora A polypeptide in the Caco-2 cell, relative to that of an equivalent, unmodified Caco-2 cell. Optionally, the Caco-2 cell may be transformed with a polynucleotide capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1 or a fragment thereof and/or a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2 or a fragment thereof and/or a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3 or a fragment thereof. Optionally, additionally, the Caco-2 cell may be transformed with a polynucleotide of SEQ ID NO: 4 or a sequence of at least 90% identity thereto or SEQ ID NO: 5 or a sequence of at least 90% identity thereto. The resulting cell is preferably a high-grade adenocarcinoma cell. More preferably, the cancer cell is a high-grade colorectal adenocarcinoma cell.

Cell cultures

Usefully, modified Caco-2 cells described herein may be provided as a cell culture. The characteristic aspects of the cells and the effects of various interventions thereon can therefore be studied in an environment which more closely approximates that of cells *in vivo*.

The cell culture may usefully be grown *in vitro* as a monolayer, for example on an artificial substrate bathed in nutrient medium. This arrangement may be desirable in some circumstances, for instance, where analysis of aspects of low- or high-grade cancer at the cellular, sub-cellular or molecular level are prioritised, for example for performing a high-throughput gene expression assay or drug screening programme.

In order to more faithfully replicate *in vivo* cell growth, behaviour, proliferation and interaction within a tissue, the modified Caco-2 cells described herein may advantageously be grown in a three-dimensional (3D) system, i.e. an artificially created environment in which cells are permitted to grow or interact with their surroundings in all three dimensions. Unlike 2D environments (e.g. a petri dish), such a 3D cell culture allows cells *in vitro* to grow in all directions, in a manner which more closely imitates *in vivo* growth. Consequently, when

grown in this three-dimensional system, the proliferating cells mature and segregate properly to form components of adult tissues analogous to their counterparts *in vivo*. This arrangement allows the events that take place in low- or high-grade adenocarcinoma, and particularly colorectal adenocarcinoma and/or the effects of various interventions (e.g. drugs) to be more accurately modelled and their effects *in vivo* to be more accurately predicted.

Such a 3D system may be achieved by methods which are well known in the art. Typically this can include provision of a 3D framework (e.g. a support structure, or matrix) on which the cells may be grown. The framework may be composed of a living material formed into a three dimensional structure. The framework may be composed of a non-living material formed into a three dimensional structure. The framework may be composed of a combination of living and non-living materials formed into a three dimensional structure. The three-dimensionality of such a framework allows for a spatial distribution which more closely approximates conditions *in vivo*, thus allowing for the formation of microenvironments conducive to cellular maturation, growth, interaction and migration. Modified Caco-2 cells described herein may be inoculated and grown on such a 3D framework.

The growth of cells in the presence of a 3D framework may be further enhanced by the addition of factors required for, or beneficial to growth and/or proliferation, e.g. growth factors and/or factors required for cell adhesion, anchorage, attachment, differentiation or proliferation. Depending on the chosen application, these additional factors may suitably be inoculated upon the 3D framework, or provided by coating the framework, by direct addition to the modified Caco-2 cells grown in the 3D system or by addition to a culture medium in which the modified Caco-2 cells are grown or by any combination thereof, such that the framework becomes populated with viable modified Caco-2 cells to form a 3D tissue structure. In addition to application of the factors to the framework itself or to the initial culture, it will be appreciated that during growth *in vitro* it may be required to additionally supplement the modified Caco-2 cells described herein, with additional factors as described above for instance to maintain or support cell growth. Additionally or alternatively, it may be desired to apply various compounds to the cells either to test their effects, for instance as cancer therapeutics or in order to bring about various changes to the cells. One example of the latter situation is the switch from low grade to high-grade adenocarcinoma, which may be brought about by the external application of a compound, such as doxycycline, to trigger an increase in gene expression in the modified Caco-2 cells making up the cell culture, in particular, for example, e.g. PLK4 or Aurora A overexpression.

The 3D framework can be configured, moulded or shaped to assume the conformation of organs and their components. In the case of colorectal adenocarcinoma, the three-dimensional framework can be shaped to assume the conformation of structures making up

the colon or the rectum, for example the intestinal epithelium. In order to more closely approximate the makeup of *in vivo* structures, the 3D framework may also be inoculated with other cells in addition to the modified Caco-2 cells described herein. Preferably, the modified Caco-2 cells are provided as a 3D organotypic culture model.

- 5 Commonly, in the case of 3D cultures, the modified Caco-2 cells as described herein will typically attach to and preferably substantially envelope the framework formed into a three-dimensional structure. When grown in this 3D system, the proliferating cells mature and segregate properly to form components of adult tissues analogous to counterparts found *in vivo*. The living modified Caco-2 tissue culture so formed provides the support, growth
10 factors, and regulatory factors necessary to sustain long-term active proliferation of modified Caco-2 cells in culture in order that aspects of their growth and development can be studied. In particular, aspects of their growth and development can be studied either in the presence or absence of additional factors such as candidate cancer-therapeutic compounds.

- Whether 2D or 3D, the nature of the substrate on which such a cell culture is grown may be
15 made of any suitable material and for example, may be solid, such as plastic (e.g. disposable or biodegradable plastics), or semi-solid, such as semisolid gels (e.g. collagen, agar). The material may be living or non-living or a combination thereof. Preferably, to better mimic *in vivo* environments, cell an extracellular matrix-based hydrogel will be used as a basis for cell cultures.

- 20 In accordance with the invention, a living cell culture comprising modified Caco-2 cells as described herein may be formed.

- Accordingly, the present invention provides a cell culture comprising more than one modified Caco-2 cell as described herein, wherein the modified Caco-2 cells, have defective
25 centrosome anchoring relative to equivalent, unmodified Caco-2 cells.

- The present invention also provides a cell culture comprising more than one modified Caco-2 cell as described herein, wherein the modified Caco-2 cells are defective in or incapable of centrosome anchoring to the cell cortex when compared to equivalent, unmodified Caco-2
30 cells.

The present invention also provides a cell culture comprising more than one modified Caco-2 cell as described herein, wherein the modified Caco-2 cells have ineffective clustering of interphase centrosomes relative to equivalent, unmodified Caco-2 cells.

The present invention also provides a cell culture comprising more than one modified Caco-2 cell as described herein, wherein the modified Caco-2 cells have defective centrosome anchoring relative to equivalent, unmodified Caco-2 cells and have supernumerary centrosomes (at a given point in the cell cycle e.g. interphase) relative to equivalent, unmodified Caco-2 cells. These features are characteristic of high-grade cancer, particularly high-grade adenocarcinoma, especially high-grade colorectal adenocarcinoma. The cells of said cell culture have extra centrosomes that are not anchored to one region of the circumference of the cell (i.e. the cell cortex) during interphase. Extra centrosomes are recognized by the presence of more than 2 centrosomes in the cell at any time. This situation contrasts with the normal situation, in which a cell has only one centrosome during early interphase and the centrosome duplicates during interphase to form a total of 2 centrosomes (one mother and one daughter centrosome). Defective anchoring of extra centrosomes is recognisable by dispersal of the centrosomes (rather than confinement of the centrosomes to one location) and subsequent assembly of a multipolar mitotic spindle (i.e. the spindle has more than 2 poles). This last feature contrasts with the normal (i.e. non-cancerous) situation where the spindle has only 2 poles (i.e. is bipolar).

The present invention also provides a cell culture comprising more than one modified Caco-2 cell, wherein the modified Caco-2 cells have reduced levels or activity of Protein Kinase C-zeta (PKCz) and/or phosphatase and tensin homologue deleted on chromosome 10 (PTEN) and/or Na⁺/H⁺ exchanger regulatory factor-1 (NHERF1) relative to equivalent, unmodified Caco-2 cells.

The present invention also provides a cell culture comprising more than one modified Caco-2 cell, wherein the modified Caco-2 cells have reduced levels or activity of Protein Kinase C-zeta (PKCz) and/or phosphatase and tensin homologue deleted on chromosome 10 (PTEN) and/or Na⁺/H⁺ exchanger regulatory factor-1 (NHERF1) relative to equivalent, unmodified Caco-2 cells; and have increased levels or activity of Polo Like Kinase 4 (PLK4) and/or Aurora A relative to equivalent, unmodified Caco-2 cells.

The present invention also provides a cell culture comprising more than one modified Caco-2 cell, wherein the modified Caco-2 cells have reduced levels or activity of Protein Kinase C-zeta (PKCz) relative to equivalent, unmodified Caco-2 cells.

The present invention also provides a cell culture comprising more than one modified Caco-2 cell, wherein the modified Caco-2 cells have reduced levels or activity of Protein Kinase C-zeta (PKCz) relative to an equivalent, unmodified Caco-2 cell; and having increased levels or activity of Polo Like Kinase 4 (PLK4) relative to equivalent, unmodified Caco-2 cells.

- 5 Preferably, the cell culture comprises a population of modified Caco-2 cells as described herein.

Preferably, the modified Caco-2 cells are provided as an 3D organotypic culture model.

- 10 The present invention also provides a method for making a cell culture of modified Caco-2 cells, comprising the steps of:

providing more than one modified Caco-2 cell as described herein; and
culturing said modified Caco-2 cells in a culture medium.

- 15 The present invention also provides a method for making a 3D cell culture of modified Caco-2 cells, comprising the steps of:

providing more than one modified Caco-2 cell as described herein;
inoculating said modified Caco-2 cells onto a three-dimensional framework; and
culturing said modified Caco-2 cells in a culture medium.

Preferably the 3D framework is provided in the culture medium.

20

The present invention also provides a method of making a 3D cell culture of modified Caco-2 cells, comprising the steps of:

- providing more than one modified Caco-2 cell as described herein;
inoculating said modified Caco-2 cells onto a three-dimensional framework; and
25 culturing said modified Caco-2 cells so that the modified Caco-2 cells attach to and substantially envelope the framework.

Preferably the 3D framework is provided in the culture medium.

The present invention also provides a cancer model system comprising:

- 30 a. a modified Caco-2 cell as described herein; or

- b. a cell culture as described herein.

Preferably the cancer model system is a 3D culture model system.

More preferably, the cancer model system is an organotypic 3D culture model system.

Preferably, the cancer model system is an adenocarcinoma model system. More preferably,
5 the cancer model system is a colorectal adenocarcinoma model system.

The present invention also provides a method for making a cell culture of modified Caco-2 cells, comprising the steps of:

- a. providing more than one modified Caco-2 cell as described herein; and

- 10 b. culturing said modified Caco-2 cells in a culture medium.

Preferably the cell culture is a 3D cell culture, and the method further comprises inoculating said modified Caco-2 cells onto a 3D framework. Preferably the 3D framework is provided in the culture medium. Preferably, the method further comprises culturing said modified Caco-2 cells so that the modified Caco-2 cells attach to and substantially envelope the framework.

15

The present invention also provides a method of producing a cancer model system for screening for a cancer-therapeutic compound, which method comprises; reducing the levels or activity of an endogenous PKC ζ and/or PTEN and/or NHERF1 protein within a Caco-2 cell such that the resulting modified Caco-2 cell has reduced levels or activities of PKC ζ
20 and/or PTEN and/or NHERF1 relative to an equivalent, unmodified Caco-2 cell. Preferably, the cancer is an adenocarcinoma. Preferably, the cancer is colorectal adenocarcinoma. Suitably, the cancer may be low-grade colorectal adenocarcinoma.

20

Preferably, the method the method further comprises transforming the Caco-2 cell with a recombinant polynucleotide capable of expressing PLK4 and/or a recombinant
25 polynucleotide capable of expressing Aurora A; such that upon expression of the recombinant polynucleotide in the Caco-2 cell, the resulting modified Caco-2 cell has increased levels or activity of PLK4 and/or Aurora A relative to an equivalent, unmodified (i.e. untransformed or transformed with empty vector control) Caco-2 cell. Preferably, the cancer is an adenocarcinoma. Preferably, the cancer is colorectal adenocarcinoma. Suitably,
30 the cancer may be high-grade colorectal adenocarcinoma.

25

30

The present invention also provides a method of screening for a cancer-therapeutic compound, comprising;

- a. providing a modified Caco-2 cell as described herein, or a cancer model system as described herein, or a cell culture as described herein; and
- 5 b. contacting the modified Caco-2 cell, cancer model or cell culture with a candidate cancer-therapeutic compound; and
- c. assaying the effect of the candidate compound on a biological activity of the modified Caco-2 cell, cancer model or cell culture.

10 Preferably, the cancer is an adenocarcinoma. Preferably, the cancer is colorectal adenocarcinoma. Suitably, the cancer may be low-grade colorectal adenocarcinoma. Suitably, the cancer may be high-grade colorectal adenocarcinoma.

The present invention also provides a method of making a cancer model system, comprising the steps of:

- 15 providing more than one modified Caco-2 cell as described herein; and
- culturing said modified Caco-2 cells in a culture medium.

20 Preferably, the cancer model system is a 3D cell culture, and the method further comprises inoculating said modified Caco-2 cells onto a 3D framework. Preferably, the method further comprises culturing said modified Caco-2 cells so that the modified Caco-2 cells attach to and substantially envelope the framework.

Preferably, the cancer is an adenocarcinoma. Preferably, the cancer is colorectal adenocarcinoma. Suitably, the cancer may be low-grade colorectal adenocarcinoma. Suitably, the cancer may be high-grade colorectal adenocarcinoma.

25 The present invention also provides a method of screening for a cancer-therapeutic compound, comprising;

- providing a modified Caco-2 cell as described herein, or a cancer model system as described herein; and
- contacting the modified Caco-2 cell, cancer model system with a candidate
- 30 cancer-therapeutic compound; and
- assaying the effect of the candidate compound on a biological activity of the modified Caco-2 cell, cancer model system.

Preferably, the cancer is an adenocarcinoma. Preferably, the cancer is colorectal adenocarcinoma. Suitably, the cancer may be low-grade colorectal adenocarcinoma. Suitably, the cancer may be high-grade colorectal adenocarcinoma.

- 5 The present invention also provides a method of determining a candidate target for cancer-therapy, wherein the method comprises the steps of:

providing a modified Caco-2 cell as described herein; and

determining the level or activity of one or more polynucleotides and/or polypeptides in the modified Caco-2 cell;

- 10 determining the level or activity of the corresponding one or more polynucleotides and/or polypeptides in an equivalent unmodified Caco-2 cell; and

comparing the level or activity of the one or more polynucleotides and/or polypeptides in the modified Caco-2 cell with that in the equivalent, unmodified Caco-2 cell;

- 15 wherein observed changes in the level or activity of the one or more polynucleotides and/or polypeptides in the modified compared to the unmodified cell is indicative of a candidate target for cancer-therapy.

Preferably, the cancer is an adenocarcinoma. Preferably, the cancer is colorectal adenocarcinoma. Suitably, the cancer may be low-grade colorectal adenocarcinoma.

- 20 Suitably, the cancer may be high-grade colorectal adenocarcinoma.

The present invention also provides a method of determining a candidate diagnostic or prognostic marker of cancer, wherein the method comprises the steps of:

providing a modified Caco-2 cell as described herein; and

- 25 determining the level or activity of one or more polynucleotides and/or polypeptides in the modified Caco-2 cell;

determining the level or activity of the corresponding one or more polynucleotides and/or polypeptides in an equivalent unmodified Caco-2 cell; and

- 30 comparing the level or activity of the one or more polynucleotides and/or polypeptides in the modified Caco-2 cell with that in the equivalent, unmodified Caco-2 cell;

wherein observed changes in the level or activity of the one or more polynucleotides and/or polypeptides in the modified compared to the unmodified cell is indicative of a diagnostic or prognostic marker of cancer.

Preferably, the cancer is an adenocarcinoma. Preferably, the cancer is colorectal adenocarcinoma. Suitably, the cancer may be low-grade colorectal adenocarcinoma. Suitably, the cancer may be high-grade colorectal adenocarcinoma.

Accordingly, the cancer model system based on modified Caco-2 cells described herein may be a model of low-grade or a high-grade adenocarcinoma, preferably, wherein the adenocarcinoma is colorectal adenocarcinoma.

Similarly, the cancer cells generated by the methods herein may be low-grade or high-grade cancer cells, preferably, wherein the cancer is an adenocarcinoma, preferably colorectal adenocarcinoma.

Equally, candidate diagnostic or prognostic markers of cancer, and candidate targets for cancer-therapy which can be determined by the methods described herein are preferably markers or candidate targets of low-grade or high-grade cancer, preferably, wherein the cancer is an adenocarcinoma, preferably colorectal adenocarcinoma.

The present invention also provides an expression vector for expression in a Caco-2 cell comprising;

a. a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1 or a fragment thereof and/or a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2 or a fragment thereof and/or a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3 or a fragment thereof; and

b. a polynucleotide sequence having a sequence of SEQ ID NO: 4 or a sequence of at least 90% thereto; and/or a sequence of SEQ ID NO: 5 or a sequence of at least 90% thereto;

wherein said expression vector is capable of expression in a Caco-2 cell, and wherein upon expression in a host Caco-2 cell results in a reduction in, or elimination of, expression or level or activity of PKC ζ and/or PTEN and/or NHERF1 and an increase in the expression or level or activity of PLK4 and/or Aurora A in said Caco-2 cell relative to that of an unmodified control Caco-2 cell. Preferably, the expression of any of the polynucleotide sequences contained therein is under the control of an inducer molecule.

The present invention also provides an expression vector for expression in a Caco-2 cell comprising;

- a. a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1 or a fragment thereof and which upon expression in a host Caco-2 cell results in a reduction in, or elimination of, expression or level or activity of native PKC ζ polypeptide in the host Caco-2 cell, relative to an equivalent, unmodified caco-2 cell; and
- b. a polynucleotide sequence having a sequence of SEQ ID NO: 4 or a sequence of at least 90% thereto and which upon expression in a host Caco-2 cell results in an increase in the expression level or activity of PLK4 polypeptide in the host Caco-2 cell, relative to an equivalent, unmodified Caco-2 cell.

Polynucleotides

Throughout, polynucleotides referred to herein, for example those targeting PKC ζ and/or PTEN and/or NHERF1 and/or those encoding PLK4 and/or Aurora A, may be isolated nucleic acid molecules and may be a DNA molecule, a cDNA molecule, an RNA molecule or synthetically produced DNA or RNA or a chimeric nucleic acid molecule. In embodiments where the polynucleotide is an RNA, it will be understood that normally uracil (U) is to be used in place of thymine (T). In particular, cDNA reference sequences are provided which encode PKC ζ (SEQ ID NO: 1), PTEN (SEQ ID NO: 2), NHERF1 (SEQ ID NO: 3), PLK4 (SEQ ID NO: 4) and Aurora A (SEQ ID NO: 5). It will be understood that where a polynucleotide is required that is capable of hybridising to an mRNA sequence corresponding to a cDNA sequence of SEQ ID NOs: 1-5 (for example in order to reduce transcript levels, downregulate levels or activity of the encoded protein, or as a probe to measure expression at the mRNA level), then the mRNA sequence may be generated by converting the relevant reference cDNA sequence, using a conversion tool that substitutes uracils for thiamines, which is widely available.

Throughout, the term "polynucleotide" as used herein refers to a deoxyribonucleotide or ribonucleotide polymer in single- or double-stranded form, or sense or anti-sense, and encompasses analogues of naturally occurring nucleotides that hybridize to nucleic acids in a manner similar to naturally occurring nucleotides. In the case of polynucleotides encoding PKCz and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A, such polynucleotides may be derived from *Homo sapiens* or may be derived from any other organism, or may be synthesised *de novo*. In all aspects of the invention which comprise transforming one or more Caco-2 cells with polynucleotides either targeting (i.e. capable of hybridising to) polynucleotides encoding (i.e.mRNAs) PKCz and/or PTEN and/or NHERF1 and/or transforming one or more Caco-2 cells with polynucleotides encoding PLK4 and/or Aurora A (i.e. overexpression constructs), the recipient Caco-2 cell may be transformed with a multiplicity of said polynucleotides. The recipient Caco-2 cell may be transiently transformed with one or more of said polynucleotides. The recipient Caco-2 cell may have one or more of said polynucleotides stably incorporated into its genome. Preferably, one or more of the polynucleotides, e.g. those encoding polynucleotides targeting PKCz and/or PTEN and/or NHERF1 (e.g. siRNA targeting those polynucleotides), may be expressed in the modified Caco-2 cell and its expression results an decrease in the expression or level of the corresponding polypeptide (i.e. PKCz and/or PTEN and/or NHERF1) compared to that of an equivalent untransformed control Caco-2 cell. Preferably, one or more of the polynucleotides, e.g. encoding PLK4 and/or Aurora A, may be expressed in the modified Caco-2 cell and its expression results an increase in the expression or level of the corresponding polypeptide (i.e. PLK4 and/or Aurora A) compared to that of an equivalent untransformed control Caco-2 cell.

The person skilled in the art will recognise that in designing appropriate siRNA sequences to reduce levels or activity of PKCz and/or PTEN and/or NHERF1 in a Caco-2 cell, the siRNA sequences should be capable of binding selectively and specifically to the mRNA transcripts corresponding to the relevant reference cDNA sequences (PKCz (SEQ ID NO: 1), PTEN (SEQ ID NO: 2), NHERF1 (SEQ ID NO: 3), PLK4 (SEQ ID NO: 4) and Aurora A (SEQ ID NO: 5)) or fragments or variants thereof. The siRNA sequence will therefore be hybridizable to that nucleotide sequence, preferably under stringent conditions, more preferably very high stringency conditions.

Polypeptide and Polynucleotide Sequence Identity

Polynucleotides or polypeptides used herein are described by reference to the respective reference polypeptide or polynucleotide sequences, i.e. those of SEQ ID NO: 1, SEQ ID NO:2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID

NO: 8, SEQ ID NO: 9 and SEQ ID NO: 10, but it will be appreciated that these reference sequences include any variant sequence having the defined percentage identity therewith. Such percentage identities include any of the following: where a reference nucleic acid or polypeptide sequence and sequences of at least a certain percentage identity are disclosed, e.g. at least 70%, then optionally the percentage identity may be different. For example: a percentage identity which is selected from one of the following: at least 70%, at least 71%, at least 72%, at least 73%, at least 74%, at least 75%, at least 76%, at least 77%, at least 78%, at least 79%, at least 80%, at least 81%, at least 82%, at least 83%, at least 84%, at least 85%, at least 86%, at least 87%, at least 88%, at least 89%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, at least 99.5%, at least 99.6%, at least 99.7%, at least 99.8% or at least 99.9%. Such sequence identity with an amino acid or nucleic acid sequence is a function of the number of identical positions shared by the sequences in a selected comparison window, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences. The popular multiple alignment program ClustalW (Thompson et al., 1994, Nucleic Acids Research, 22, 4673-4680; Thompson et al., 1997, Nucleic Acids Research, 24, 4876-4882) is a preferred way for generating multiple alignments of polypeptides, proteins, polynucleotides (comprising RNA, DNA or synthetic nucleic acids) in accordance with the invention. Suitable parameters for ClustalW may be as follows: For DNA alignments: Gap Open Penalty = 15.0, Gap Extension Penalty = 6.66, and Matrix = Identity. For protein alignments: Gap Open Penalty = 10.0, Gap Extension Penalty = 0.2, and Matrix = Gonnet. In all aforementioned aspects of the present invention, amino acid residues may be substituted conservatively or non-conservatively. Conservative amino acid substitutions refer to those where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not alter the functional properties of the resulting polypeptide. Similarly it will be appreciated by the skilled reader that nucleic acid sequences may be substituted conservatively or non-conservatively without affecting the function of the polypeptide. Conservatively modified nucleic acids are those substituted for nucleic acids which encode identical or functionally identical variants of the amino acid sequences. It will be appreciated by the skilled reader that each codon in a nucleic acid (except AUG and UGG; typically the only codons for methionine or tryptophan, respectively) can be modified to yield a functionally identical molecule. Accordingly, each silent variation (i.e. synonymous codon) of a polynucleotide or polypeptide, which encodes a polypeptide of the present invention, is implicit in each described polypeptide sequence.

The invention is based on decreasing the expression or activity of PKC ζ and/or PTEN and/or NHERF1, alone, or in combination with increasing the expression or activity of PLK4 and/or Aurora A in Caco-2 cells. The invention is based on decreasing the expression or activity of PKC ζ and/or PTEN and/or NHERF1 protein, alone, or in combination with increasing the expression or activity of PLK4 and/or Aurora A protein in Caco-2 cells. Modulating the levels or activity of a polypeptide encoded by a nucleic acid molecule may be achieved by various means well known in the art. Decreasing the expression or activity of the relevant protein in may involve, for example suppression of expression at the transcript level, for instance by siRNA knockdown, reducing transcript or polypeptide stability, and/or reducing the biological activity of the protein. On the other hand, increasing the expression or activity of the relevant protein may involve elevating polypeptide or transcript levels by over-expression or increasing the stability thereof and/or elevating in the biological activity of the protein.

Optionally, increased or decreased expression may be achieved, for example, by elevating mRNA levels encoding said polypeptide by placing the nucleotide under the control of a strong promoter sequence or altering the gene dosage by providing a cell with multiple copies of said gene or its complement. Alternatively, the stability of the mRNA encoding said polypeptide may be modulated to alter the steady state levels of an mRNA molecule; this may conveniently be achieved via alteration to the 5' or 3' untranslated regions of the mRNA. Similarly, the production of a polypeptide may be modified by altering the efficiency of translational processing, increasing or decreasing protein stability or by altering the rate of post translational modification (e.g. proteolytic cleavage) or secretion.

Commonly, where a cell naturally expresses said PKC ζ and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A, modification of their expression may be achieved by altering the expression pattern of the native gene(s) and/or the processing and/or production of the polypeptide. This may be achieved by any suitable method, including, but not limited to altering transcription of the gene, and/or translation of the mRNA into polypeptide, and post-translational modification of the polypeptide. Altering the expression pattern of a native gene may be achieved by placing it under control of a heterologous regulatory sequence, which is capable of directing the desired expression pattern of the gene. Suitable regulatory sequences may be placed 5' and/or 3' of the endogenous gene and may include, but are not limited to promoter sequences, terminator fragments, polyadenylation sequences or enhancer sequences operably linked to the sequences of interest.

Alteration, whether activation or repression of the expression pattern of the native PKC ζ and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A gene(s) and/or the processing

and/or production of the polypeptide(s) may also be achieved by using CRISPR-Cas technology, for example, by providing a catalytically inactive Cas9 (dCas9) fusion to synthetically repress expression of PKC ζ and/or PTEN and/or NHERF1 and/or activate expression of PLK4 and/or Aurora A. In both approaches, a single guide RNA (sgRNA) may be designed to direct the dCas9 repressor or activator to a chosen genomic location, i.e. a genomic sequence specific to the target polypeptide. Potential target locations include promoter regions, regulatory regions, and early coding regions. To achieve transcriptional repression, dCas9 can be used by itself (repressing transcription through steric hindrance) or as part of a dCas9–KRAB transcriptional repressor fusion protein. Alternatively, transcriptional activation of PLK4 and/or Aurora A can be achieved by various approaches, for instance using the VP64 transcriptional activator. One approach is a dCas9–VP64 fusion protein. Another, aimed at signal amplification, requires fusion of dCas9 to a repeating array of peptide epitopes, which modularly recruit multiple copies of single-chain variable fragment (ScFv) antibodies fused to transcriptional activation domains. A further approach is a dCas9–VP64 fusion protein together with a modified sgRNA scaffold with an MS2 RNA motif loop. This MS2 RNA loop recruits MS2 coat protein (MCP) fused to additional activators such as p65 and heat shock factor 1 (HSF1). In this way, expression of PKC ζ and/or PTEN and/or NHERF1 can be knocked-down and that of PLK4 and/or Aurora A can be elevated as desired. A number of plasmids which are designed to be manipulated for the purpose are freely available from plasmid repositories, such as <https://www.addgene.org/crispr/>.

Optionally, the overall levels of PKC ζ and/or PTEN and/or NHERF1 in modified Caco-2 cells described herein are decreased compared to those of control Caco-2 cells. The overall levels of PKC ζ and/or PTEN and/or NHERF1 of modified Caco-2 cells described herein may be decreased in the range 5 fold to 1000 fold relative to control Caco-2 cells (e.g. genetically equivalent but unmodified Caco-2 cells); optionally decreased 10 fold, 15 fold, 20 fold, 30 fold, 40 fold, 50 fold, 60 fold, 70 fold, 80 fold, 90 fold, 100 fold, 150 fold, 200 fold, 300 fold, 400 fold, 500 fold, 600 fold, 700 fold, 800 fold, 900 fold or 1000 fold those of control cells (e.g. genetically equivalent but unmodified Caco-2 cells).

Suppression of expression levels or activity can be achieved by a range of means, which are well known in the art and include, for example, modification of the encoding genomic DNA in order to disrupt the biological activity, structure, characteristics of the expression product, the rate of expression or the manner of expression control. Such modifications may be achieved by a variety of techniques which are now considered to be standard in the art, and may include, but are not limited to mutations, insertions, deletions and substitutions of one or more nucleotides. This may be achieved by for example, chemical mutagenesis or via one of

the available DNA editing technologies, for example, Zinc Finger Nucleases (ZFNs), TALENs or CRISPR Cas-based genome editing. Such modification may produce non-contiguous nucleotide sequences corresponding to Gene IDs: 5590 or 5728 or 9368 to produce non-functional PKCz and/or PTEN and/or NHERF1 protein in the recipient cell respectively. It may be desirable to generate single, double or triple mutants of these genes. Similarly, a reduction in expression levels or activity of these polypeptides may be achieved by editing the encoding DNA or mRNA, repression of target genes using CRISPR-based technologies or nucleases e.g. Zinc Finger Nucleases, siRNA knockdown or knock-out, inactivation of the gene by targeted homologous recombination or mutation to produce truncated or otherwise non-functional proteins or by functional inhibition or inactivation of the protein itself.

Most conveniently, PKCz and/or PTEN and/or NHERF1 expression is suppressed by siRNA knockdown. For example, the siRNA constructs used in the examples were as follows: to suppress PKCz (in this case targeting PKCz Sequence ID Z15108.1): Dharmacon Smart-Pool On-Target plus human PKCz siRNA for PKCs (Catalog No L-003526-00-0005, ThermoFisher Scientific, Dublin, Ireland) and nontargeting control siRNA oligonucleotides (catalog no. D-001810-01-05, Thermo Fisher Scientific, Dublin, Ireland) were used, to suppress PKCz. To knockdown PTEN (in this case targeting the PTEN coding region NM_000314):

5'-CCGG**CGTATACAGGAACAATATTG**CTCGAG**CAATATTGTTCTGTATACGC**-3' [SEQ ID NO: 11]; and reverse, 5'-AATT**GCGTATACAGGAACAATATTG**CTCGAG**CAATATTGTTCTGTATACG**-3' [SEQ ID NO: 12]. To

suppress NHERF1, SMARTpool: ON-TARGETplus SLC9A3R1 siRNA, L-012688-0010 were used. This targets the sequence NM_004252.4. However, the person skilled in the art will appreciate that equally effective siRNAs can be designed to different regions of the transcript of the relevant polynucleotide. It will therefore be appreciated that siRNAs targeting various regions of the transcript may be used in accordance with the present invention. However, the person skilled in the art will recognise that in designing appropriate siRNAs to reduce the expression or activity of the chosen polypeptide, it is required that the siRNA sequences be capable of binding selectively and specifically to their target transcripts; i.e. those mRNA sequences corresponding to the cDNA sequences provided by nucleotide accession numbers Z15108.1 (PKCz) or NM_000314.6 (PTEN) or NM_004252.4 (NHERF1); or fragments or variants thereof.

Additionally or alternatively, it may be desirable to use an inhibitor to reduce the functional activity of the relevant protein, for instance a PKCz pseudosubstrate inhibitor (PKCzI). In some conditions, e.g. cells that are already transiently transfected, a double transfection to achieve siRNA knockdown can be stressful to cells. In these cases it may be desirable to

use an inhibitor to reduce the functional activity of the relevant protein, for instance a PKCz pseudosubstrate inhibitor (PKCzi). One that has been extensively validated in various cell types is P1614 Sigma; MDL number MFCD03458229. In the case of NHERF1, one example of a suitable inhibitor, is a cell-permeant disruptor peptide of the ezrin binding domain in NHERF1 (KERAHQKRSSKRAPQMDWSKKNELFSNL) (SEQ ID NO: 13). Many suitable inhibitors are available and could be selected accordingly by the skilled person.

Generation of the events characteristic of high-grade cancer, in particular those characteristic of high-grade adenocarcinoma and especially of colorectal adenocarcinoma, is achieved by a reduction in the levels or activity of PKCz and/or PTEN and/or NHERF1, in combination with increased levels or activity of PLK4 and/or Aurora A. This may be achieved in a variety of ways by techniques which are well known in the art. A convenient strategy to achieve high levels of protein is to drive high-levels of transcript accumulation in host cells (e.g. Caco-2 cells), by using a strong promoter (either constitutive or in response to an inducer, e.g. doxycycline or DMSO). Alternatively or additionally, the activity of the PLK4 and/or Aurora A protein may be increased in host cells (e.g. Caco-2 cells). Preferably, high levels of expression of PLK4 and/or Aurora A may be generated by inducible overexpression. Preferably, PLK4 overexpression is generated by inducible overexpression.

Preferably, the overall levels of PLK4 and/or Aurora A polypeptide in modified Caco-2 cells are increased compared to those of control Caco-2 cells. The overall levels of PLK4 and/or Aurora A of modified Caco-2 cells of the invention may be increased in the range 5 fold to 1000 fold relative to control cells (e.g. genetically equivalent but unmodified Caco-2 cells); optionally 10 fold, 15 fold, 20 fold, 30 fold, 40 fold, 50 fold, 60 fold, 70 fold, 80 fold, 90 fold, 100 fold, 150 fold, 200 fold, 300 fold, 400 fold, 500 fold, 600 fold, 700 fold, 800 fold, 900 fold or 1000 fold those of control cells (e.g. genetically equivalent but unmodified Caco-2 cells).

Generation of Modified Caco-2 Cells

Caco-2 cells transformed with a polynucleotide or expression construct encoding PLK4 and/or Aurora A may be produced by standard techniques for the genetic manipulation of cells which are known in the art. DNA may be introduced into Caco-2 cells using any suitable technology, such as transfection, particle or microprojectile bombardment, electroporation, microinjection, electrophoresis, direct injection or other forms of direct DNA uptake. Alternatively, a combination of different techniques may be employed to enhance the efficiency of the transformation process. According to the present invention, Caco-2 cells may be stably or transiently transformed as desired, wherein stable transformation refers to polynucleotides which become incorporated into the host chromosomes such that the host

genetic material may be permanently and heritably altered and the transformed cell may continue to express traits caused by this genetic material, even after several generations of cell divisions. Transiently transformed Caco-2 cells refer to cells which contain heterologous DNA or RNA, and are capable of expressing the trait conferred by the heterologous genetic material, without having fully incorporated that genetic material into the cell's DNA. Heterologous genetic material may be incorporated into nuclear or organellar (e.g. mitochondrial) genomes as required to suit the application of the invention. Where cells are transformed with more than one polynucleotide it is envisaged that combinations of stable and transient transformations are possible, if required. More commonly, Caco-2 cells of the invention may be stably or transiently transformed with polynucleotides targeting PKCz and/or PTEN and/or NHERF1 and/or those encoding PLK4 and/or Aurora A. The polypeptide sequences and polynucleotides used in the present invention may be isolated or purified. By "purified" is meant that they are substantially free from other cellular components or material, or culture medium. "Isolated" means that they may also be free of naturally occurring sequences which flank the native sequence, for example in the case of nucleic acid molecule, isolated may mean that it is free of 5' and 3' regulatory sequences.

According to the present invention, for use in reducing the expression or activity of PKCz and/or PTEN and/or NHERF1 and optionally increasing the expression or activity of PLK4 and/or Aurora A in modified Caco-2 cells compared with corresponding polypeptides in an unmodified Caco-2 cell, polynucleotide sequences described herein may be integrated into an expression cassette comprising a regulatory sequence to generate expression of the relevant polynucleotides and/or polypeptides in Caco-2 cells. Depending on the desired application, it may be desired to, for example, integrate polynucleotide sequences into an expression cassette comprising a regulatory sequence to express the relevant genes in one part of the cell and subsequently be transported to another part of the cell. Preferably, the regulatory sequences are designed to be operably linked to the relevant polynucleotides, in order to direct expression in a manner according to the present invention.

Expression Cassettes

The polynucleotides as described herein (e.g. those encoding antisense or siRNA constructs targeting PKCz and/or PTEN and/or NHERF1 either alone, or in combination with those encoding PLK4 and/or Aurora A) and/or one or more regulatory sequences are preferably provided as part of an expression cassette (e.g. an expression vector) in order that expression of the relevant polynucleotide can be carried out in a Caco-2 cell of the invention. In some instances the expression cassettes will also comprise polynucleotide sequences targeting PKCz and/or PTEN and/or NHERF1 and/or those encoding PLK4 and/or Aurora A (for example an siRNA construct targeting PKCz and a PLK4OE construct). Suitable

expression cassettes (or vectors) will vary according to the recipient host cell and suitably may incorporate regulatory elements which allow desired expression in a Caco-2 cell of the invention and preferably which facilitate high-levels of expression. Such regulatory sequences may be capable of influencing transcription or translation of a gene or gene product, for example in terms of initiation, accuracy, rate, stability, downstream processing and mobility. Suitable expression cassettes for use in the present invention may be constructed by standard techniques known in the art, to comprise 5' and 3' regulatory sequences, including, but not limited to promoter sequences, terminator fragments, polyadenylation sequences or enhancer sequences operably linked to the sequences of interest. Such elements may be included in the expression construct to obtain the optimal expression and function of the polynucleotides (and/or any polypeptides that they encode) in the recipient Caco-2 cell. In addition, polynucleotides encoding, for example, selectable markers and reporter genes may be included. The expression cassette preferably also contains one or more restriction sites, to enable insertion of the nucleotide sequence and/or a regulatory sequence into the recipient (host) cell genome, at pre-selected loci. Also provided on the expression cassette may be transcription and translation initiation regions, to enable expression of the incoming genes, transcription and translational termination regions, and regulatory sequences. These sequences may be native to the cell being transformed, or may be heterologous. The expression cassettes may be a bi-functional expression cassette which functions in multiple hosts.

In particular, the present invention provides an expression vector comprising;

a. a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1 or a fragment thereof and/or a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2 or a fragment thereof and/or a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3 or a fragment thereof; and b. a polynucleotide sequence having a sequence of SEQ ID NO: 4 or a sequence of at least 90% thereto; and/or a sequence of SEQ ID NO: 5 or a sequence of at least 90% thereto; wherein said expression vector is capable of expression in a Caco-2 cell, and wherein upon expression in a Caco-2 cell causes a decrease in the expression of PKCz and/or PTEN and/or NHERF1 and an increase in the expression of PLK4 and/or Aurora A in said Caco-2 cell relative to that of an unmodified control Caco-2 cell.

Preferably the present invention provides an expression vector comprising;

a. a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1 or a fragment thereof; and

b. a polynucleotide sequence having a sequence of SEQ ID NO: 4 or a sequence of at least 90% thereto;

wherein said expression vector is capable of expression in a Caco-2 cell, and wherein upon expression in a Caco-2 cell results in a decrease in the expression of PKC ζ and an increase in the expression of PLK4 in said Caco-2 cell relative to that of an unmodified control Caco-2 cell.

Regulatory Sequences

A regulatory sequence is a nucleotide sequence which is capable of influencing transcription or translation of a gene or gene product, for example in terms of initiation, accuracy, rate, stability, downstream processing and mobility. Examples of regulatory sequences include promoters, 5' and 3' UTR's, enhancers, transcription factor or protein binding sequences, start sites and termination sequences, ribosome binding sites, recombination sites, polyadenylation sequences, sense or antisense sequences. They may be DNA, RNA or protein. The regulatory sequences may be derived from any organism or may be synthetic. The regulatory sequences may be derived from eukaryotic, prokaryotic, achaeal or viral organisms and may be, for example, human-, plant-, bacteria-, fungal- or virus derived. To achieve the most faithful and/or effective expression in the host (i.e. Caco-2) cells, preferably any regulatory sequences may be derived from the same species as the cell being modulated (e.g. a *homo sapiens*- derived sequence in the case of Caco-2 cells).

Promoters

Typically, promoters controlling the expression of the polynucleotides disclosed herein may be endogenous or heterologous. Suitable promoters may be constitutive, whereby they direct expression under most environmental conditions or developmental stages. Alternatively, suitable promoters may be developmental stage specific, or inducible (i.e. initiating transcription in response to a stimulus). Optionally, the promoter may be inducible, to direct expression in response to environmental, chemical or developmental cues, such as, for instance temperature, light, chemicals, and other stimuli. In some instances it may be desired to combine the strong expression of a constitutive promoter and rely on a signal peptide to direct accumulation of the polypeptide in response to a given stimulus. Where cells are transformed with more than one polynucleotide it is envisaged that different promoters may be used with each polynucleotide, if required, to optimise expression.

Advantageously, such an approach may be utilised, for example, to provide an inducible switch to initiate the characteristic events of low-grade adenocarcinoma, for example by placing a PCKz siRNA construct (comprising a sequence capable of specifically hybridising to PKCz transcript and bringing about a reduction in expression or activity of the PKCz polypeptide) under the control of an inducible promoter. Advantageously, such an approach may be utilised, for example, to provide an inducible switch between a low-grade adenocarcinoma cell and a high-grade adenocarcinoma cell e.g. by placing a PCKz siRNA construct (comprising a sequence capable of specifically hybridising to PKCz transcript and bringing about a reduction in expression or activity of the PKCz polypeptide) under the control of a constitutive promoter and PLK4 under the control of an inducible one. Optionally, to provide an additional degree of flexibility, the sequences capable of bringing about the events characteristic of low-grade adenocarcinoma (i.e. those capable of reducing the level or activity of PKCz and/or PTEN and/or NHERF1) may be inducible and those sequences capable of bringing about the events characteristic of high-grade adenocarcinoma (i.e. those capable of increasing the level or activity of PLK4 and/or Aurora A) may be inducible. In such a system, the inducer molecules of each set of sequences, or indeed, for each individual sequence may be different. Said sequences may optionally be provided as part of one or more expression vectors. Such an approach may advantageously allow diagnostic or prognostic markers to be determined which are characteristic of different grades or stages of cancer, for example by comparison of the expression pattern, level, or activity of polynucleotides or polypeptides before and after induction. Additionally, a controllable switch allows the efficacy of candidate cancer-therapeutic agents, compounds, treatment and/or dosage regimes to be determined for different stages of cancer. Suitable promoter sequences may include but are not limited to those of the cytomegalovirus (CMV) general expression promoter. It will be clear to the skilled person that other promoters may be equally effective. These promoters may be derived from *Homo sapiens* or any other suitable organism. Typically, suitable promoters will be those capable of driving high levels of expression in a Caco-2 cell background. Preferably, the PCKz siRNA construct is under the control of a CMV promoter. Preferably, PLK4 expression is under the control of a doxycycline-inducible CMV promoter. Preferably, inducible overexpression of PLK4 and/or Aurora A transgenes may be obtained using an overexpression system based on the lentiviral vectors pLenti-CMV-TetR-Blast (17492, Addgene) and pLenti-CMV/TO-Neo-Dest (17292, Addgene), where elevated-levels of expression are induced by application of doxycycline. It is envisaged that one or more siRNA constructs, targeting, for example, either different transcript variants of derived from the same gene and/or different targets e.g PKCz and/or PTEN and/or NHERF1 may be provided in a single expression vector whether alone

or in combination with expression constructs comprising sequences capable of increasing the level or activities of PLK4 and/or Aurora A in modified Caco-2 cells.

Signal Peptides

5 Polynucleotides referred to herein, for example those targeting PKC ζ and/or PTEN and/or NHERF1 and/or those encoding PLK4 and/or Aurora A, will preferably be provided in an expression vector or expression cassette. The expression cassette comprising said polynucleotides may also comprise sequences coding for a transit or signal peptide, to drive the protein encoded by the heterologous polynucleotide to a desired cellular location. Such
10 transit peptides are well known to those of ordinary skill in the art, and may include single transit peptides, as well as multiple transit peptides obtained by the combination of sequences coding for at least two transit peptides.

Subsequently, in some instances, the expression levels of the PKC ζ and/or PTEN and/or
15 NHERF1 and/or PLK4 and/or Aurora A protein in modified Caco-2 cells may be determined. Subsequently, in some instances, the activity levels of the PKC ζ and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A protein in modified Caco-2 cells may be determined. In some instances, it may be possible to directly determine functional expression, e.g. as with GFP or by enzymatic action of the protein of interest (POI) to generate a detectable
20 optical signal. However, in some instances it may be chosen to determine physical expression, e.g. by antibody probing, and rely on separate test to verify that physical expression is accompanied by the required function. Optionally, PKC ζ and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A expression may be detectable by a high-throughput screening method, for example, relying on detection of an optical signal. For this purpose, it
25 may be necessary for the protein of interest (POI) to incorporate a tag, or be labelled with a removable tag, which permits detection of expression. Such a tag may be, for example, a fluorescence reporter molecule translationally-fused to the POI, e.g. Green Fluorescent Protein (GFP), Yellow Fluorescent Protein (YFP), Red Fluorescent Protein (RFP), Cyan Fluorescent Protein (CFP) or mCherry. Such a tag may provide a suitable marker for
30 visualisation of functional PKC ζ and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A expression since its expression can be simply and directly assayed by fluorescence measurement in modified Caco-2 cells. It may be an enzyme which can be used to generate an optical signal. Tags used for detection of expression may also be antigen peptide tags. A tag employed for detection of expression may usefully be cleavable from the POI. Other
35 kinds of label may be used to mark the nucleic acid including organic dye molecules, radiolabels and spin labels which may be small molecules.

Expression Vector Delivery

The expression vectors comprising polynucleotides targeting PKC ζ and/or PTEN and/or NHERF1 and/or those encoding PLK4 and/or Aurora A can be delivered in several different ways into the recipient Caco-2 cell. Optionally, the vectors may be delivered into the recipient Caco-2 cell by methods which are well known in the art, for example lipofectamine - mediated transfection. Alternatively, the vectors may be delivered into the recipient Caco-2 cell by electroporation with in vitro transcript. In another alternative, the vectors may be delivered into the recipient Caco-2 cell by bombardment of cells with e.g. gold particles coated with the plasmids. Transfection of the Caco-2 cells in accordance with the methods of the invention may involve a single application of the vector to the cell. However, it will be understood that treatment may alternatively involve multiple applications of the same vector or composition or indeed combinations of the modified expression vectors disclosed herein. Where multiple (i.e. two or more) different vectors are applied to the same cell, these may be applied simultaneously, separately (in any order) or sequentially. The vectors may be delivered into the recipient Caco-2 cell by any combination of the methods disclosed herein.

Host Cells

In accordance with the present invention, suitable cells are human cells. In preferred aspects of the invention, host cells are Caco-2 cells. Depending on whether a low-grade or high-grade cancer model is required, the cells, in particular Caco-2 cells, may be engineered to have altered levels of expression or activity of PCK ζ and/or PTEN and/or NHERF1 alone or in combination with PLK4 and/or Aurora A.

In accordance with the invention the Caco-2 cells may be used as a host system for knock-down or knock-out of PCK ζ and/or PTEN and/or NHERF1 activity or expression alone, or in combination with, increased levels of expression of PLK4, and/or Aurora A.

If a low-grade cancer model is required, Caco-2 cells may be engineered to have reduced levels or activity of PCK ζ and/or PTEN and/or NHERF1 when compared to a genetically equivalent but unmodified control Caco-2 cell. This may conveniently be achieved, for example, by knock-down or knock-out of PCK ζ and/or PTEN and/or NHERF1 expression. Most conveniently, if a low-grade cancer model is required, one or more Caco-2 cells may be engineered to have reduced levels or activity of PCK ζ , for example by knock-down or knock-out of PCK ζ expression when compared to those of an equivalent but unmodified control Caco-2 cell. This provides an effective low-grade cancer model, in particular a low-grade colorectal adenocarcinoma model with a minimal level of genetic adjustment.

Where a high-grade cancer model is required, Caco-2 cells may be engineered to have reduced levels or activity of PCKz and/or PTEN and/or NHERF1 in combination with increased levels or activity of PLK4, and/or Aurora A when compared to those of an equivalent but unmodified control Caco-2 cell. This may conveniently be achieved, for example, by knock-down or knock-out of PCKz and/or PTEN and/or NHERF1 expression and over-expression of PLK4, and/or Aurora A. Most conveniently, if a high-grade cancer model is required, one or more Caco-2 cells may be engineered to have reduced levels or activity of PCKz and increased levels or activity of PLK4, when compared to those of an equivalent but unmodified control Caco-2 cell. This provides an effective high-grade cancer model, in particular a high-grade colorectal adenocarcinoma model with a minimal level of genetic adjustment.

The overall levels of PCKz and/or PTEN and/or NHERF1 may be decreased in the range 2 fold to 500 fold relative to equivalent but unmodified control Caco-2 cells; optionally at least 2 fold, at least 3 fold, at least 4 fold, at least 5 fold, at least 6 fold, at least 7 fold, at least 8 fold, at least 9 fold, at least 10 fold, at least 20 fold, at least 30 fold, at least 40 fold, at least 50 fold, at least 60 fold, at least 70 fold, at least 80 fold, at least 90 fold, 20 at least 100 fold, at least 150 fold, at least 200 fold, at least 250 fold, at least 300 fold, at least 350 fold, at least 400 fold, at least 450 fold or 500 fold that of equivalent but unmodified control Caco-2 cells.

The overall levels of PLK4, and/or Aurora A may be increased in the range 2 fold to 500 fold relative to equivalent but unmodified control Caco-2 cells; optionally at least 2 fold, at least 3 fold, at least 4 fold, at least 5 fold, at least 6 fold, at least 7 fold, at least 8 fold, at least 9 fold, at least 10 fold, at least 20 fold, at least 30 fold, at least 40 fold, at least 50 fold, at least 60 fold, at least 70 fold, at least 80 fold, at least 90 fold, 20 at least 100 fold, at least 150 fold, at least 200 fold, at least 250 fold, at least 300 fold, at least 350 fold, at least 400 fold, at least 450 fold or 500 fold that of equivalent but unmodified control Caco-2 cells.

The overall activity of PCKz and/or PTEN and/or NHERF1 may be decreased in the range 2 fold to 500 fold relative to equivalent but unmodified control Caco-2 cells; optionally at least 2 fold, at least 3 fold, at least 4 fold, at least 5 fold, at least 6 fold, at least 7 fold, at least 8 fold, at least 9 fold, at least 10 fold, at least 20 fold, at least 30 fold, at least 40 fold, at least 50 fold, at least 60 fold, at least 70 fold, at least 80 fold, at least 90 fold, 20 at least 100 fold, at least 150 fold, at least 200 fold, at least 250 fold, at least 300 fold, at least 350 fold, at least 400 fold, at least 450 fold or 500 fold that of equivalent but unmodified control Caco-2 cells.

The overall activity of PLK4, and/or Aurora A may be increased in the range 2 fold to 500 fold relative to equivalent but unmodified control Caco-2 cells; optionally at least 2 fold, at least 3 fold, at least 4 fold, at least 5 fold, at least 6 fold, at least 7 fold, at least 8 fold, at least 9 fold, at least 10 fold, at least 20 fold, at least 30 fold, at least 40 fold, at least 50 fold, at least 60 fold, at least 70 fold, at least 80 fold, at least 90 fold, at least 100 fold, at least 150 fold, at least 200 fold, at least 250 fold, at least 300 fold, at least 350 fold, at least 400 fold, at least 450 fold or 500 fold that of equivalent but unmodified control Caco-2 cells.

Usefully, each of the PCKz and/or PTEN and/or NHERF1 polypeptides disclosed herein is of broad application and may be used independently or in combination to generate a low-grade adenocarcinoma model in one or more Caco-2 cells. Preferably, the expression level or activity of PCKz alone is reduced compared to an equivalent but unmodified control Caco-2 cell.

Usefully, PLK4, and/or Aurora A may also be harnessed independently or in combination in order to generate a high-grade cancer model in one or more Caco-2 cells, wherein the level or activity of either or both is/are increased in a Caco-2 cell background which has reduced expression level or activity of PCKz and/or PTEN and/or NHERF1. Preferably, the expression level or activity of PLK4 is increased in a Caco-2 cell background where the expression level or activity of PCKz is reduced compared to an equivalent but unmodified control Caco-2 cell. Preferably, to efficiently generate a high-grade cancer model, in particular a high-grade colorectal adenocarcinoma model, PLK4 is over-expressed in a PCKz-deficient Caco-2 cell line.

Normally, the expression level of PKCz and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A in a Caco-2 cell will be analysed and compared with one or more reference values. Preferably, the mRNA transcript level of PKCz and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A in a Caco-2 cell will be analysed and compared with one or more reference values. Preferably, the protein level of PKCz and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A in a Caco-2 cell will be analysed and compared with one or more reference values.

Throughout, polynucleotides or polypeptides in the Caco-2 cells are said to be differentially expressed (i.e. increased or decreased) where their expression levels are significantly up- or down-regulated compared with those of one or more control Caco-2 cells or one or more reference values. This may include scaling of expression levels in relation to sample mean and sample variance. Clearly, variation in the sensitivity of individual polynucleotides or polypeptides, and cells means that different levels of confidence may be attached to each polynucleotide or polypeptide. Polynucleotides or polypeptides encoding PKCz and/or

PTEN and/or NHERF1 and/or PLK4 and/or Aurora A may be said to be significantly up-regulated (or elevated) or down-regulated (or reduced) they exhibit a 2-fold change compared with one or more control Caco-2 cells or one or more reference values. Preferably, said polynucleotides or polypeptides will exhibit a 3-fold change or more compared with one or more control Caco-2 cells or one or more reference values. More preferably polynucleotides or polypeptides of the invention will exhibit a 4-fold change or more compared with one or more control Caco-2 cells or one or more reference values. That is to say, in the case of decreased expression level of PKC ζ and/or PTEN and/or NHERF1 (down-regulation relative to reference values), the polynucleotide or polypeptide level will be half that or less than half that of the reference value or that observed in the one or more control Caco-2 cells. Preferably, the polynucleotide or polypeptide level will be less than a quarter of the level of the one or more reference values or that in the one or more Caco-2 cells. More preferably, the polynucleotide or polypeptide level will be less than 1/8 of the level of the one or more reference values or that in the one or more Caco-2 cells. Conversely, in the case of increased expression level of PLK4 and Aurora A (up-regulation relative to reference values), the polynucleotide or polypeptide level will be double or more than double that of the reference value or that observed in the one or more control Caco-2 cells. Preferably, the polynucleotide or polypeptide level will be more than 3 times the level of the one or more reference values or that in the one or more Caco-2 cells. More preferably, the polynucleotide or polypeptide level will be more than 4 times the level of the one or more reference values or that in the one or more Caco-2 cells.

Reference Values

Throughout, the term "reference value" may refer to a pre-determined reference value, for instance specifying a confidence interval or threshold value for the expression of the relevant polynucleotide or polypeptide. Alternatively, the reference value may be derived from the expression level of a corresponding polynucleotide or polypeptide in a 'control' biological sample, for example a positive (e.g. one or more low- or high-grade Caco-2 cells) or negative (e.g. one or more genetically equivalent, unmodified Caco-2 cells) control. Positive controls might include, for example one or more low- or high-grade Caco-2 cells or one or more a cancerous cells, e.g. cancerous human intestinal epithelial cells. Negative controls might include, for example one or more genetically equivalent, unmodified Caco-2 cells, or one or more genetically equivalent Caco-2 cells transformed with a equivalent but 'empty' vector, or one or more healthy (i.e. non-cancerous) cell, e.g. non-cancerous human intestinal epithelial cells. Preferably, control cells such as equivalent, unmodified Caco-2 cells will be grown or cultured in identical conditions to the cells being tested (i.e. modified Caco-2 cells)

and will be sampled at identical time points and/or stages of the cell cycle to ensure an accurate comparison.

Furthermore, the reference value may be an 'internal' standard or range of internal standards, for example a known concentration of a protein, transcript, label or compound.

5 Alternatively, the reference value may be an internal technical control for the calibration of expression values or to validate the quality of the sample or measurement techniques. This may involve a measurement of one or several transcripts or proteins within the sample which are known to be constitutively expressed or expressed at a known level (e.g. an invariant level). Accordingly, it would be routine for the skilled person to apply these known
10 techniques alone or in combination in order to quantify the level of a given polynucleotide or polypeptide in a Caco-2 cell relative to standards or other transcripts or proteins or in order to validate the quality of the Caco-2 cell sample, the assay or statistical analysis.

Reduced levels of PKCz and/or PTEN and/or NHERF1 expression in a modified Caco-2 cell when compared with one or more reference values or reference cells may suitably be
15 discerned at the transcript (mRNA) and/or protein level. Most conveniently, reduced levels of PKCz and/or PTEN and/or NHERF1 expression in modified Caco-2 cells when compared with one or more reference values or reference cells are detectable at the transcript (mRNA) level.

Elevated levels of PLK4 and/or Aurora A expression in a modified Caco-2 cell when
20 compared with one or more reference values or reference cells may suitably be discerned at the transcript (mRNA) and/or protein level. Most conveniently, elevated levels of PLK4 and/or Aurora A expression in modified Caco-2 cells when compared with one or more reference values or reference cells are detectable at the transcript (mRNA) level.

Suitably, the relevant polynucleotides or polypeptides chosen to determine expression levels
25 of PKCz and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A in a Caco-2 cell are selected from the group consisting of: protein; and nucleic acid molecule encoding the protein, preferably a nucleic acid molecule, and in particular an mRNA molecule.

It is preferred that the levels of the relevant polynucleotides or polypeptides in a Caco-2 cell are investigated using specific binding partners. Suitably the binding partners may be
30 selected from the group consisting of: complementary nucleic acids; aptamers; antibodies or antibody fragments. Suitable classes of binding partners will be apparent to the skilled person.

Suitably, the levels of the relevant polynucleotides or polypeptides in a Caco-2 cell may be detected by direct assessment of binding between the target molecules and binding
35 partners.

Conveniently, the levels of the relevant polynucleotides or polypeptides in a Caco-2 cell are detected using a reporter moiety attached to a binding partner. Preferably, the reporter moiety is selected from the group consisting of: fluorophores; chromogenic substrates; and chromogenic enzymes.

- 5 Equally the activity of the PKC ζ and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A protein may be assessed by methods which will be apparent to those skilled in the art.

Binding Partners

10 Expression levels of PKC ζ and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A in a biological sample (i.e. in Caco-2 cells) may be investigated using binding partners which bind or hybridize specifically to the polynucleotides or polypeptides or a fragment thereof. In relation to the present invention the term 'binding partners' may include any ligands, which are capable of binding specifically to the relevant polynucleotides or polypeptides and/or variants thereof with high affinity. Said ligands include, but are not limited to nucleic acids
15 (DNA or RNA), proteins, peptides, antibodies, synthetic affinity probes, carbohydrates, lipids, artificial molecules or small organic molecules such as drugs. In certain embodiments the binding partners may be selected from the group comprising: complementary nucleic acids; aptamers; antibodies or antibody fragments. In the case of detecting mRNAs, nucleic acids represent highly suitable binding partners.

20 In the context of the present invention, a binding partner specific to a polynucleotide or polypeptide should be taken as requiring that the binding partner should be capable of binding to at least one such target molecule in a manner that can be distinguished from non-specific binding to molecules that are not targets. A suitable distinction may, for example, be based on distinguishable differences in the magnitude of such binding.

25 In preferred embodiments of the methods or devices of the invention, the target molecule for detection is a nucleic acid, preferably an mRNA molecule, and the binding partner is selected from the group comprising; complementary nucleic acids or aptamers.

Suitably the binding partner is a nucleic acid molecule (typically DNA, but it can be RNA) having a sequence which is complementary to the sequence the relevant mRNA or cDNA
30 against which it is targeted. Such a nucleic acid is often referred to as a 'probe' (or a reporter or an oligo) and the complementary sequence to which it binds is often referred to as the 'target'. Probe-target hybridization is usually detected and quantified by detection of fluorophore-, silver-, or chemiluminescence-labeled targets to determine relative abundance of nucleic acid sequences in the target.

Probes can be from 25 to 1000 nucleotides in length. However, lengths of 30 to 100 nucleotides are preferred, and probes of around 50 nucleotides in length are commonly used successfully in complete transcriptome analysis.

While the determination of suitable probes can be difficult, e.g. in very complex arrays, there are many commercial sources of complete transcriptome arrays available, and it is routine to develop bespoke arrays to detect any given set of specific mRNAs using publically available sequence information. Commercial sources of microarrays for transcriptome analysis include Illumina and Affymetrix. Nucleotide probe sequences may be designed to any sequence region of the target transcripts (corresponding to the cDNA sequences of accession numbers Z15108.1 (PKCz) or NM_000314.6 (PTEN) or NM_004252.4 (NHERF1) or NM_014264.4 (PLK4) or NM_198433.2 (Aurora A)) or a variant thereof. The person skilled in the art will appreciate that equally effective probes can be designed to different regions of the transcript and that the effectiveness of the particular probes chosen will vary, amongst other things, according to the platform used to measure transcript abundance and the hybridization conditions employed. It will therefore be appreciated that probes targeting various regions of the transcript may be used in accordance with the present invention.

In other suitable embodiments of the invention, the target may be a protein, and the binding partner is selected from the group comprising; antibodies, antibody fragments or aptamers. Polynucleotides encoding any of the specific binding partners of polynucleotides or polypeptides recited above may be isolated and/or purified nucleic acid molecules and may be RNA or DNA molecules.

Throughout, the term "polynucleotide" as used herein refers to a deoxyribonucleotide or ribonucleotide polymer in single- or double-stranded form, or sense or anti-sense, and encompasses analogues of naturally occurring nucleotides that hybridize to nucleic acids in a manner similar to naturally occurring nucleotides. Such polynucleotides may be derived from Homo sapiens, or may be synthetic or may be derived from any other organism. Commonly, polypeptide sequences and polynucleotides used as binding partners in the present invention may be isolated or purified. By "purified" is meant that they are substantially free from other cellular components or material, or culture medium. "Isolated" means that they may also be free of naturally occurring sequences which flank the native sequence, for example in the case of nucleic acid molecule, isolated may mean that it is free of 5' and 3' regulatory sequences. In a preferred embodiment the nucleic acid is mRNA. There are numerous suitable techniques known in the art for the quantitative measurement of mRNA transcript levels in a given biological sample. These techniques include but are not limited to; "Northern" RNA blotting, Real Time Polymerase Chain Reaction (RTPCR), Quantitative Polymerase Chain Reaction (qPCR), digital PCR (dPCR), multiplex PCR,

Reverse Transcription Quantitative Polymerase Chain Reaction (RT-qPCR), branched DNA signal amplification or by high-throughput analysis such as hybridization microarray, Next Generation Sequencing (NGS) or by direct mRNA quantification, for example by "Nanopore" sequencing. Alternatively, "tag based" technologies may be used, which include but are not limited to Serial Analysis of Gene Expression (SAGE). Commonly, the levels of mRNA transcript in a given Caco-2 cell sample may be determined by hybridization to specific complementary nucleotide probes on a hybridization microarray or "chip", by Bead Array Microarray technology or by RNA-Seq where sequence data is matched to a reference genome or reference sequences. In a preferred embodiment, where the target nucleic acid is mRNA, the levels of target transcript(s) will be determined by PCR. Preferably mRNA transcript abundance will be determined by qPCR, dPCR or multiplex PCR. More preferably, transcript abundance will be determined by multiplex-PCR. Nucleotide primer sequences may be designed to any sequence region of the target transcripts (corresponding to the cDNA sequences of accession numbers Z15108.1 (PKCz) or NM_000314.6 (PTEN) or NM_004252.4 (NHERF1) or NM_014264.4 (PLK4) or NM_198433.2 (Aurora A)) or a variant thereof. The person skilled in the art will appreciate that equally effective primers can be designed to different regions of the transcript or cDNA of the relevant polynucleotide, and that the effectiveness of the particular primers chosen will vary, amongst other things, according to the platform used to measure transcript abundance, the biological sample (i.e. the Caco-2 cells) and the hybridization conditions employed. It will therefore be appreciated that primers targeting various regions of the transcript may be used in accordance with the present invention. However, the person skilled in the art will recognise that in designing appropriate primer sequences to detect expression of the chosen polynucleotide, it is required that the primer sequences be capable of binding selectively and specifically to the transcripts or cDNA sequences corresponding to the target sequence (i.e. nucleotide accession numbers Z15108.1 (PKCz) or NM_000314.6 (PTEN) or NM_004252.4 (NHERF1) or NM_014264.4 (PLK4) or NM_198433.2 (Aurora A)) or fragments or variants thereof.

Many different techniques known in the art are suitable for detecting binding of the target sequence and for high-throughput screening and analysis of protein interactions. According to the present invention, appropriate techniques include (either independently or in combination), but are not limited to; co-immunoprecipitation, bimolecular fluorescence complementation (BiFC), dual expression recombinase based (DERB) single vector system, affinity electrophoresis, pull-down assays, label transfer, yeast two-hybrid screens, phage display, in vivo crosslinking, tandem affinity purification (TAP), ChIP assays, chemical cross-linking followed by high mass MALDI mass spectrometry, strep-protein interaction experiment (SPINE), quantitative immunoprecipitation combined with knock-down (QUICK),

proximity ligation assay (PLA), bio-layer interferometry, dual polarisation interferometry (DPI), static light scattering (SLS), dynamic light scattering (DLS), surface plasmon resonance (SPR), fluorescence correlation spectroscopy, fluorescence resonance energy transfer (FRET), isothermal titration calorimetry (ITC), microscale thermophoresis (MST),
5 chromatin immunoprecipitation assay, electrophoretic mobility shift assay, pull-down assay, microplate capture and detection assay, reporter assay, RNase protection assay, FISH/ISH co-localization, microarrays, microsphere arrays or silicon nanowire (SiNW)-based detection. Where protein levels are to be quantified, preferably the interactions between the binding partner and protein will be analysed using antibodies with a fluorescent reporter attached. In
10 certain embodiments of the invention, the expression level of a particular protein may be detected by direct assessment of binding of the protein to its binding partner. Suitable examples of such methods in accordance with this embodiment of the invention may utilise techniques such as electro-impedance spectroscopy (EIS) to directly assess binding of binding partners (e.g. antibodies) to target proteins (e.g. PKC ζ and/or PTEN and/or NHERF1
15 and/or PLK4 and/or Aurora A proteins).

In certain embodiments of the present invention the binding partner may be an antibody, or antibody fragment, and the detection of the target molecules utilises an immunological method. In certain embodiments of the methods or devices, the immunological method may be an enzyme-linked immunosorbent assay (ELISA) or utilise a lateral flow device.

20 A method of the invention may further comprise quantification of the amount of the target molecules indicative of expression of the PKC ζ and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A that is present in a Caco-2 cell. Suitable methods of the invention, in which the amount of the target molecule present has been quantified, and the volume of the sample is known, may further comprise determination of the concentration of the target
25 molecules present in the Caco-2 cell sample.

Reporter moieties

In preferred embodiments of the present invention the expression levels of the protein in a biological sample may be determined. In some instances, it may be possible to directly
30 determine expression, e.g. as with GFP or by enzymatic action of the protein of interest (POI) to generate a detectable optical signal. However, in some instances it may be chosen to determine physical expression, e.g. by antibody probing, and rely on separate test to verify that physical expression is accompanied by the required function.

In preferred embodiments of the invention, the expression levels of a particular
35 polynucleotide or polypeptide will be detectable in a biological sample (i.e. Caco-2 cells) by a

high-throughput screening method, for example, relying on detection of an optical signal, for instance using reporter moieties. For this purpose, it may be necessary for the specific binding partner (or the sequence encoding the protein itself e.g. PLK4) to incorporate a tag, or be labelled with a removable tag, which permits detection of expression. Such a tag may be, for example, a fluorescence reporter molecule translationally-fused to the protein of interest (POI), e.g. Green Fluorescent Protein (GFP), Yellow Fluorescent Protein (YFP), Red Fluorescent Protein (RFP), Cyan Fluorescent Protein (CFP) or mCherry. Such a tag may provide a suitable marker for visualisation of target expression since its expression can be simply and directly assayed by fluorescence measurement in vitro or on an array. Alternatively, it may be an enzyme which can be used to generate an optical signal. Tags used for detection of expression may also be antigen peptide tags. Similarly, reporter moieties may be selected from the group consisting of fluorophores; chromogenic substrates; and chromogenic enzymes. Other kinds of label may be used to mark a nucleic acid binding partner including organic dye molecules, radiolabels and spin labels which may be small molecules.

Preferably, the levels of a polynucleotide or several polynucleotides of interest will be quantified by measuring the specific hybridization of a complementary nucleotide probe to the polynucleotide of interest under high-stringency or very high-stringency conditions.

Preferably, probe-target hybridization will be detected and quantified by detection of fluorophore-, silver-, or chemiluminescence-labelled probes to determine relative abundance of nucleic acid sequences in the sample. Alternatively, levels of mRNA transcript abundance can be determined directly by RNA sequencing or nanopore sequencing technologies.

The methods or devices of the invention may make use of molecules selected from the group consisting of: the protein of interest (e.g. PKCz and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A); and nucleic acid encoding the protein.

Probes and Hybridization Conditions

The person skilled in the art would regard it as routine to design nucleotide or polypeptide probes to any sequence region of the relevant transcripts (corresponding to the cDNA sequences: Z15108.1 (PKCz) or NM_000314.6 (PTEN) or NM_004252.4 (NHERF1) or NM_014264.4 (PLK4) or NM_198433.2 (Aurora A)) or polypeptides (accession numbers CAA78813.1 (PKCz) or AAD13528.1 (PTEN) or AAH49220.1 (NHERF1) or NP_055079 (PLK4) or BC002499.2 (Aurora A)) or a fragment or variant thereof. This is also the case with nucleotide primers used where detection of expression levels is determined by PCR-based technology. The person skilled in the art will appreciate that equally effective probes

can be designed to different regions of the transcript, and that the effectiveness of the particular probes chosen will vary, amongst other things, according to the platform used to measure transcript abundance and the hybridization conditions employed. It will therefore be appreciated that probes targeting various regions of the transcript may be used in accordance with the present invention.

Of course the person skilled in the art will recognise that in designing appropriate probe sequences to detect expression levels of PKC ζ and/or PTEN and/or NHERF1 and/or PLK4 and/or Aurora A in a Caco-2 cell, it is required that the probe sequences be capable of binding selectively and specifically to the transcripts or cDNA sequences of the corresponding reference sequences or fragments or variants thereof. The probe sequence will therefore be hybridizable to that nucleotide sequence, preferably under stringent conditions, more preferably very high stringency conditions. The term "stringent conditions" may be understood to describe a set of conditions for hybridization and washing and a variety of stringent hybridization conditions will be familiar to the skilled reader. Hybridization of a nucleic acid molecule occurs when two complementary nucleic acid molecules undergo an amount of hydrogen bonding to each other known as Watson-Crick base pairing. The stringency of hybridization can vary according to the environmental (i.e. chemical/physical/biological) conditions surrounding the nucleic acids, temperature, the nature of the hybridization method, and the composition and length of the nucleic acid molecules used. Calculations regarding hybridization conditions required for attaining particular degrees of stringency are discussed in Sambrook et al. (2001, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY); and Tijssen (1993, Laboratory Techniques in Biochemistry and Molecular Biology— Hybridization with Nucleic Acid Probes Part I, Chapter 2, Elsevier, NY). The T_m is the temperature at which 50% of a given strand of a nucleic acid molecule is hybridized to its complementary strand. In any of the references herein to hybridization conditions, the following are exemplary and not limiting:

Very High Stringency (allows sequences that share at least 90% identity to hybridize)

Hybridization: 5x SSC at 65°C for 16 hours

Wash twice: 2x SSC at room temperature (RT) for 15 minutes each

Wash twice: 0.5x SSC at 65°C for 20 minutes each

High Stringency (allows sequences that share at least 80% identity to hybridize)

Hybridization: 5x-6x SSC at 65°C-70°C for 16-20 hours

Wash twice: 2x SSC at RT for 5-20 minutes each

Wash twice: 1x SSC at 55°C-70°C for 30 minutes each Low Stringency (allows sequences that share at least 50% identity to hybridize)

5

Hybridization: 6x SSC at RT to 55°C for 16-20 hours

Wash at least twice: 2x-3x SSC at RT to 55°C for 20-30 minutes each. Diagnostic Devices and Kits

10 Reference Sequences

PKCz

As used herein "PKCz" designates "Protein Kinase C-zeta". A reference cDNA sequence corresponding to full-length human PKCz mRNA transcript is available from the GenBank database under accession number Z15108, version Z15108.1 [Gene ID: 5590]; [24-Jan-15 2018]. [SEQ ID NO:1]

PTEN

As used herein "PTEN" designates "phosphatase and tensin homologue deleted on chromosome 10". A reference cDNA sequence corresponding to full-length human PTEN mRNA transcript is available from the GenBank database under accession number 20 NM_000314 version NM_000314.6; [Gene ID: 5728]; [21 Jan 2018]. [SEQ ID NO:2]

NHERF1

As used herein "NHERF1" designates "Na⁺/H⁺ exchanger regulatory factor-1". A reference cDNA sequence corresponding to full-length human NHERF1 mRNA transcript is available from the NCBI-GenBank database under accession number NM_004252; version 25 NM_004252.4; [Gene ID: 9368]; [24 Jan 2018]. [SEQ ID NO: 3]

PLK4

As used herein "PLK4" designates "Polo Like Kinase 4". A reference cDNA sequence corresponding to full-length human PLK4 mRNA transcript is available from the NCBI-GenBank database under accession number NM_014264, version NM_014264.4 (18 Jan 30 2018). [SEQ ID NO: 4]

Aurora A

As used herein "Aurora A" designates "serine/threonine-protein kinase 6". A reference cDNA sequence corresponding to full-length human Aurora A mRNA transcript is available from the NCBI-GenBank database under accession number NM_198433, version NM_198433.2; [Gene ID: 20878]; [23 Jan 2018]. [SEQ ID NO: 5]

- 5 All accession and version numbers of the reference sequences of polynucleotides disclosed herein were obtained from the NCBI-GenBank database (Flat File Release 223.0) available at <https://www.ncbi.nlm.nih.gov/genbank/> as of 29 Jan 2018.

Below are reference polynucleotide and amino acid sequences used in accordance with the invention.

10

PKCz (Protein kinase C zeta) *Homo sapiens*

Protein [GenBank: CAA78813.1]

1 megsggrvrl kahyggdifi tsvdaattfe elceevrdmc rlhqhpitl kwvdsegdpc
 61 tvssqmelee aflrarqcrd egliihvfps tpeqpglpcp gedksiyrrg arrwrklyra
 15 121 nghlfqakrf nrraycgqcs eriwglarqg yrcinckllv hkrchglvpl tcrkhmdsvm
 181 psqpepvddk nedadlpsee tdgiayiss rhkdsikdds edlkpvidgm dgikisqglg
 241 lqdfdlirvi grgsyakvll vrlkndqiy amkvvkkelv hddedidwvq tekhvfeqas
 301 snpflvglhs cfqttsrlfl vieyvnggdI mfhmqrrkl peeharfya eicialnflh
 361 ergiiyrdlk ldnvldadg hikltdygmC keglpggdt stfcgtpnyi apeilrgeey
 20 421 gfsvdwwalg vlmfemmagr spfdiitdn dmntedylfQ vilekpirip rflsvkashv
 481 lkgflnkdpk erlgrcpqtg fsdikshaff rsidwdlleK kqalppfpq itddyglndf
 541 dtqftsepvq ltpddedaik ridqsefegf eyinpllst eesv [SEQ ID NO: 6]

PKCz cDNA sequence GenBank: Z15108.1

25 1 cccaagatgg aaggagcgcg cgcccgcgtc cgccctcaagg cgcattacgg gggggacatc
 61 ttcataccca gcgtggacgc cgccacgacc ttcgaggagc tctgtgagga agtgagagac
 121 atgtgtcgtc tgcaccagca gcacccgctc accctcaagt ggggtggacag cgaaggtgac
 181 ccttgccacgg tgcctcccca gatggagctg gaagaggctt tccgctggc cgcctcagtc
 241 agggatgaag gcctcatcat tcatgttttc ccgagcacc ctgagcagcc tggcctgcca
 30 301 tgtccgggag aagacaaatc tatctaccgc cggggagcca gaagatggag gaagctgtac
 361 cgtgccaaacg gccacctctt ccaagccaag cgctttaaca ggagagcgta ctgcggtcag
 421 tgcagcgaga ggatatgggg cctcgcgagg caaggctaca ggtgcatcaa ctgcaaactg
 481 ctggtccata agcgtgcca cggcctcgtc ccgctgacct gcaggaagca tatggattct
 541 gtcattgcctt cccaagagcc tccagtagac gacaagaacg aggacgccga ccttccttcc
 35 601 gaggagacag atggaattgc ttacatttcc tcatcccga agcatgacag cattaagac
 661 gactcggagg accttaagcc agttatcgat gggatggatg gaatcaaat ctctcagggg
 721 cttgggctgc aggactttga cctaatacaga gtcacgggc gcgggagcta cgccaaggtt
 781 ctcttggtgc ggttgaagaa gaatgaccaa atttacgcca tgaagtgggt gaagaaagag
 841 ctggtgcatg atgacgagga tattgactgg gtacagacag agaagcagct gtttgagcag
 40 901 gcatccagca accccttctt ggtcggatta cactcctgct tccagacgac aagtcggttg
 961 ttcttggtca ttgagtacgt caacggcggg gacctgatgt tccacatgca gaggcagagg
 1021 aagctccctg aggagcacgc caggttctac gcggccgaga tctgcatcgc cctcaacttc
 1081 ctgcacgaga gggggatcat ctacagggac ctgaagctgg acaacgtcct cctggatgag
 1141 gacgggcaca tcaagctcac agactacggc atgtgcaagg aaggcctggg ccctgggtgac
 45 1201 acaacgagca ctttctgcgg aaccccgaa tacatcgccc ccgaaatcct gcggggagag
 1261 gagtacgggt tcagcgtgga ctggtgggag ctgggagtc tcatgtttga gatgatggcc
 1321 gggcgctccc cgttcgacat catcaccgac aacccggaca tgaacacaga ggactacctt

```

1381 ttccaagtga tcttgagaa gcccatccgg atcccccggt tctgtccgt caaagcctcc
1441 catgttttaa aaggattttt aaataaggac cccaaagaga ggctcggctg ccggccacag
1501 actggatttt ctgacatcaa gtcccacgog ttcttccgca gcatagactg ggacttgctg
1561 gagaagaagc aggcgctccc tccattccag ccacagatca cagacgacta cggctctggac
5 1621 aactttgaca cacagttcac cagcgagccc gtgcagctga cccagacga tgaggatgcc
1681 ataaagagga tgcaccagtc agagttcgaa ggctttgagt atatcaaccc attattgctg
1741 tccaccgagg agtcgggtgtg aggccgctgt cgtctctgtc gtggacacgc gtgattgacc
1801 ctttaactgt atccttaacc accgcatatg catgccaggc tgggcacggc tccgagggcg
1861 gccagggaca gacgcttgcg ccgagaccgc agagggaagc gtcagcgggc gctgctggga
10 1921 gcagaacagt ccctcacacc tggcccggca ggcagcttcg tgctggagga acttgctgct
1981 gtgcctgctg cgcggcggat ccgcggggac cctgccgagg gggctgtcat gcggtttcca
2041 aggtgcacat ttccacgga aacagaactc gatgcactga cctgtccgc caggaaagtg
2101 agcgtgtagc gtcctgagga ataaaatgtt ccgatgaaaa aaaaaa [SEQ ID NO: 1]

```

15 PKCz reference gDNA sequence

Gene ID: 5590; Location - Chromosome 1, NC_000001.11 (2050470..2185395)

PTEN

PTEN amino acid sequence [AAD13528.1]

```

20 1 mtaiikeivs rnkrryqedg fdldltiyp niiamgfpa rlegvyrnni ddvvrflnsk
    61 hknhykiynl caerhydtak fncrvaqyfp edhnppql el ikpfcedldq wlseddnhva
    121 aihckagkgr tgvmicayll hrgkflkaqe aldfygevrt rdkkgvtips qrryvyvysy
    181 llknhldyrp vallfhkmmf etipmfsggt cnpqfvvcql kvkiyssnsg ptrredkfmy
    241 fefpqpplvc gdikveffhk qnkmlkkdkm fhfwvntffi pgpeetsekv engslcdqei
25 301 dsicsierad ndkeylvltl tkndldkank dkanryfspn fkvklyftkt veepsnpeas
    361 sstsvtpdvs dnepdhyrys dttdsdpene pfdedqhtqi tkv [SEQ ID NO: 7]

```

PTEN cDNA sequence [NM_000314. Version 6]

```

30 1 cctccccctcg cccggcgagg tcccgctcgc ctctcgtctg cctccccgct cccctcggtc
    61 ttccgaggcg cccgggctcc cggcgcgagg cgggaggggg cgggcaggcc ggccggcggt
    121 gatgtggcgg gactctttat gcgctgcggc aggatacgcg ctccggcgct ggacgcgact
    181 gcgctcagtt ctctctctc ggaagctgca gccatgatgg aagtttgaga gttgagccgc
    241 tgtgaggcga ggccgggctc aggcgagggg gatgagagac ggccggcgcc gcggcccgga
    301 gcccctctca gcgcctgtga gcagccgcgg gggcagcgcc ctccggggag cggccggcct
35 361 gcggcgggcg cagcgggcgg gtttctcgcc tctcttctgt ctttcttaac cgtgcagcct
    421 cttcctcggc ttctcctgaa agggaaggtg gaagccgtgg gtcggggcgg gagccggctg
    481 aggcgcggcg gcggcgggcg cacctcccgc tctgggagcg ggggggagaa gcggcgggcg
    541 cggcgggcgc ggcggtgca gctccaggga gggggtctga gtcgctgtc accatttcca
    601 gggctgggaa cgccggagag ttggtctctc cccttctact gcctccaaca cggcgggcgg
40 661 ggccggcgga catccaggga cccgggcccg ttttaaacct cccgtccgcc gccgcgcac
    721 ccccggtggc ccgggctcgg gaggccggcg gcggaggcag ccgttcggag gattattcgt
    781 cttctcccca ttccgctgcc gccgctgcca ggcctctggc tgctgaggag aagcaggccc
    841 agtcgctgca accatccagc agccggcgca gcagccatta cccggtgctg gtccagagcc
    901 aagcgggcgg agagcgaggg gcatcagcta ccgccaaagt cagagccatt tccatcctgc
45 961 agaagaagcc ccgccaccag cagcttctgc catctctctc ctctttttt ttcagccaca
    1021 ggctcccaga catgacagcc atcatcaaag agatcgtagg cagaaacaaa aggagatatc
    1081 aagaggatgg attcgactta gacttgacct atatttatcc aaacattatt gctatgggat
    1141 ttctgcaga aagacttgaa ggcgtataca ggaacaatat tgatgatgta gtaaggtttt
    1201 tggattcaaa gcataaaaac cattacaaga tatacaatct ttgtgctgaa agacattatg
50 1261 acaccgcaa atttaattgc agagttgcac aatatccttt tgaagaccat aaccaccac
    1321 agctagaact tatcaaaccc ttttgtgaag atcttgacca atggctaagt gaagatgaca
    1381 atcatgttgc agcaattcac tgtaaagctg gaaaggagcg aactggtgta atgatatgtg

```

	1441	catatatttatt	acatcggggc	aaattttttaa	aggcacaaga	ggccctagat	ttctatgggg
	1501	aagtaaggac	cagagacaaa	aagggagtaa	ctattcccag	tcagaggcgc	tatgtgtatt
	1561	attatagcta	cctgttaaag	aatcatctgg	attatagacc	agtggcactg	ttgtttcaca
5	1621	agatgatgtt	tgaaactatt	ccaatgttca	gtggcggaac	ttgcaatcct	cagtttgtgg
	1681	tctgccagct	aaaggtgaag	atatattcct	ccaattcagg	accacacga	cgggaagaca
	1741	agttcatgta	ctttgagttc	cctcagccgt	tacctgtgtg	tggtgatc	aaagtagagt
	1801	tcttccacaa	acagaacaag	atgctaaaaa	aggacaaaat	gtttcacttt	tgggtaaata
	1861	cattcttcat	accaggacca	gaggaaacct	cagaaaaagt	agaaaatgga	agtctatgtg
10	1921	atcaagaaat	cgatagcatt	tgcagtatag	agcgtgcaga	taatgacaag	gaatatctag
	1981	tactttacttt	aacaaaaaat	gatcttgaca	aagcaataaa	agacaaagcc	aaccgatact
	2041	tttctccaaa	ttttaaggtg	aagctgtact	tcacaaaaac	agtagaggag	ccgtcaaate
	2101	cagaggctag	cagttcaact	tctgtaacac	cagatgttag	tgacaatgaa	cctgatcatt
	2161	atagatatcc	tgacaccact	gactctgac	cagagaatga	accttttgat	gaagatcagc
	2221	atacacaaa	tacaaaagtc	tgaatttttt	tttatcaaga	gggataaaac	accatgaaaa
15	2281	taaacttgaa	taaactgaaa	atggaccttt	tttttttttaa	tggcaatagg	acattgtgtc
	2341	agattaccag	ttataggaac	aattctcttt	tcctgaccac	tcttgtttta	ccctatacat
	2401	ccacaggggt	ttgacacttg	ttgtccagtt	gaaaaaagg	tgtgtagctg	tgtcatgtat
	2461	ataccttttt	gtgtcaaaa	gacattttaa	attcaattag	gattaataaa	gatggcactt
20	2521	tcccgttttt	ttccagtttt	ataaaaagtg	gagacagact	gatgtgtata	cgtaggaatt
	2581	ttttctcttt	gtgttctgtc	accaactgaa	gtggctaaag	agctttgtga	tatactgggt
	2641	cacatcctac	ccctttgcac	ttgtggcaac	agataagttt	gcagttggct	aagagaggtt
	2701	tccgaagggt	tttgctacat	tctaattgcat	gtattcgggt	taggggaatg	gagggaatgc
	2761	tcagaaagga	aataatttta	tgctggactc	tggaccatat	accatctcca	gctattttaca
25	2821	cacacctttc	tttagcatgc	tacagttatt	aatctggaca	ttcgaggaat	tggccgctgt
	2881	cactgcttgt	tgtttgcgca	ttttttttta	aagcatattg	gtgctagaaa	aggcagctaa
	2941	aggaagtga	tctgtatttg	ggtacaggaa	tgaaccttct	gcaacatctt	aagatccaca
	3001	aatgaaggga	tataaaaaata	atgtcatagg	taagaaacac	agcaacaatg	acttaaccat
	3061	ataaatgtgg	aggctatcaa	caaagaatgg	gcttgaaaca	ttataaaaaa	tgacaatgat
30	3121	ttattaaata	tgttttctca	attgtaacga	cttctccatc	tcctgtgtaa	tcaaggccag
	3181	tgttaaaatt	cagatgctgt	tagtacctac	atcagtcaac	aacttacct	tattttacta
	3241	gtttttcaatc	ataataacctg	ctgtggatgc	ttcatgtgct	gcctgcaagc	ttcttttttc
	3301	tcattaaata	taaaataattt	tgtaatgctg	cacagaaaatt	ttcaatttga	gattctacag
	3361	taagcgtttt	ttttctttga	agatttatga	tgcacttatt	caatagctgt	cagccgttcc
35	3421	acccttttga	ccttacacat	tctattacaa	tgaattttgc	agttttgcac	attttttaaa
	3481	tgtcattaac	tgtttagggaa	ttttacttga	atactgaata	catataatgt	ttatattaaa
	3541	aaggacattt	gtgttaaaaa	ggaaattaga	gttgacagta	actttcaatg	ctgcacacaa
	3601	aaaaagaca	tttgattttt	cagtagaaat	tgtcctacat	gtgctttatt	gatttgctat
	3661	tgaaagaata	gggttttttt	tttttttttt	tttttttttt	ttaaatgtgc	agtgttgaat
40	3721	catttcttca	tagtgctccc	ccgagttggg	actagggctt	caatttctact	tcttaaaaaa
	3781	aatcatcata	tatttgatat	gcccagactg	catacgattt	taagcggagt	acaactacta
	3841	ttgtaaagct	aatgtgaaga	tattattaaa	aagggttttt	tttccagaaa	tttgggtgtc
	3901	tcaaattata	ccttcacctt	gacatttgaa	tatccagcca	ttttgtttct	taatggata
	3961	aaattccatt	ttcaataact	tattggtgct	gaaattgttc	actagctgtg	gtctgacct
45	4021	gttaattttac	aaatacagat	tgaataggac	ctactagagc	agcatttata	gagtttgatg
	4081	gcaaatagat	taggcagaac	ttcatctaaa	atattcttag	taaaataatg	tgacacgttt
	4141	tcacatacctt	gtcagtttca	ttcaacaatt	tttaaatttt	taacaaagct	cctaggtatt
	4201	acacatttat	atttaaacat	tgataatag	agtattgatt	gattgctcat	aagttaaatt
	4261	ggtaaagtta	gagacaacta	ttctaaccac	tcaccattga	aattttatatg	ccacctgtgc
50	4321	tttcataaaa	gctgaaaatt	gttacctaaa	atgaaaaatca	acttcatgtt	ttgaagatag
	4381	ttataaatat	tgttctttgt	tacaatttcg	ggcaccgcat	attaaaacgt	aactttattg
	4441	ttccaatatg	taacatggag	ggccaggtca	taaataatga	cattataatg	ggcttttgca
	4501	ctgttattat	ttttcctttg	gaatgtgaag	gtctgaatga	gggttttgat	tttgaatgtt
	4561	tcaatgtttt	tgagaagcct	tgcttacatt	ttatggtgta	gtcattggaa	atggaaaaat
55	4621	ggcattatat	atattatata	tataaatata	tattatacat	actctcctta	ctttatttca
	4681	gttaccatcc	ccatagaatt	tgacaagaat	tgctatgact	gaaaggtttt	cgagtcctaa
	4741	ttaaaacttt	atttatggca	gtattcataa	ttagcctgaa	atgcattctg	taggtaattc
	4801	ctgagttttc	ggaatatttt	cttagacttt	ttggatgtgc	agcagcttac	atgtctgaag
	4861	ttacttgaag	gcatcacttt	taagaaagct	tacagttggg	ccctgtacca	tcccaagtcc
60	4921	ttttagactc	ctcttgaaca	tgtttgccat	acttttataa	gggtagttga	ataaatagca
	4981	tcaccattct	ttgctgtggc	acaggttata	aacttaagtg	gagtttaccg	gcagcatcaa
	5041	atgttttcagc	tttaaaaaat	aaaagtaggg	tacaagttaa	atgttttagtt	ctagaaattt

	5101	tgtgcaatat	gttcataacg	atggctgtgg	ttgccacaaa	gtgcctcgtt	tacctttaaa
	5161	tactgttaat	gtgtcatgca	tgcagatgga	aggggtggaa	ctgtgcacta	aagtgggggc
	5221	tttaactgta	gtatttggca	gagttgcctt	ctacctgcca	gttcaaaaagt	tcaacctgtt
5	5281	ttcatataga	atatatatac	taaaaaat	cagtctgtta	aacagcctta	ctctgattca
	5341	gcctcttcag	atactcttgt	gctgtgcagc	agtggctctg	tgtgtaaatg	ctatgcactg
	5401	aggatacaca	aaaataccaa	tatgatgtgt	acaggataat	gcctcatccc	aatcagatgt
	5461	ccatttgtta	ttgtgtttgt	taacaaccct	ttatctctta	gtgttataaa	ctccacttaa
	5521	aactgattaa	agtctcattc	ttgtcattgt	gtgggtgttt	tattaaatga	gagtttataa
10	5581	ttcaaattgc	ttaagtccat	tgaagtttta	attaatgggc	agccaaatgt	gaatacaaaag
	5641	ttttcagttt	ttttttttcc	tgctgtcctt	caaagcctac	tgtttaaaaa	aaaaaaaaaa
	5701	aaaaaacatg	gcctgagagt	agagtatctg	tctactcatg	tttaattaag	gaaaaacact
	5761	tatttttagg	gcttttagtca	tcacttcata	aattgtataa	gcacattaaa	tagcgttcta
	5821	gtcctgaaaa	agtccaagat	tcttagaaaa	ttgtgcata	ttttattatg	acagatgttt
15	5881	gaagataa	ccccagaatg	gatttgatac	tttagatttc	aattttgtgg	cttttgtcta
	5941	ttattctgta	ctctgccatc	agcatatgga	aagcttcatt	tactcatcat	gacttgtgcc
	6001	atataaaaa	tgatatttgc	gaatagtcta	aaggactttt	tgtacttgaa	tttaatcatg
	6061	ttgtttctaa	tattcttaaa	agcttgaaga	ctaaagcata	tcctttcaac	aaagcatagt
	6121	aaggtaataa	gaaagtgtag	tttgtacaag	tgtaaaaaaa	ataaagtaga	caatgttaca
20	6181	gtgggactta	ttatttcaag	tttacatttt	ctccatgtaa	ttttttaaaa	agtaaatgaa
	6241	aaaatgtgca	ataatgtaaa	atatgaagtg	tatgtgtaca	cacattttat	ttttcggtat
	6301	cttgggtata	cgtatgggtg	aaaactatac	tggagtctaa	aagtattcta	atttataaga
	6361	agacattttg	gtgatgtttg	aaaaatagaa	atgtgctagt	tttgttttta	tatcatgtcc
	6421	tttgtacggt	gtaatatgag	ctggcttggg	tcagtaaatg	ccatcaccat	ttccattgag
25	6481	aattttaaac	tcaccagtgt	ttaatatgca	ggcttccaaa	ggcttatgaa	aaaaatcaag
	6541	acccttaaat	ctagttaatt	tgctgctaac	atgaaaactc	ttggttcttt	tatttttgcc
	6601	agataattag	acacacatct	aaagcttagt	cttaaatggc	ttaagtgtag	ctattgatta
	6661	gtgctgttgc	tagttcagaa	agaaatgttt	gtgaatggaa	acaagaatat	tcagtccaaa
	6721	ctggttgaag	gacagtacct	gaaaaccagg	aaacaggata	atggaaaaag	tcttttaaaag
30	6781	atgaaatggt	ggagccaact	ttcttataga	attaattgta	tggtgctata	gaaagcctaa
	6841	ttatgtttgc	ttatttttga	gagcatatta	ttcttttatg	accataatct	tgctgttttt
	6901	ccatcttcca	aaagatcttc	cttctaatat	gtatatcaga	atgtgggtag	ccagtccagac
	6961	aaattcatat	tggttggtag	ctttaaaaag	tttgtaatgt	gaagacagga	aaggacaaaa
	7021	tagtttgctt	tggtggtagt	actctggttg	ttaagctagg	tattttgaga	ctacttcccc
35	7081	atcacacaaa	caataaaaata	atcactcata	atcctatcac	ctggagacat	agccatcggt
	7141	aatatgttag	tgactataca	atcatgtttt	cttctgtata	tccatgtata	ttctttaaaa
	7201	atgaaattta	tactgtacct	gatctcaaag	cttttttagct	tagtatatct	gtcatgaatt
	7261	tgtaggatgt	tccattgcat	cagaaaacgg	acagtgattt	gattactttc	taatgccaca
	7321	gatgcagatt	acatgtagtt	attgagaatc	ctttcgaatt	cagtggctta	atcatgaatg
40	7381	tctaaatatt	gttgacatta	ggatgatata	tgtaaaattaa	agttacattt	gtttagcata
	7441	gacaagctta	acattgtaga	tgtttctctt	caaaaaatcat	cttaaacatt	tgcatttgga
	7501	atttgtgttaa	atagaatgtg	tgaaacactg	tattagttaa	cttcatcacc	tttctacttc
	7561	cttatagttt	gaacttttca	gtttttgtag	ttcccaaaaca	gttgcctaat	ttagagcaaa
	7621	ttaattttaac	acctgccaaa	aaaaggctgc	tgttggctta	tcagttgtct	ttaaattcaa
45	7681	atgctcatgt	gactttttatc	acatcaaaaa	atatttcatt	aatgattcac	ctttagctct
	7741	gaaaattacc	gcgttttagta	attatagtgg	gcttataaaa	acatgcaact	ctttttgata
	7801	gttatattgag	aatttttggtg	aaaaatat	agctgagggc	agtatagaac	ttataaacca
	7861	atatattgat	attttttaaaa	cattttttaca	tataagtaaa	ctgccatctt	tgagcataac
	7921	tacattttaa	aataaagctg	cataattttta	aatcaagtgt	ttaacaagaa	tttatatttt
50	7981	ttattttttta	aaattaaaaa	taattttatat	ttcctctgtt	gcatgaggat	tctcatctgt
	8041	gcttataatg	gttagagatt	ttatttgtgt	ggaatgaagt	gaggcttgta	gtcatggttc
	8101	tagtgtttca	gtttgccaaag	tctgtttact	gcagtgaagt	tcatcaaagt	tttcagtgtg
	8161	gttttctgta	gcctatcatt	tactggctat	ttttttatgt	acacctttag	gattttctgc
	8221	ctactctatc	cagttgtcca	aatgatatcc	tacattttac	aaatgccctt	tcagtttcta
55	8281	ttttcttttt	ccattaaatt	gccctcatgt	cctaagtgtc	agtttgaag	tgtgtgtgtg
	8341	tgtgtctgtg	tgtgtgtgaa	tttgattttc	aagagtgtc	gacttccaat	ttgagagatt
	8401	aaataattta	attcaggcaa	acatttttca	ttggaatttc	acagttcatt	gtaatgaaaa
	8461	tgttaatcct	ggatgacctt	tgacatacag	taatgaatct	tggatattaa	tgaatttgtt
	8521	agtagcatct	tgatgtgtgt	tttaatgagt	tatttttcaa	gttgtgcatt	aaaccaagt
60	8581	tggcatactg	gaagtgttta	tatcaagttc	cattttggcta	ctgatggaca	aaaaatagaa
	8641	atgccttcct	atggagagta	tttttccttt	aaaaaattaa	aaaggttaat	tattttgact
	8701	aaaaaaaaaa	aaaaaaaa	[SEQ ID NO: 2]			

PTEN reference gDNA sequence

Gene ID: 5728

AF067844.1

5

NHERF1

Protein GenBank: AAH49220.1

1 agrglvcgpl saprgsrrpt vpgtpaclar paaqgfsaal pvrwtgrrag psrpvpigtp
 61 sraadpsqge msadaaagap lprlcclekg pngygfhlg ekgklgqyir lvepgspaek
 10 121 agllagdriv evngenveke thqqvvsrir aalnavrllv vdpetdeqlq klvgqvreel
 181 lraqeapgga eppaaaevqg agneneprea dkshpeqrel rprlctmkkg psgygfnlhs
 241 dkskpgqfir svdpdspaea sglraqdriv evngvcmegk qhgdvvsair aggdetkllv
 301 vdretdeffk kcrvipsqeh lngplvpvft ngeiqkensr ealaeaaales prpalvrsas
 361 sdtseelnsq dsppkqdsta psstsssdpi ldfnislama kerahqkrss krapqmdwsk
 15 421 knelfsnl [SEQ ID NO: 8]

SLC9A3R1 (NHERF1) cDNA sequence

[NM_004252.4]

1 ggactctggg acgctcagac gccgcgcggg gcggggattg gtctgtggtc ctctctcggc
 20 61 tcctcgcggc tcgcgcggcg cgacgggtcc tgggacacct gcttgcttgg ccggtccggc
 121 ggctcagggc ttctctgctg cgctcccggt tcgctggacg ggaagaaggc ctgggccgtc
 181 cgcgtccgctc cccatcggaa ccccaagtcg cgccgctgac ccgtcgcagg gcgagatgag
 241 cgcggacgca gcggccgggg cgccctgcc ccggctctgc tgcttgaga agggctccgaa
 301 cggctacggc ttccacctgc acggggagaa gggcaagttg ggccagtaca tccggctggg
 25 361 ggagcccggc tcgcgcggcg agaaggcggg gctgctggcg ggggaccggc tgggtggagg
 421 gaacggcgaa aacgtggaga aggagaccca ccagcaggtg gtgagccgca tccgcgccgc
 481 actcaacgcc gtgcgcctgc tgggtggtcga ccccgagacg gacgagcagc tgcagaagct
 541 cggcgtccag gtccgagagg agctgctgcg cgcccaggaa gcgccggggc aggccgagcc
 601 gccggccgcc gccgaggtgc agggggctgg caacgaaaat gagcctcgcg aggccgacaa
 30 661 gagccaccgc gagcagcgcg agcttcggcc tcggctctgt accatgaaga agggccccag
 721 tggctatggc ttcaacctgc acagcgacaa gtccaagcca ggccagttca tccggtcagt
 781 ggaccacgac tccccgctg aggttcagg gctccgggcc caggatcgca ttgtggagg
 841 gaacggggc tgcatggagg ggaagcagca tggggacgtg gtgtccgcca tcagggtg
 901 cggggacgag accaagctgc tgggtggtgga cagggaaact gacgagttct tcaagaaatg
 35 961 cagagtgatc ccactcagc agcacctgaa tgggtcccct cctgtgccct tcaccaatgg
 1021 ggagatacac aaggagaaca gtcgtgaagc cctggcagag gcagccttgg agagccccag
 1081 gccagccctg gtgagatccg cctccagtga caccagcgag gagctgaatt cccaagacag
 1141 cccccaaaaa caggactcca cagcgccctc gtctacctcc tcctccgacc ccactcctaga
 1201 cttcaacatc tccctggcca tggccaaaaga gagggccac cagaaacgca gcagcaaacg
 40 1261 ggccccgcag atggactgga gcaagaaaaa cgaactcttc agcaacctct gagcgccctg
 1321 ctgccacca gtgactggca gggccgagcc agcattccac cccacctttt tccttctccc
 1381 caattactcc cctgaatcaa tgtacaaatc agcaccaca tcccccttct tgacaaatga
 1441 tttttctaga gaactatgtt cttccctgac tttagggaag gtgaatgtgt tcccgctctc
 1501 ccgcagtcag aaaggagact ctgcctccct cctcctcact gactgcctca tcctaccggg
 45 1561 tgtccctttg ccaccctgcc tgggacatcg ctggaacctg caccatgccg ggatcatggg
 1621 accaggcgag agggcaccct cccttcctcc cccatgtgat aaatgggtcc agggctgatc
 1681 aaagaactct gactgcagaa ctgccgctct cagtggacag ggcatctgtt accctgagac
 1741 ctgtggcaga cacgtcttgt tttcatttga tttttgttaa gactgcagta ttgcagagtc
 1801 tagaggaatt tttgtttcct tgattaacat gattttcctg gttgtttacat ccaggggcatg

1861 gcagtggcct cagccttaaa cttttgttcc tactcccacc ctcagcgaac tgggcagcac
 1921 ggggaggggt tggctacccc tgcccatccc tgagccaggt accaccattg taaggaaaca
 1981 ctttcagaaa ttcagctggg tcttccaaac ctttcaaaaa aaaaaaaaaa aa

// [SEQ ID NO: 3]

5

NHERF1 gDNA sequence

Gene name –SLC9A3R1 Gene ID: 9368

Chromosomal location - 17q25.1

10 PLK4

PLK4 Protein Accession No (NP_055079)

1 matcigekie dfkvgnllgk gsfgavyrae sihtglevai kmidkkamyk agmvqrvqne
 61 vkihcqlkhp silelynyfe dsnyvylve mchngemnry lknrvkpfse nearhfmhqi
 121 itgmlylhsh gilhrdlts nlltrnmni kiadfglatq lkmphekhvt lcgtpnyisp
 15 181 eiatsahgl esdvwslgcm fytlligrpp ftdtdvknti nkvvldyem psflsieakd
 241 lihqlrrnp adrlslssvl dhpfmrsnss tkskdltgve dsidsghati staitassst
 301 sisgslfdkr rlligqplpn kmtvfpknks stdfsssgdg nsfytqwgng etsnsgrgrv
 361 iqdaeerphs rylrrayssd rsgtsnsqsq aktytmerch saemlsvskr sgggeneery
 421 sptdnnanif nffektsss sgsferpdnn qalsnhlcpg ktpfpadpt pqtetvqqwf
 20 481 gnlqinahlr ktteydsisp nrdfqghpdl qkdtknawt dtkvkknsda sdnahsvkqk
 541 ntmkymtalh skpeiiqqec vfgsdplseq sktrgmepw gyqnrtrsi tsplvahrk
 601 pirqtkkav vsildsevc velvkeyasq eyvkevlqis sdgntitiyy pnggrgfpla
 661 drppsptdni srysfndlpe kywrkyqyas rfvlvrsk pkityfrya kcilmenspg
 721 adfevwfydg vkihtktdfi qviektgksy tlksesevns lkeeikmymd haneghrcl
 25 781 alesiiisee rktrsapffp iigrkpgst sspkalsppp svdsnyptre rasfnrmvmh
 841 saasptqapi lpsmvtneg lgtttasgt dissnslkdc lpksaqlks vfvknvgwat
 901 qltsgavwvq fndgsqlvvq agvssisyts pngqttryge neklpdyikq klqclssill
 961 mfsnptpnfh

// [SEQ ID NO: 9]

30

PLK4 cDNA Sequence (NM_014264.4)

1 agcctatccg gagcgatcca tctcgttacg tcaccaccag cctagctcgg acggcaagcg
 61 gcgggagatt ttcaaaatgg gagcccagag gcaccgccc ggctcggaa ggtgtcaggg
 121 agaactttcc gtggtttcag cgctcgtcgc tggagcggcg gtttagagag ccgagcctga
 35 181 tgggcgccaa ggccggctgg ctgcttgag cgctgcctcg aagggactgc gtgaaggag
 241 ctaatccgga gaaccagagc cagagcctgg aaatatggcg acctgcatcg gggagaagat
 301 cgaggatatt aaagttagaa atctgcttgg taaaggatca tttgctgggtg tctacagagc
 361 tgagtccatt cacactgggt tggaagtggc aatcaaaatg atagataaga aagccatgta
 421 caaagcagga atggtacaga gagtccaaaa tgaggtgaaa atacattgcc aattgaaaca
 40 481 tccttctatc ttggagcttt ataactattt tgaagatagc aattatgtgt atctggtatt
 541 agaaatgtgc cataatggag aaatgaacag gtatctaaag aatagagtga aacccttctc
 601 agaaaatgaa gctcgacact tcatgcacca gatcatcaca gggatgttgt atcttcattc
 661 tcatggtata ctacaccggg acctcacact ttctaacctc ctactgactc gtaatatgaa
 721 catcaagatt gctgattttg ggctggcaac tcaactgaaa atgccacatg aaaagcacta
 45 781 tacattatgt ggaactccta actacatttc accagaaatt gccactcgaa gtgcacatgg
 841 ccttgaatct gatgtttggg ccctgggctg tatgttttat acattactta tcgggagacc
 901 acccttcgac actgacacag tcaagaacac attaaataaa gtagtattgg cagattatga
 961 aatgccatct tttttgtcaa tagaggccaa ggaccttatt caccagttac ttcgtagaaa
 1021 tccagcagat cgtttaagtc tgtcttcagt attggaccat ccttttatgt cccgaaattc
 50 1081 ttcaacaaaa agtaaagatt taggaactgt ggaagactca attgatagtg ggcagccac

```

1141 aattttctact gcaattacag cttctttccag taccagtata agtggtagtt tatttgacaa
1201 aagaagactt ttgattgggtc agccactccc aaataaaaatg actgtatttc caaagaataa
1261 aagttcaact gattttttctt cttcaggaga tggaaacagt ttttatactc agtggggaaa
5 1321 tcaagaaacc agtaatatgtg gaaggggaag agtaattcaa gatgcagaag aaaggccaca
1381 ttctcgatac cttcgtagag cttattcctc tgatagatct ggcacttcta atagtcagtc
1441 tcaagcaaaa acatatacaa tggaaacgatg tcaactcagca gaaatgcttt cagtgtccaa
1501 aagatcagga ggaggtgaaa atgaagagag gtactcacc acagacaaca atgccaacat
1561 ttttaacttc tttaaagaaa agacatccag tagttctgga tcttttgaaa gacctgataa
1621 caatcaagca ctctccaatc atctttgtcc aggaaaaact ccttttccat ttgcagaccc
10 1681 gacacctcag actgaaaccg tacaacagtg gtttggaat ctgcaaataa atgctcattt
1741 aagaaaaact actgaatatg acagcatcag cccaaaccgg gacttccagg gccatccaga
1801 tttgcagaag gacacatcaa aaaatgcctg gactgataca aaagtcaaaa agaactctga
1861 tgcttctgat aatgcacatt ctgtaaaaca gcaaaatacc atgaaatata tgactgcact
1921 tcacagttaa cctgagataa tccaacaaga atgtgttttt ggctcagatc ctctttctga
15 1981 acagagcaag actaggggta tggagccacc atgggggttat cagaatcgta cattaagaag
2041 cattacatct ccgttggttg ctacacaggtt aaaaccaatc agacagaaaa ccaaaaaggc
2101 tgtggtgagc atacttgatt cagaggaggt gtgtgtggag cttgtaaagg agtatgcac
2161 tcaagaatat gtgaaagaag ttcttcagat atctagtgat ggaaatacga tcactattta
2221 ttatccaaat ggtggtagag gttttcctct tgctgataga ccaccctcac ctactgacaa
20 2281 catcagtagg tacagctttg acaatttacc agaaaaatac tggcgaaaat atcaatatgc
2341 ttccagggtt gtacagcttg taagatctaa atctcccaa atcacttatt ttacaagata
2401 tgctaaatgc attttgatgg agaattctcc tgggtgctgat tttgagggtt ggttttatga
2461 tggggtaaaa atacacaaaa cagaagattt cattcagggtg attgaaaaga caggggaagtc
2521 ttacacttta aaaagtgaaa gtgaagttaa tagcttgaaa gaggagataa aaatgtatat
25 2581 ggaccatgct aatgaggggtc atcgtattttg ttttagcactg gaatccataa tttcagaaga
2641 ggaaaggaaa actaggagtg ctcccttttt cccaataatc ataggaagaa aacctggtag
2701 tactagtcca cctaaggcct tatcacctcc tccttctgtg gattcaaatt acccaacgag
2761 agagagagca tctttcaaca gaatgggtcat gcatagtgct gcttctccaa cacaggcacc
2821 aatccttaat ccctctatgg ttacaaatga aggacttggt cttacaacta cagcttctgg
30 2881 aacagacatc tcttctaata gtctaaaaga ttgtcttcct aaatcagcac aacttttgaa
2941 atctgttttt gtgaaaaatg ttggttgggc tacacagtta actagtggag ctgtgtgggt
3001 tcagtttaat gatgggtccc agttgggtgt gcaggcagga gtgtcttcta tcagttatac
3061 ctacacaaat ggtcaaacaa ctagggtatgg agaaaatgaa aaattaccag actacatcaa
3121 acagaaatta cagtgtctgt cttccatcct tttgatgttt tctaatacga ctctaattt
35 3181 tcattgatta aaactccttt cagacatata agtttaataa ataactttt tgttgacttt
3241 caagtaaagt gatttttttt aatttaacat aaagtcttca gaaagccttt ctatgaaaga
3301 attttaacct ataagttaa ggatgtattc tgagagaaca aagcagaatg aaacttgagt
3361 cacttactaa atatagtga tataaaatag aacacctgac tttgctctta gaccataacc
3421 ccggaactta ctatgttcat atatttgtat tgaacaatct tttaaaagca aaaatgtaaa
40 3481 tgatgtgtag tttatttgtg cttttattgt tttccctgct tctcagacat gttgagaatc
3541 atggacaaaa cctgctggaa ttttggaatt tttgaagatg taaataatgt gtatttatgt
3601 tataagtaac atatgtaaac atgtatatatt gttttatatt tatttttgta acaccagtgt
3661 ctgatgaaac atttttgcaa atgcatttta taaaaaata aatatagtga taagtacat
3721 tatcttttga ttcatttaat taaatactta tttttaaata acttaccagt aaactcactt
45 3781 tttaaatttt gttgcctgtt gaggagccaa ttaaaattta aatattaatt ttgcaaatgt
3841 taaatacatt gtttctctat tatctgaaa

```

// [SEQ ID NO: 4]

PLK4 reference gDNA

50

Gene ID 10733 (Chromosome 4)

Aurora A (Aurora kinase A)

55

Aurora A Protein (Accession No BC002499.2)

1 mdrskencis gpvkatapvg gpkrvltq fpcqnplpv nsgqaqrvlcp snssqrvplq

61 aqklvssshkp vqnqkqkqlq atsvphpvsr plnntqkskq plpsapennp eeelaskqkn

121 eeskkqwal edfeigrplg kgkfgnvyla rekqskfila lkvlfkaql kagvehqlrr

181 eveiqshlrh pnrlrygyf hdatrvylil eyaplgtyvr elqklskfde qrtatyitel
 241 analysychsk rvihrdikpe nllgsagel kiadfgwsvh apssrrtlc gtldylppem
 301 iegrmhdekv dlwslgvlcy eflvgkppfe antyqetykr isrveftfpd fvtegardli
 361 srlkhnpqsq rpmlrevleh pwitansskp sncqnkesas kqs [SEQ ID NO: 10]

5

Aurora A cDNA sequence (NM_198433.2)

1 cttaaaccgcg actcaaggcg tcgggtttgt tgtcaaccaa tcacaaggca gcctcgctcg
 61 agcgcaggcc aatcggtttt ctagctagag ggttttaactc ctatttataaa agaagaacct
 121 ttgaattcta acggctgagc tcttggaaga cttgggtcct tgggtcgagc gtgggagccg
 10 181 acgggtgggt agaccgtggg ggatatctca gtggcgagc aggacggcgg ggacaagggg
 241 cggctggctg gagtggcgga gcgtaagtc cctgtcgtgt tctcgcgtcc ctgagtgtcc
 301 ttggcgctgc cttgtgcccg ccagcgcct ttgcatccgc tctgggacac cgaggcgccc
 361 tgtaggatac tgcttggtac ttattacagc tagagggtct cactccattg cccaggccag
 421 agtgcgggga tatttgataa gaaacttcag tgaaggccgg gcgcggtggc tcatgccctg
 15 481 aatcccagca ttttcggagg ccgaggctgg agtgcaatgg tgtgatctca gctcactgca
 541 acctctgctt cctgggttta agtgattctc ctgcctcagc ctcccagata gctgggatta
 601 caggcatcat ggaccgatct aaagaaaact gcatttcagg acctgttaag gctacagctc
 661 cagttggagg tccaaaacgt gttctcgtga ctcagcaatt tccttgctag aatccattac
 721 ctgtaaatag tggccaggct cagcgggtct tgtgtccttc aaattcttcc cagcgcattc
 20 781 ctttgcaagc acaaaagctt gtctccagtc acaagccggg tcagaatcag aagcagaagc
 841 aattgcaggc aaccagtgtc cctcatcctg tctccaggcc actgaataac acccaaaaga
 901 gcaagcagcc cctgccatcg gcacctgaaa ataactctga ggaggaactg gcatcaaaac
 961 agaaaaatga agaatacaaaa aagaggcagt gggctttgga agactttgaa attggtcgcc
 1021 ctctgggtaa aggaaagtgt ggtaagtgtt atttggcaag agaaaagcaa agcaagttaa
 25 1081 ttctggctct taaagtgtta tttaaagctc agctggagaa agccggagtg gagcatcagc
 1141 tcagaagaga agtagaaata cagtccacc ttcggcatcc taatattctt agactgtatg
 1201 gttattttcca tgatgctacc agagtctacc taattctgga atatgcacca cttggaacag
 1261 tttatagaga acttcagaaa ctttcaaagt ttgatgagca gagaactgct acttatataa
 1321 cagaattggc aaatgccctg tcttactgtc attcgaagag agttattcat agagacatta
 30 1381 agccagagaa cttacttctt ggatcagctg gagagcttaa aattgcagat tttgggtggg
 1441 cagtacatgc tccatcttcc aggaggacca ctctctgtgg caccctggac tacctgcccc
 1501 ctgaaatgat tgaaggctcg atgcatgatg agaaggtgga tctctggagc cttggagttc
 1561 tttgctatga atttttagtt ggggaagcctc cttttgaggc aaacacatac caagagacct
 1621 acaaaagaat atcacgggtt gaattcacat tccctgactt tgtaacagag ggagccaggg
 35 1681 acctcatttc aagactgttg aagcataatc ccagccagag gccaatgctc agagaagtac
 1741 ttgaacaccc ctggatcaca gcaaattcat caaaaccatc aaattgcca aacaaagaat
 1801 cagctagcaa acagtcttag gaatcgtgca gggggagaaa tccttgagcc agggctgcca
 1861 tataacctga caggaacatg ctactgaagt ttattttacc attgactgct gccctcaatc
 1921 tagaacgcta cacaagaaat atttgtttta ctcagcaggt gtgccttaac ctccctattc
 40 1981 agaaagctcc acatcaataa acatgacact ctgaagtga agtagccacg agaattgtgc
 2041 tacttatact gggtcataat ctggaggcaa ggttcgactg cagccgcccc taagcctgt
 2101 gctaggcatg gtgtcttcac aggaggcaaa tccagagcct ggctgtggg aaagtacca
 2161 ctctgcccctg accccgatca gttaaaggagc tgtgcaataa ccttcctagt acctgagtga
 2221 gtgtgtaact tattgggttg gcgaagcctg gtaaagctgt tggaatgagt atgtgattct
 45 2281 ttttaagtat gaaaataaag atatatgtac agacttgat tttttctctg gtggcattcc
 2341 ttttaggaatg ctgtgtgtct gtccggcacc ccggtaggcc tgattgggtt tctagtcctc
 2401 ctttaaccact tatctcccat atgagagtgt gaaaaatagg aacacgtgct ctacctccat
 2461 ttagggtattt gcttgggata cagaagaggc catgtgtctc agagctgtta agggcttatt
 2521 tttttaaaac attggagtca tagcatgtgt gtaaaactta aatatgcaaa taaataagta
 50 2581 tctatgtcta aaaaaaaaa aaaaa [SEQ ID NO: 5]

Detailed Description

Brief Description of the Figures

The invention will now be described in detail with reference to examples and with reference to the accompanying drawings, in which:

5

Figure 1 shows the dynamics of ezrin cap formation

A Coimmunoprecipitation (CoIP) assays of total ezrin binding to NHERF1 in Caco-2 cells transfected by control nontargeting (NT) siRNA or PKCz siRNA. **B** Ezrin p-T567 cortical recruitment in Caco-2 cells transfected by control (NT) or PKCz siRNA. Assays at 3.5 h after plating. Reduced cortical intensity of Ezrin p-T567 shown in PKCz siRNA transfected cells. **C** NHERF1 cortical recruitment in Caco-2 cells transfected by control (NT) or PKCz siRNA. . Reduced cortical intensity of NHERF1 shown in PKCz siRNA transfected cells. For all cortical recruitment studies, cells were synchronized in G₀/G₁. At 3.5 h, most cells expressed the S phase marker EdU. Fifty EdU expressing cells were randomly selected and assessed in triplicate for each experimental condition. **D [i]** Co-immunoprecipitation assays of total ezrin/NHERF1 interaction in Caco-2 cells treated with a scrambled peptide (control) or the ezrin/NHERF1 peptide binding inhibitor (Ez/Nhe pbi). **D [ii]** Confocal assays of ezrin p-T567 cortical enrichment in Caco-2 cells treated with scrambled peptide (control) or Ez/Nhe pbi, at 3.5 h after plating. Reduced cortical intensity of Ezrin p-T567 shown in Ez/Nhe pbi treated cells. **E** Confocal assays of ezrin p-T567 cortical enrichment in Caco-2 cells transfected by control nontargeting (NT) siRNA versus NHERF1 siRNA KD, at 3.5 h after plating. Reduced cortical intensity of Ezrin p-T567 shown in NHERF1 siRNA transfected cells **F[i]** Effects of CK-666 treatment (100 μ M (Yi *et al. Nat Cell Biol* 2011; **13**: 1252-1258)) on ezrin cortical recruitment at 3.5 h (column 1) and on ezrin and actin cortical cap formation at 14 h (columns 2-5). Note CK-666 treatment does not affect Ezrin cortical recruitment at 3.5 h after synchronization but suppresses actin-driven Ezrin cap formation at 14h. **F[ii]** **A** Summary effects of CK-666 versus vehicle only control on ezrin p-T567 cortical recruitment at 3.5 h, p = NS and **B** on ezrin cap formation at 14 h; **p=0.01. Analysis by paired Student's t test. Staining - DAPI (blue) for nuclear DNA shown in cell centres (indicated by solid arrow in grayscale views), ezrin p-T567 (red; interrupted arrow) NHERF1 (red; interrupted arrow), EdU (green; solid arrow), actin (green; interrupted arrow), Ezrin or actin cap formation indicated by text, scale bar = 20 μ m.

Figure 2 shows effects of ezrin/NHERF1 interaction on multicellular morphogenesis

(A) Intestinal organoids - Left and right panels were stained to show mitotic spindle architecture and lumen formation respectively. Left panels show bipolar spindle orientation in control and Ez/Nhe pbi-treated organoids. High power spindle views (insets with thick border) show the orientation angles (interrupted white arrows) of spindle planes (double headed white arrows) towards gland centres [GC]. Right panels show lumen formation and epithelial configurations in control versus Ez/Nhe pbi-treated organoids. Multiple lumens and early epithelial stratification are indicated by solid and interrupted white arrows respectively in Ez/Nhe pbi-treated organoid cultures. **B** Summary spindle angles relative to GCs in control versus Ez/Nhe pbi-treated organoids shown in **A**, $**p < 0.01$; paired Student's t test (n=30 mitotic cells per experimental group). **C** Organotypic 3D Caco-2 cultures. Left and right panels are stained to show mitotic spindle architecture and lumen formation respectively. Left panels show bipolar spindle orientation in control and Ez/Nhe pbi-treated Caco-2 cultures. High power spindle views (insets with thick border) show orientation angles (interrupted white arrows) of spindle planes (double headed white arrows) towards gland centres [GC]. Right panels show lumen formation and epithelial configurations in control vs Ez/Nhe pbi-treated Caco-2 cultures. Multiple lumens and epithelial stratification are indicated by solid and interrupted white arrows respectively in Ez/Nhe pbi-treated Caco-2 cultures. **D** Summary spindle angles relative to GCs in control versus Ez/Nhe pbi-treated Caco-2 cultures shown in **C**, $**p < 0.01$. Analyses by paired Student's t test (n=30 mitotic cells per experimental group). Staining - DAPI (blue), α -tubulin (green), phalloidin (green), ezrin p-T567 (red in colour images as indicated in Figure 1). Scale bar = 20 μ m. Assays at 4 days of culture.

Figure 3 shows effects of PKCz on mitotic spindle architecture in cells with extra centrosomes

A Centrosome clustering (insets, top left of panel with border) and spindle architecture in control, PKCzI-treated (1 μ M) or PKCz siRNA transfected Caco-2 cells. **B** Summary mitotic spindle architecture data in control versus PKCzI-treated Caco-2 cells shown in **A**; Bipolar, $**p = 0.001$; Multipolar, $**p = 0.002$ and in control (NT siRNA) versus PKCz siRNA transfected Caco-2 cells shown in **A**; $**p = 0.002$ for bipolar and multipolar. **C** Centrosome clustering (inset) in control versus NHERF1 siRNA-transfected Caco-2 cells. **D** Summary spindle architecture data in control versus NHERF1 siRNA-transfected Caco-2 cells shown in **C**; Bipolar, $**p = 0.002$; Multipolar, $**p = 0.001$. **E** Centrosome clustering (insets) and spindle architecture in control, PLK4OE and PLK4OE + PKCzI-treated Caco-2 cells. **F** Summary multipolar spindle architecture data in control, Caco-2 versus PLK4OE versus PLK4OE + PKCzI-treated cells shown in **E**; Caco-2 versus PLK4OE versus PLK4OE + PKCzI-treated cells shown in **E** $**p < 0.001$; Control versus PLK4OE = NS, ANOVA, Tukey's

post hoc test (n = 100 mitotic cells in triplicate, expressed as %). Monopolar or indeterminate mitotic figures were counted but not analysed. Staining - DAPI (blue), pericentrin (red), α -tubulin (green) in colour images as indicated in Figure 1) Scale bars = 20 μ m.

5 Figure 4 shows effects of PKCz on chromosome segregation in cells with extra centrosomes

A Chromosome (Chr) 1 (green in colour images, indicated by solid arrows in grayscale) and 2 (red in colour images, indicated by interrupted arrows in grayscale) signals in control Caco-2 transfected with empty vector only versus PLK4OE versus PLK4OE + PKCz siRNA transfected Caco-2 cells. Chromosome fluorophores were counterstained against the DAPI DNA stain (blue). Note 2 x Chr1 and 2 x Chr2 signals in control and PLK4OE cells but 3 x Chr1 signals in Caco-2 PLK4OE + PKCz siRNA transfected cells. The inventors analysed 20 spreads per experimental condition. The inventors found > 2 Chr1 and/or >2 Chr2 signals in 1/20 control Caco-2 and Caco-2 PLK4OE spreads each. However, 4/20 Caco-2 PLK4OE-PKCz siRNA spreads showed > 2 Chr1 and/or >2 Chr2 signals on FISH assay. **B** Chromosome 19 (red in colour images, indicated by narrow solid arrows in grayscale) signals in control Caco-2 versus PLK4OE versus PLK4OE + PKCz siRNA transfected Caco-2 cells (n=30 cells per spread). The inventors found >2 Chr 19 signals in 9/30 control Caco-2, 10/30 PLK4OE and 14/30 Caco-2 PLK4OE + PKCz siRNA transfected Caco-2 cells. **C** Total chromosome number per spread in control Caco-2 versus PLK4OE versus PLK4OE + PKCz siRNA transfected Caco-2 cells, **p <0.01; ANOVA; Tukey's *post hoc* test; control Caco-2 versus PLK4OE = NS (n = 10 spreads per experimental condition). **D** Chromosome 19 signals and micronuclei in control Caco-2 versus PLK4OE versus PLK4OE + PKCz siRNA transfected Caco-2 cells. Micronuclei (blue with DAPI stain in colour images) lacking or containing a Chr 19 signal (red in colour images) are indicated by lined or solid white arrowheads, respectively. **E** Summary of Chr19 signals per micronucleus in control Caco-2 versus PLK4OE versus PLK4OE + PKCz siRNA transfected Caco-2 cells, **p=0.011; ANOVA with Tukey's *post hoc* test (Caco-2 control versus Caco-2 PLK4OE = NS). Scale bar = 20 μ m.

30

Figure 5 shows relationships between mitotic spindle geometry and multicellular morphology in 3D organotypic CRC cultures

A Confocal assays of spindle architecture (insets) and multicellular morphology in 3D organotypic cultures. Control Caco-2 cultures with appropriately orientated bipolar spindles (left panel) had regular 3D morphology with single central lumens surrounded by a uniform apical membrane and columnar epithelial monolayers. SiRNA knockdown of PKCz in 3D Caco-2 PLK4OE cultures induced multipolar spindle formation and solid cell-filled 3D

35

structures with dispersed apical membrane foci. These cultures either lacked any lumen (middle panel) or had aberrant noncentric lumens lying outwith gland centres, surrounded by atypical epithelium (right panel). Cells with multipolar spindles extended across the basal interface with extracellular matrix (ECM) more frequently than cells with bipolar spindles (see **Figure 5D**). **B** Nuclear “roundness” scores in glands with bipolar versus multipolar spindles, *** $p < 0.001$; paired Student’s t test. A score of 1 MRU denotes a perfect circle (Filippi-Chiela *et al. PLoS One* 2012; **7**: e42522) (n= 60 cells from glands containing bipolar or multipolar spindles). **C** Range of nuclear size in glands with bipolar versus multipolar spindles ($p < 0.01$; Levene’s test; n=30 bipolar or multipolar cells). **D** Summary extension of cells with bipolar or multipolar spindles across the ECM interface (denoted by interrupted line). Distances between spindle midpoints and the ECM interface were assessed. Positive or negative values were assigned for direction of extensions into or away from the ECM respectively. Positive distance values (into ECM) in multipolar versus bipolar spindles = 14/50 versus 1/50; * $p = 0.03$; paired Student’s t test. Staining - DAPI blue, for nuclear DNA; p-PKC red, for apical membranes, α -tubulin green, for microtubules in colour images. Assays at 4 days of 3D culture.

Figure 6 shows translational studies in archival colorectal cancer

A IHC assay of NHERF1 apical expression in archival CRC. Sections at 50X objective magnification with scores of 2, 1 and 0 respectively. Apical localization of NHERF1 is indicated by arrows in high power insets (thick borders). **B** Cells with 2 or >2 Aurora A spindle pole signals in archival CRC sections indicative of bipolar or multipolar spindles respectively (Herz *et al. Mol Carcinog* 2012; **51**: 696-710). Objective magnification x63. Staining - DAPI (blue), Aurora A (green) in colour images. **C** Relationship between multipolar spindle formation (mitotic cells with >2 Aurora A spindle pole signals) and apical NHERF1 intensity ($r = -0.452$); ** $p = 0.007$; Pearson’s test; Aurora A spindle pole signals assessed by IF in 180 ± 72 mitotic cells per tumour section in 35 CRCs. **D** Relationship between multipolar spindle frequency and cancer grade, ** $p = 0.005$. ANOVA; Tukey’s *post hoc* test. Multipolar spindles assessed as % of all mitotic figures. Scale bar = 20 μ m.

Figure 7 shows Dynamics of ezrin cap formation

A Quantification of total ezrin/NHERF1 binding shown in **Figure 1A** after PKC ζ SiRNA KD normalised to control nontargeting (NT) SiRNA = 0.66 ± 0.038 ; * $p = 0.0123$. **B** Immunoblots in Caco-2 cells after PKC ζ pseudosubstrate inhibitor (PKC ζ I 1 μ M (Jagan *et al. Neoplasia* 2013; **15**: 1218-1230) treatment or PKC ζ siRNA transfection vs vehicle only or nontargeting siRNA controls. Quantification of ezrin p-T567 ADU in Western blots shown in **Figure 7B** after PKC ζ I treatment (**C**) or PKC ζ siRNA KD (**D**) normalised to control.

C = 0.45 ± 0.05 ; ** $p = 0.009$; **D** = 0.54 ± 0.04 ; ** $p = 0.009$. **E** Schematic of ezrin cap formation, centrosome anchoring, replication and clustering during interphase. Ezrin (red in colour images) is recruited from the cytosol to the cortex, where it becomes progressively restricted to form the ezrin cap that binds astral microtubules (green in colour images). This anchoring process stabilizes the interphase centrosome (orange in colour images) for normal (single curved arrow) or abnormal (double curved arrow) replication and clusters any extra (encircled) centrosomes (Hebert *et al. Genes Dev* 2012; **26**: 2709-2723). Ezrin p-T567 **F(i)** and total ezrin **F(ii)** cortical recruitment and cap formation in Caco-2 cells at 3.5h and 14 h respectively after plating. **G** %Caco-2 cells with ezrin p-T567 cortical recruitment at 3.5 hours shown in **Figure 1B** (Control NT SiRNA vs PKCz SiRNA = 76.7 ± 4.1 vs $26.0 \pm 3.0\%$; * $p = 0.012$). **H** %Caco-2 cells with NHERF1 cortical recruitment at 3.5 hours shown in **Figure 1C** (Control NT SiRNA vs PKCz SiRNA = 84 ± 4.1 vs $34.7 \pm 2.9\%$; ** $p = 0.001$). **I** Merlin cortical localization in control or PKCzI treated Caco-2 cells at 14h after plating. **J** Effects of scrambled peptide control or Ez/Nhe pbi on total ezrin/NHERF1 binding shown in **Figure 1D(i)** Data expressed as fold change over scrambled peptide control = 0.66 ± 0.038 ; * $p = 0.012$. **K** Effects of peptide treatment on % cells with ezrin p-T567 cortical enrichment at 3.5h after plating shown in **Figure 1D (ii)** (Control vs Ez/Nhe pbi = 73.3 ± 4.21 vs $24.7 \pm 3.5\%$; ** $p = 0.0035$). **L** Immunoblot of NHERF1 expression after transfection of NHERF1 SiRNA vs nontargeting (NT) control SiRNA. **M** Confocal assays of ezrin and NHERF1 localization at 14h after plating in Caco-2 cells. NHERF1 does not localize at a cap. **N** Effects of PKCz siRNA KD (left panels) or Ez/Nhe pbi treatment (left panels) on ezrin cap formation in Caco-2 cells at 14h after plating. Analyses by paired Student's t test. Staining - DAPI (blue), ezrin p-T567 (red), merlin (red), total ezrin (green in colour images). Scale bars - 20 μm .

Figure 8 shows summary effects of ezrin/NHERF1 interaction on multicellular morphogenesis

A Summary single lumen formation in control vs Ez/Nhe pbi-treated organoids = 70.0 ± 5.7 vs $33.3 \pm 3.3\%$; * $p = 0.03$ shown in **Figure 2A** ($n = 30$ organoids per experimental condition in triplicate, expressed as %). **B** Nuclear "roundness" scores in control vs Ez/Nhe pbi treated organoids = 0.7456 ± 0.017 vs 0.674 ± 0.018 (measured roundness units [MRU]); * $p = 0.02$. **C** Nuclear area in control and Ez/Nhe pbi treated organoids shown in **Figure 2A** = 3326 ± 272 vs $3349 \pm 354 \mu\text{m}^2$; $p = \text{NS}$.

D Summary single lumen formation in control vs Ez/Nhe pbi-treated Caco-2 cultures shown in **Figure 2C** = 37.3 ± 2.4 vs $16.7 \pm 1.7\%$; ** $p = 0.004$. ($n = 30$ Caco-2 glands per experimental condition in triplicate, expressed as %). **E** Nuclear "roundness" scores in control vs Ez/Nhe pbi treated Caco-2 cultures shown in **Figure 2C** = 0.86 ± 0.014 vs 0.52 ± 0.02 MRU; *** p

<0.001. **F** Nuclear area in control and Ez/Nhe pbi treated Caco-2 cultures shown in **Figure 2C** = 1207 ± 61 vs $2198 \pm 159 \mu\text{m}^2$; * $p=0.012$; $n = 100$ cells per experimental condition in triplicate. All analyses by paired Student's t test.

5 Figure 9 shows summary effects of PKCz on mitotic spindle architecture in cells with extra centrosomes

A Summary clustering in control and PKCz siRNA transfected Caco-2 cells with >2 centrosomes, shown in Figure 3A = 40.33 ± 2.60 vs $21.0 \pm 2.08\%$; ** $p = 0.001$. **B** Centrosome clustering (inset) in control vs PKCzI-treated U2OS cells. **C** Summary centrosome clustering for control vs PKCzI treated U2OS cells with >2 centrosomes in B = 61.7 ± 4.40 vs $27.0 \pm 4.58\%$; * $p = 0.05$. **D** Summary centrosome (Csm) clustering in control vs NHERF1 siRNA-transfected Caco-2 cells with >2 centrosomes shown in **Figure 3C** = 42.7 ± 2.9 vs $21.3 \pm 1.3\%$; * $p = 0.03$. **E** Centrosome clustering (inset) in control vs NHERF1 siRNA transfected B549 cells. **F** Summary centrosome clustering in control vs NHERF1 siRNA-transfected B549 cells with >2 centrosomes shown in E = 60 ± 1.1 vs $23.3 \pm 1.7\%$; ** $p = 0.005$. **G** Mitotic spindle architecture in control vs Ez/Nhe pbi-treated Caco-2 cells. **H** Summary spindle architecture in control vs Ez/Nhe pbi-treated Caco-2 cells shown in G = Bipolar = 66.0 ± 5.3 vs $38.0 \pm 3.46\%$; * $p=0.01$; Multipolar = 0 vs $25.3 \pm 2.40\%$; ** $p < 0.001$. **I** Doxocycline-inducible PLK4 overexpression (PLK4OE) in Caco-2 and HCT116 cells. **J** Summary doxocycline-inducible PLK4 ADU in Caco-2 and HCT116 cells shown in I. **K** Centrosome number in control and PLK4 overexpressing (PLK4OE) Caco-2 and HCT116 cells. Assays at 24h after doxocycline-driven PLK4 overexpression. **L** % control vs PLK4OE Caco-2 or HCT116 cells with extra centrosomes shown in K; Caco-2 = 5.7 ± 1.2 vs $51.3 \pm 4.06\%$; ** $p = 0.006$; HCT116 = $0.7 \pm .7$ vs $50.7 \pm 3.7\%$; ** $p=0.007$. Clustering assessed $n=100$ cells with > 2 centrosomes in triplicate. Spindle architecture assessed in 100 mitotic cells in triplicate, expressed as a percentage. Analyses by paired Student's t test (A,C,D,F, L) or ANOVA (H). Staining - DAPI (blue), pericentrin (red); α -tubulin (green in colour images).

30 Figure 10 shows effectiveness of siRNA PKCz knockdown

A PKCz expression after siRNA knockdown in Caco-2 PLK4OE cells.

B Fold change in PKCz ADU induced by PKCz siRNA transfection in Caco-2 PLK4OE cells shown in **A**, normalized against control = 0.5 ± 0.08 ADU; ** $p=0.008$; paired Student's t test.

35 Figure 11 shows spindle architecture in control versus Ez/Nhe pbi-treated Caco-2 cultures and associated quantitative data

A Spindle architecture (insets) in control vs Ez/Nhe pbi-treated cultures. **B** Summary spindle architecture shown in **A**; Bipolar = 64.3 ± 3.5 vs $39 \pm 3.70\%$; $**p = 0.008$; Multipolar = 0 vs $19.3 \pm 2.9\%$; $**p = 0.003$. (100 mitotic cells assessed in triplicate in each experimental condition). Analysis by ANOVA. Staining DAPI (blue), pericentrin (red), α -tubulin (green in colour images).

Figure 12 shows a graphical summary of the effects of defective centrosome anchoring on evolution of CRC morphological and/or genomic phenotypes

(i) Ezrin cap (red in colour images) anchoring of the interphase centrosome through astral microtubule binding (green in colour images). Thus stabilized, the centrosome replicates to generate 2 normal or extra (encircled) centrosomes (Hebert *et al. Genes Dev* 2012; **26**: 2709-2723). **(ia1 and ia2)** show bipolar spindle assembly and normal orientation with normal **(ia1)** or clustered **(ia2)** centrosomes. **(ib-ic)** shows correct bipolar spindle orientation, normal cleavage furrows, apical membrane (AM; red) alignment, luminal secretion (arrows), multicellular assembly and gland formation (Jaffe *et al. J Cell Biol* 2008; **183**: 625-633). **(id-ie)** show representative images in the culture model and normal colon.

(ii) Ezrin cap deficiency with a single interphase centrosome. Impaired centrosome anchoring drives bipolar spindle misorientation **(iia)**, apical membrane AM (red in colour images) misalignment and aberrant cell division planes **(iib)** that collectively induce cribriform morphology **(iic)** in 3D cultures **(iid)**. Cribriform morphology shown in low-grade CRC **(iie)**.

(iii) Ezrin cap deficiency combined with extra centrosomes. Dispersal of multiple, unanchored centrosomes promotes transient multipolar spindle formation **(iiia)**, conversion to misorientated, pseudobipolar spindles by error-prone metaphase clustering mechanisms, accompanied by chromosome lag (Ganem *et al. Nature* 2009; **460**: 278-282) **(iiib1)**. A few cells with multipolar spindles undergo multipolar division to generate pleomorphic progeny **(iiib2)**. Segregation error associated with metaphase clustering **(iiia to iiib1 conversion)** promotes chromosomal instability (CIN). [Note whole chromosome (Chr) aneuploidy shown by 3 x Chr1 signals (green in colour images)]. CIN arising from these mechanisms is accompanied by nuclear pleomorphism, gross multicellular perturbation and extrusion of genomically unstable cells across basal interface with ECM, in the cartoon **(iiic)** and culture model **(iiid)**. Insets show bipolar or multipolar spindle architecture. Extrusion of malignant cells from main epithelial mass in high grade CRC shown in **(iiie)**.

Examples

Example 1: Dynamics of ezrin cap formation

5 Because PKC ζ regulates ezrin accumulation in the blastomere cell cortex and ezrin/NHERF1 interaction controls cortical retention of ezrin, the inventors investigated PKC ζ control of ezrin/NHERF1 interaction and ezrin cortical recruitment in Caco-2 cells. To investigate PKC ζ regulation of ezrin/NHERF1 interaction, the inventors immunoprecipitated NHERF1 from control and PKC ζ siRNA knockdown (KD) lysates. SiRNA PKC ζ KD led to reduced co-immunoprecipitations of total ezrin by NHERF1 (Figure 1A and Figure 7A). PKC ζ siRNA KD or treatment by a PKC ζ pseudosubstrate inhibitor (PKC ζ I) suppressed ezrin phosphorylation at T567 (Figure 7B-D), a key conformational switch that enables ezrin/NHERF1 binding and ezrin cortical enrichment. After cortical recruitment, ezrin becomes progressively restricted to form a pericentrosomal cap that anchors the interphase centrosome or clusters supernumerary centrosomes (Figure 7E). To study ezrin cortical recruitment and temporal restriction of ezrin within the cortex to form the cap, the inventors synchronized Caco-2 cells in G₀/G₁ by serum starvation and conducted confocal microscopy assays. Here the inventors have shown ezrin cortical recruitment and cap formation at 3.5 h and 14 h after plating respectively (Figure 7F [i] and [ii]). In addition, cortical dynamics of active ezrin p-T567 parallels that of total ezrin (Figure 7F [i] and [ii]). Since ezrin p-T567 was more easily detectable, it was assessed in most confocal experiments. SiRNA PKC ζ KD inhibited cortical recruitment of ezrin p-T567 (Figure 1B and Figure 7G) and NHERF1 (Figure 1C and Figure 7H). The inventors used the S phase marker 5-ethynyl-2'-deoxyuridine (EdU) to ensure equivalent cell cycle phases in experimental groups for NHERF1 studies (Figure 1C and Figure 7H). Merlin and ezrin are closely related and PKC ζ I treatment suppressed merlin cortical recruitment (Figure 7I). PKC ζ I treatment reduced the percentage of Caco-2 cells showing cortical localization of merlin from $32.7 \pm 2.90\%$ in control cells to $13.3 \pm 2.0\%$ after PKC ζ I treatment; $p = 0.02$. To investigate the role of ezrin/NHERF1 interaction in ezrin cortical recruitment, the inventors used a specific ezrin/NHERF1 peptide binding inhibitor (Ez/Nhe pbi) (Stokka *et al. Biochem J* 2010; **425**: 381-388) and conducted NHERF1 siRNA KD studies. Ez/Nhe pbi treatment inhibited the interaction between total ezrin and NHERF1 (Figure 1 D[i] and Figure 7J) and suppressed ezrin p-T567 cortical recruitment (Figure 1 D[ii] and Figure 7K). Furthermore, NHERF1 siRNA KD (Figure 7L) also inhibited ezrin p-T567 cortical recruitment (Figure 1E). Transfection of NHERF1 siRNA induced a fold reduction of NHERF1 protein expression of 0.55 ± 0.07 ; $p = 0.03$ and reduced the percentage of cells with ezrin p-T567 cortical

recruitment from $73.3 \pm 4.7\%$ in control cells transfected with nontargeting (NT) siRNA) to $34.6 \pm 4.1\%$ in NHERF1 siRNA transfectants, $p=0.018$.

Following ezrin cortical enrichment, mechanisms dependent upon actin filaments, merlin and α -catenin drive ezrin cortical restriction to form the pericentrosomal cap. Because ezrin localization within the microvillus cortex can be driven by Arp2/3-mediated actin treadmilling, the inventors suppressed Arp2/3 using the specific inhibitor CK-666 (Yi *et al. Nat Cell Biol* 2011; **13**: 1252-1258). Here the inventors have shown that CK-666 treatment did not affect ezrin cortical recruitment at 3.5 h but inhibited formation of both the ezrin cap and the actin cap at 14 h (Figure 1F[i], F[ii]). Both merlin and NHERF1 decorated the entire circumference of the cell cortex at 14 h (Figure 7I, M). Because siRNA PKC ζ KD or disruption of ezrin/NHERF1 binding by peptide inhibitor treatment or siRNA NHERF1 KD (Figure 7L) suppressed ezrin cortical recruitment (Figures 1B, D[ii], E and Figure 7G, K), the inventors tested effects of these interventions on ezrin cortical cap formation. All of these interventions suppressed ezrin cap formation (data shown for PKC ζ siRNA and Ez/Nhe pbi treatment only; Figure 7N). The percentages of cells with ezrin cap formation were $66.70 \pm 4.41\%$ (control) versus $32.0 \pm 2.0\%$ (PKC ζ siRNA) versus $7.0 \pm 1.53\%$ (Ez/Nhe pbi treatment); $p = 0.008$ or $p = 0.009$, respectively. Collectively, these data support a two-stage process of ezrin cap formation. Firstly, PKC ζ promotes ezrin phosphorylation to enhance ezrin/NHERF1 binding and ezrin cortical recruitment. Secondly, processes dependent on merlin and Arp2/3 drive ezrin cortical restriction. Perturbation of either ezrin cortical recruitment or restriction impede ezrin cap formation.

Example 2: Effects of ezrin/NHERF1 interaction on multicellular morphogenesis

The ezrin cap controls mitotic spindle dynamics that guide multicellular morphogenesis by well-characterized biological mechanisms. Here the inventors investigated ezrin/NHERF1 interactive effects on spindle dynamics and multicellular assembly in physiological and cancer models. The inventors used organoids formed from primary intestinal cells (Fatehullah *et al. Nat Cell Biol* 2016; **18**: 246-254) and 3D Caco-2 organotypic CRC model systems (Jagan *et al. Oncogene* 2013; **32**: 1305-1315). Acute perturbation of ezrin/NHERF1 interaction by peptide inhibitor treatment induced common effects of bipolar spindle misorientation, multi-lumen formation and epithelial stratification (Figure 2A-D and Figure 8A,D) that collectively induce cribriform multicellular morphology. To assess effects on cellular phenotypes, the inventors assessed nuclear roundness scores and nuclear size as indicators of pleomorphism (Saito *et al. J Pathol Inform* 2016; **7**: 36). While suppression of ezrin/NHERF1 interaction reduced nuclear roundness scores in both models (Figure 8B,E), it affected nuclear size only in the Caco-2 cells (Figure 8C,F). While ezrin and NHERF proteins have important roles in organization of cell membrane domains,

cell-cell and cell-extracellular matrix (ECM) communication, our studies reveal that ezrin/NHERF1 interaction is fundamental to morphogenic trajectories involving epithelial shape, configuration, spatial rearrangements and luminogenesis in physiological and cancer states through control of bipolar mitotic spindle orientation.

5

Example 3: Effects of PKCz on mitotic spindle architecture in cells with extra centrosomes

Increased centrosome number is a common cancer characteristic. Variable percentages of cells in Caco-2 and other cancer lines contain extra centrosomes. In Caco-2
10 cells, clustering of extra centrosomes at the ezrin cap enables bipolar spindle assembly. To investigate the role PKCz in these processes, the inventors conducted functional inhibition and/or siRNA knockdown studies against endpoints of multipolar spindle frequency and/or centrosome clustering. The inventors investigated Caco-2, U2OS and B549 cancer cells that are known to cluster extra centrosomes. PKCz treatment or PKCz siRNA knockdown
15 promoted development of multipolar mitotic spindles in Caco-2 cells (Figure 3A,B). These interventions also suppressed centrosome clustering not only Caco-2 cells but also in U2OS cells (Figure 9A-C). To investigate the role of ezrin/NHERF1 interactions downstream of PKCz in these processes, the inventors conducted siRNA knockdown studies or inhibited protein-protein interactions. SiRNA NHERF1 KD inhibited centrosome clustering (Figure 3C and Figure 9D) and promoted multipolar spindle architecture in Caco-2 cells (Figure 3C,D).
20 In keeping with this finding, siRNA NHERF1 KD also suppressed centrosome clustering in B549 cells (Figure 9E,F). Direct suppression of ezrin/NHERF1 interaction by peptide inhibitor treatment also prevented clustering and induced multipolar spindle formation in Caco-2 cells (Figure 9G,H).

25 Although extra centrosomes are causally implicated in multipolar spindle formation, the relationship appears non-linear. To investigate the association between extra centrosomes, PKCz and spindle defects, the inventors forced centrosome amplification in Caco-2 and in chromosomally stable, near diploid HCT116 cells by stable overexpression of PLK4 (Kleylein-Sohn *et al. Dev Cell* 2007; **13**: 190-202)(Figure 9 I,J). PLK4 overexpression
30 (PLK4OE) increased the percentages of Caco-2 and HCT116 cells with extra centrosomes (Figure 9K,L). While PLK4 overexpression (PLK4OE) caused only a modest increase in the frequency of multipolar spindle formation in Caco-2 cells, PLK4OE combined with PKCz functional inhibition induced a substantively higher frequency of the multipolar spindle phenotype (Figure 3E,F). Taken together, these data show that PKCz ameliorates effects of
35 centrosome amplification via ezrin/NHERF1 interactions, interphase centrosome clustering at the ezrin cap and suppression of multipolar spindle formation in cancer cells.

Example 4: Effects of PKCz on chromosome segregation in cells with extra centrosomes

The properly assembled bipolar mitotic spindle coordinates intracellular forces that drive equal genome partitioning. Conversely, multipolar spindle formation invokes segregation error either by progression through anaphase or by activation of different centrosome clustering mechanisms during metaphase that associate with merotelic attachments, missegregation and chromosomal instability (CIN). To investigate the role of PKCz on chromosome segregation in cells with extra centrosomes, the inventors conducted PKCz siRNA knockdown studies in stable PLK4 overexpressing Caco-2 (Caco-2 PLK4OE) cells. The inventors conducted fluorescent *in situ* hybridization (FISH) assays of chromosome 1 and 2 because these chromosomes are large and suitable for assessment of chromosomal rearrangements. Conversely, the inventors also studied chromosome 19 because it is small, frequently found in micronuclei and contains CRC susceptibility loci (Carvajal-Carmona *et al. Hum Mol Genet* 2011; **20**: 2879-2888). Here the inventors have shown that siRNA KD of PKCz in Caco-2 PLK4OE cells (Figure 10A,B) induced aneuploidy of chromosome 1 (Figure 4A), chromosome 19 (Figure 4B) and increased total chromosome number (Figure 4C). The inventors found micronuclei in 24/300 control versus 28/300 PLK4OE versus 42/300 PLK4OE + PKC siRNA transfected Caco-2 cells. Some micronuclei contained chromosome 19 signals (Figure 4D). SiRNA PKCz KD increased chromosome 19 signals within micronuclei (Figure 4E) in Caco-2 PLK4OE cells. Taken together, these data indicate that PKCz knockdown in cells with extra centrosomes induces errors in genome partitioning including CIN and chromosome missegregation into micronuclei.

Example 5: Relationships between mitotic spindle geometry and multicellular morphology in 3D organotypic CRC cultures

In 3D cancer models, morphological adaptations are partly driven by bipolar mitotic spindle dynamics through control of abscission, cytokinesis and multicellular assembly. Conversely, associations between multipolar spindle architecture and cancer morphology remain unclear. In this study, interventions that induced formation of multipolar spindles in Caco-2 PLK4OE cell monolayers also did so in 3D Caco-2 cultures. For example, siRNA PKCz KD or Ez/Nhe pbi treatment induced multipolar spindle formation in 3D Caco-2 PLK4OE glandular structures [glands] (Figure 5A and Figure 11A,B). Multipolar spindle formation induced by siRNA PKCz KD associated with development of heterogeneous multicellular morphology. Phenotypic alterations included solid 3D multicellular structures with absent lumens or glands with multiple or noncentric lumens as well as atypical epithelial organization (Figure 5A). These changes were accompanied by nuclear pleomorphism evidenced by reduced nuclear roundness scores (Figure 5B) and a wide range of nuclear

size (Figure 5C) in 3D glands. Furthermore, cells with multipolar spindles frequently extended across the basal interface with extracellular matrix (ECM) in 3D cultures (Figure 5A,D). While Ez/Nhe pbi treatment induced multipolar spindle formation it also suppressed growth of 3D Caco-2 PLK4OE glands and thus hampered analysis of multicellular morphology. Collectively, our studies show that multipolar spindle formation induced by PKCz knockdown in cells with extra centrosomes induced CIN, nuclear pleomorphism, aberrant multicellular morphology and spatial outgrowth of genomically unstable cells across the ECM interface. In combination, these phenotypes are evocative of aggressive, high-grade CRC.

Example 6: Translational studies in archival colorectal cancer

Our findings link perturbations of interphase centrosome anchoring at the cell cortex to cancer phenotype anomalies in human CRC model systems. To integrate analyses from CRC models with primary human tumours, the inventors conducted immunohistochemical (IHC) and immunofluorescence (IF) studies in archival CRC tissues. Apical NHERF1 intensity provides robust readout of PKCz morphogenic activity in 3D cultures and has been used previously as indirect readout in archival CRCs (Jagan *et al. Oncogene* 2013; **32**: 1305-1315). Here the inventors investigated NHERF1 intensity in two CRC sample collections. Sample (A) comprised 35 whole tumour sections and 5 matched normal mucosa specimens and sample (B) was a tissue microarray (TMA) comprising 309 tumour CRC specimen cores, derived from 92 CRCs (Deevi *et al. Oncotarget* 2016; **7**: 49042-49064). In sample A, the inventors assayed apical NHERF1 intensity by IHC (Figure 6A) and mitotic spindle architecture by Aurora A IF assays of spindle poles (Figure 6B) (Herz *et al. Mol Carcinog* 2012; **51**: 696-710). Apical NHERF1 intensity inversely related to the frequency of mitotic cells with multipolar spindle architecture (Figure 6C), defined by >2 Aurora A positive spindle pole signals (Figure 6B), in CRC sections. In view of the small size of sample A, the inventors also conducted semi-quantitative assays of apical and total NHERF1 intensity by IHC in tissue microarrays of sample B. The inventors found a positive correlation between total and apical NHERF1 intensity ($r = 0.504$; $p < 0.01$), although apical NHERF1 scores had a stronger inverse relationship with lymph node metastases ($p < 0.001$; data not shown). The inventors found that multipolar spindle frequency defined by Aurora A IF in CRC tissue sections, directly related to tumour grade (Figure 6D). Hence, these studies show that defective apical localization of NHERF1, a key component of centrosome anchoring machinery, associates with multipolar spindle architecture and high-grade morphology in human CRC.

In summary, the inventors have shown that impaired cortical control of single or supernumerary interphase centrosomes provide unifying rationale for cell shape, genome

segregation and multicellular pattern phenotypes that characterize low- or high-grade colorectal cancer (a graphic summary is shown in Figure 12).

Example 7: Summary

Mechanisms that integrate morphological and genomic phenotypes in colorectal cancer (CRC) represent a fundamental knowledge gap in pathology. In this study, the inventors have shown that PKC ζ couples genome segregation to multicellular assembly by control of interphase centrosome anchoring. Furthermore, the inventors reveal genomic, cytological and morphological consequences of perturbation (summarized in Figure 12).

Within the cell cortex, a polarized ezrin cap promotes anchoring and/or clustering of interphase centrosomes, guides mitotic spindle orientation and multicellular assembly. Here the inventors identify discrete steps in ezrin cap formation. Ezrin phosphorylation at T567 unmask binding domains, enhances ezrin/NHERF1 interaction and ezrin cortical enrichment. The inventors have shown that PKC ζ promotes ezrin T567 phosphorylation and increases ezrin/NHERF1 interaction. While ezrin/NHERF1 binding is required for maintenance of active ezrin at the cell cortex, the role of this molecular interaction in ezrin cortical recruitment remained unclear. This study shows that suppression of ezrin/NHERF1 interaction by peptide inhibitor treatment (Ez/Nhe pbi) blocked ezrin cortical recruitment. Furthermore, these effects on ezrin cortical recruitment were phenocopied by siRNA knockdown of NHERF1. Merlin (NF2) shares common ancestry with ezrin and interacts with both ezrin and NHERF1. The inventors further show that functional inhibition of PKC ζ suppresses merlin cortical recruitment. Collectively, these findings indicate that PKC ζ promotes ezrin phosphorylation, enhances ezrin/NHERF1 binding and promotes cortical enrichment of ezrin, NHERF1 and merlin.

In this study, the inventors have shown cortical recruitment and cap formation of ezrin p-T567 and total ezrin at 3.5 h and 14 h respectively after plating. Within the cortex, ezrin restriction to form the cap depends upon actin and merlin but is independent of Myosin II motor activity. Unlike merlin, ezrin directly binds actin via a conserved C-terminal binding domain and ezrin and actin cap formation develop in parallel. The inventors have shown that inhibition of the actin nucleator Arp2/3 by CK-666 treatment suppressed formation of ezrin and actin caps without affecting ezrin cortical recruitment.

In addition to genome partitioning, spindle dynamics also provide spatial directives for morphogenic processes. To aid understanding of their role in cancer, the inventors investigated morphological trajectories in physiological and cancer models and focussed on commonalities as well as differences. Suppression of ezrin/NHERF1 interaction by inhibitory peptide treatment induced bipolar spindle misorientation, aberrant epithelial stratification and multi-lumen formation in both organoid and Caco-2 CRC models. These phenomena lead to

cribriform morphology (CM) over longer term culture intervals of up to 12 days. While CM is regarded as a marker of malignant transformation in human colon (Brown *et al. J Clin Pathol* 2016; **69**: 292-299), this study shows early features of this morphology in normal intestinal organoids when the ezrin/NHERF1 interaction was disrupted. Hence, these data suggest that cribriform morphogenesis is a consequence of bipolar spindle misorientation but is not necessarily restricted to malignant cells. Without wishing to be bound by any particular theory, the inventors suspect that in cancer, CM may reflect cumulative mutational silencing of core intrinsic regulators of mitotic spindle orientation. In the cancer model, acute perturbation of the ezrin/NHERF1 interaction had greater effects on nuclear roundness scores and nuclear area, possibly in association with intrinsic unidentified mutations.

In cells with extra centrosomes, the ezrin cap promotes clustering during interphase to enable bipolar spindle assembly. Conversely, failure of interphase centrosome clustering allows centrosome dispersal and multipolar spindle formation. To investigate the role of PKCz in centrosome clustering and spindle architecture, the inventors forced centrosome amplification by PLK4 overexpression (PLK4OE)(Kleylein-Sohn *et al. Dev Cell* 2007; **13**: 190-202). Caco-2 cells accommodated a large rise in PLK4-induced centrosome number with only a small increase in multipolar spindle formation. Conversely, functional inhibition of PKCz led to increased multipolar spindle frequency. To further investigate connectivity between PKCz, ezrin and NHERF1, the inventors conducted siRNA knockdown studies or inhibited ezrin/NHERF1 interaction by peptide treatment. NHERF1 knockdown or inhibitory peptide treatment suppressed ezrin cap formation, prevented centrosome clustering and promoted the multipolar spindle phenotype. Hence, PKCz controls mitotic spindle architecture by regulation of ezrin/NHERF1 linkage and ezrin cortical dynamics.

Genome transmission and mitotic spindle assembly are intrinsically linked. This study shows that impaired cortical anchoring of the normal interphase centrosome promotes bipolar spindle misorientation, abnormal epithelial configuration and mislocalization of apical membrane markers. While these changes promote development of cribriform morphology, they also enabled error-free chromosome segregation in our model system. In contrast to previous work in a *Drosophila* model, transition to multipolar spindle formation in the Caco-2 system promoted chromosomal instability (CIN). Without wishing to be bound by any particular theory, the inventors suspect that whole mis-segregated chromosomes may become incorporated within micronuclei in inverse proportion to their size and may drive chromothripsis, a major mutagenic phenomenon. The inventors have shown that PLK4OE combined with PKCz knockdown increased whole chromosome number as well as the number of chromosome 19 signals within micronuclei. These data indicate that effective clustering of extra centrosomes during interphase by PKCz-dependent cortical machinery

inhibits multipolar spindle formation, suppresses CIN and chromosomal misincorporation into micronuclei that can trigger complex genomic rearrangements.

In cancer, aberrant morphology is classified within grading systems to provide the best-established predictors of clinical outcome. Key features of high-grade aggressive cancer include gross nuclear pleomorphism, aberrant mitotic figures and loss of glandular architecture. Here the inventors have shown that suppression of PKCz-dependent cortical machinery in cells with extra centrosomes drove development of these high-grade cancer phenotypes in 3D cell model systems. All Caco-2 PLK4OE glands with multipolar spindles induced by siRNA PKCz KD developed very abnormal morphology. None had a single central lumen and most comprised solid, cell-filled structures with widely dispersed apical membrane foci. These foci varied in size, probably from differences in transapical secretion. In larger foci, the ectopic apical membrane enclosed discernible lumens that the inventors termed noncentric since they were not situated in gland centres. The inventors described glands as cell-filled if they contained only cells and no lumen-like structures. Ez/Nhe pbi treatment drove multipolar spindle formation but also appeared to suppress growth of Caco-2 PLK4OE glands and thus hampered interpretation of multicellular morphology. Precise mechanisms of Ez/Nhe pbi growth suppression remain unclear but could be related to robust inhibition of ezrin cortical recruitment. Malignant cell detachment from the main tumour mass is a key metastatic process and is a common histological feature in high-grade cancer sections. In a Drosophila tumour model, cell extrusion across basement membrane and early invasion can be driven by chromosomal instability (CIN). In accord with those findings, the inventors have shown that multipolar spindle formation promotes CIN and cell extension across extracellular matrix interface, in organotypic 3D CRC cultures.

Correlative analyses in archival human tumour samples may shed light on experimental discoveries. Apical NHERF1 intensity provides readout of PKCz morphogenic signalling in 3D organotypic Caco-2 glands and has previously been used as indirect readout in paraffin fixed tissues. In the present study, apical NHERF1 IHC intensity inversely related to multipolar spindle formation, defined by Aurora A IF. Furthermore, multipolar spindle frequency directly associated with aberrant multicellular morphology of high-grade CRC. Hence, without wishing to be bound by any particular theory, the inventors suspect that defective PKCz cortical signalling reflected by low apical NHERF1 intensity may underlie mitotic errors and aberrant multicellular morphology that characterize aggressive CRC.

This study uncovers core molecular machinery that controls centrosome anchoring, mitotic spindle geometry, genome segregation and multicellular assembly. Without wishing to be bound by any particular theory, perturbation of these processes by diverse oncogenic pressures may provide a phylogenetic basis for branched evolution of genomic and morphological phenotypes underlying cancer trajectories to more aggressive subtypes.

Example 8: General Materials and Methods

5 **a. Reagents and antibodies** All laboratory chemicals were purchased from Sigma-Aldrich, Dorset, England unless otherwise stated.

10 **b. Organotypic and organoid cultures** Caco-2, BT-549 and U2OS cells were obtained from ATCC, Middlesex, England. Caco-2 cells were grown in three-dimensional (3D) organotypic cultures. Organoids of normal intestinal epithelium were isolated as the inventors have previously described (Tait *et al.* Cell Transplant 1994; 3: 33-40; Patel *et al.* Gut 1996; **38**: 679-686) and cultured in Matrigel matrix (Corning Inc, NY, USA Product No #354230) by a modification of previously described method (Sato *et al.* Gastroenterology 2011; 141: 1762-1772). Caco-2 were also grown as monolayers as were other cell types.

15 **c. Stable and transient transfections** The inventors carried out mammalian SiRNA and plasmid DNA transfections using RNAiMAX and X-tremeGENE transfection reagents (ThermoFisher, Dublin, Ireland) respectively, as the inventors have previously described (Javadi *et al.* Elife 2017; **6**). Lentiviral vector transfections were conducted using Lipofectamine 2000 (ThermoFisher, Dublin, Ireland) according to manufacturer's protocols. Stable clones were selected in Blasticidine (ThermoFisher, Dublin, Ireland). Overexpression of PLK4 encoded by the lentiviral system was induced by doxycycline treatment (Shearer & Saunders. *Genes Cells* 2015; **20**: 1-10).

25 **d. Inhibition of intracellular protein:protein interactions** To study biological effects of ezrin-NHERF1 interactions, cells were incubated with a cell-permeant disruptor peptide of the ezrin binding domain in NHERF1 (KERAHQKRSSKRAPQMDWSKKNELFSNL) (Stokka *et al.* Biochem J 2010; **425**: 381-388) or a control nontargeting peptide (KERAHQKRSSKRAPQMDASKANELASNL). The peptides were synthesized by EZBiolab, Carmel IN, USA.

30 **e. Fluorescent in situ hybridization (FISH) assays of chromosome segregation** In separate experiments, two colour FISH assays were performed in separate experiments using centromeric probes for chromosomes 1 (green) and 2 (red) or chromosome 19 (red) in separate experiments (Carl Zeiss, Cambridge UK, XCP

Human WCP probes). Assays of chromosome 19 mis-segregation into micronuclei were also conducted (Leach & Jackson-Cook *Mutation research* 2001; **495**: 11-19).

f. **Confocal imaging** Assays of cell cortex dynamics, centrosome disposition, mitotic spindle orientation and geometry, nuclear pleomorphism and multicellular patterns were conducted using a Leica SP5 confocal microscope, on a HCX PL APO lambda blue 63x 1.40 oil immersion objective at 1x or 2x zoom as the inventors have previously described (Javadi *et al. Elife* 2017; **6**).

g. **Human tumour samples** Anonymized formalin-fixed, paraffin-embedded (FFPE) colorectal primary tumours from previously described study cohorts (Deevi *et al. Oncotarget* 2016; **7**: 49042-49064) were released from the Northern Ireland Biobank (NIB), which has ethical approval to collect, store and distribute anonymized tissue samples to researchers by an approved protocol.

h. Full details of cell culture methods, transfection, FISH assays, Ez/Nhe pbi and control peptide sequences, confocal imaging, human tumour samples and Ethical approval reference numbers are provided below:

Example 9: Reagents and antibodies

Antibodies used in this study were mouse anti-PKCz (Abcam, Cambridge UK), rabbit anti-phospho-PKCz [T560] (Abcam), rabbit anti-PLK4 antibody (Abcam), mouse anti-NHERF1, (Lifespan Biosciences, Seattle USA), mouse anti-ezrin (Abcam), rabbit anti-phospho-ezrin [T567] (Abcam), rabbit anti-pericentrin (Abcam), mouse anti- α -tubulin (Abcam), FITC-Phalloidin and FITC-tubulin (Sigma-Aldrich, Dorset, UK), mouse anti-neurofibromin (also known as Merlin - Santa Cruz, Heidelberg, Germany) and mouse anti-GAPDH (glyceraldehyde-3-phosphate dehydrogenase (Abcam). These primary antibodies were used where appropriate in conjunction with Li-Cor IR Dye680 (anti-rabbit) and IR Dye800 (anti-mouse) secondary antibodies, for use with the Li-Cor Infra-Red imaging systems (Li-Cor Biosciences, Lincoln, Nebraska, USA) in Western blots or with Alexa Fluor 568 (anti-rabbit) and Alexa Fluor 488 (anti-mouse) (Molecular probes, Invitrogen, Carlsbad, CA, USA) for fluorescence or confocal microscopy. Anti- Aurora A antibodies (clone 4 BD Transduction Laboratories™, Heidelberg, Germany) were used in 3D immunofluorescence microscopy of human tissues. RNAiMAX and X-tremeGENE transfection reagents were purchased from Thermofisher, Dublin, Ireland and Roche, Basel, Switzerland, respectively. We suppressed PKCz by siRNA knockdown in most experiments. In some conditions however, e.g. cells that are already transiently transfected, a double transfection to achieve siRNA knockdown can

be stressful to cells. We therefore used a PKC ζ pseudosubstrate inhibitor [PKC ζ I] that has been extensively validated in various cell types (Dekanty *et al. Proc Natl Acad Sci U S A* 2012; **109**: 20549-20554; Yang & Hinds *Mol Cell* 2003; **11**: 1163-1176; Duensing *et al. Environ Mol Mutagen* 2009; **50**: 741-747) (P1614 Sigma; MDL number MFCD03458229) for all PKC ζ functional inhibition studies. The cell-permeant disruptor peptide of the ezrin binding domain in NHERF1 [Ezrin/NHERF1 peptide binding inhibitor (Ez/Nhe pbi) and control peptides] were synthesized by EZBiolab, Carmel IN, USA. The cell permeant disruptor peptide and control sequences were KERAHQKRSSKRAPQMDWSKKNELFSNL and KERAHQKRSSKRAPQMDASKANELASNL respectively, as previously described (Shearer & Saunders. *Genes Cells* 2015; **20**: 1-10). HPLC and Mass spectroscopy analyses indicated >95% purity. The peptides were supplied as a lyophilized powder and then dissolved in PBS, in accord with manufacturer's instructions.

Example 10: Cell culture

Caco-2, BT-549 and U2OS were obtained from ATCC, Middlesex, England. HCT116 cells were a gift from Dr Tod Waldman (Georgetown Q34 University, USA). Caco-2 cells were cultured in DMEM supplemented with 10% fetal calf serum (FCS), 1mM non-essential amino acids and 1mM L-glutamine at 37°C in 5% CO₂. HCT116 cells were cultured in McCoy's 5A medium supplemented with 10% foetal calf serum, 1mM sodium pyruvate and 1mM L-glutamine. U2OS and BT-549 cells were cultured in DMEM with 10% FCS. In 3D cultures Caco-2 cells and subclones stably transfected with empty vector (EV) only or doxocycline (Doxy) - inducible polo-like kinase 4 (PLK4) (see supplementary material) were cultured embedded in "Matrigel" matrix (BD Biosciences, Oxford, UK), as previously described (Jagan *et al. Neoplasia* 2013; **15**: 1218-1230; Chambers & Bretscher *Biochemistry* 2005; **44**: 3926-3932). In brief, 6 x 10⁴ trypsinized cells were mixed with Matrigel (40%) in a final volume of 100 μ l, which was plated into each well of an 8 chamber slide, allowed to solidify for 30 min at 37°C and subsequently overlaid with 400 μ l of media/well.

Example 11: Oligonucleotides

Dharmacon SmartPool siRNA oligonucleotides targeted against PKC ζ , NHERF1 and non-targeting control siRNAs were purchased from ThermoFisher Scientific, Dublin, Ireland.

Example 12: Cell transfection

Small interfering RNA (siRNA) and DNA transfections were carried out using RNAiMAX and X-tremeGENE transfection reagents (ThermoFisher, Dublin, Ireland) respectively. Cells were plated at 2x10⁵ cells/35 mm dish for 24 h, then transfected with 10 μ M siRNA or 500ng DNA/2x10⁵ cells for all respective siRNA oligonucleotides or DNA constructs. Transfections

were carried out according to the manufacturer's protocols. Cells were incubated with RNA/RNAiMAX or DNA/X-tremeGENE lipofectamine complexes for 24 or 48 h, before lysis and probing, as previously described (Jagan *et al. Neoplasia* 2013; **15**: 1218-1230).

5 **Example 13: Protein extraction and Western blotting**

Proteins were resolved using gel electrophoresis, followed by blotting onto nitrocellulose membranes. Membranes were probed using antibodies as indicated in the text.

Example 14: Co-immunoprecipitation (Co-IP)

- 10 Cells were lysed on ice in buffer containing 100 mM Tris-HCl, pH 7.5, 1% Triton X-100, 5 mM EDTA, 50 mM NaCl, 5 mM NaF, 1 mM Na₃VO₄ and protease inhibitor. Cell lysates were centrifuged (for 10 min at 15,000g) and protein concentrations were measured by the BCA method. 1000 µg of protein was precleared overnight with control IgG and 15 µl of Protein A/G Sepharose beads (Santa Cruz, Dallas, Texas, USA). The protein was then
- 15 immunoprecipitated with the appropriate antibody-beads conjugate and incubated on a rotating wheel for 2 h. The beads were collected by centrifugation and washed five times in wash buffer (50 mM HEPES, pH 7.4, 1 % Triton X-100, 0.1 %, SDS, 150 mM NaCl, 1 mM Na₃VO₄). The beads were subsequently resuspended in 40µl Laemmli sample buffer supplemented with 3 µl of 1 M DTT. Suspensions were boiled for 10 min, centrifuged and
- 20 the supernatant was used for gel electrophoresis.

Example 15: Lentiviral vectors

- To generate the inducible PLK4 overexpression system we used lentiviral vectors pLenti-CMV-TetR-Blast (17492, Addgene) and pLenti-CMV/TO-Neo-Dest (17292, Addgene), as
- 25 previously described (Shumilov *et al. Nat Commun* 2017; **8**: 14257)(Gift from Dr S. Godinho, Bart's Cancer Institute, London). Stable clones were treated with a dose range of doxycycline for 48 h to optimise transgene expression as previously described (Shumilov *et al. Nat Commun* 2017; **8**: 14257), for use in subsequent experiments.

30 **Example 16: Imaging of cell monolayers**

- For assays of cortical protein expression and ezrin cap formation, cells were cultured in 6 well plates. For transfections, cells were incubated in transfection reagents containing test or control constructs for 24 h. Cells were then harvested and placed in Matrigel coated 8 well multichambers and imaged at intervals up to 14 h after plating, as previously outlined
- 35 (Hebert *et al. Genes Dev* 2012; **26**: 2709-2723). For treatments, cells were cultured in 6 well plates as outlined above then harvested and placed in Matrigel coated 8 well multichambers containing test or control treatments. Imaging was conducted at intervals up to 14 h after

plating as outlined above. For assays of centrosome and mitotic spindle dynamics, cells were cultured in 6 well plates as outlined above then synchronized by double thymidine block, as previously outlined. Cells were then transfected as outlined above or treated and cultured for 24 h. Cells were fixed with 2% paraformaldehyde (PFA) then immunostained with appropriate antibodies at 4°C overnight. Cortical protein recruitment was defined by confocal demonstration of protein localization to the cell cortex beneath the cell surface, in excess of the protein quantity dispersed throughout the cytoplasm (Li *et al. Nat Commun* 2017; **8**: 14866). Ezrin cap formation was defined as a local ezrin accumulation to form a cap-like structure at one pole of the cell cortex of spherical cells, prior to the first cell division (Hebert *et al. Genes Dev* 2012; **26**: 2709-2723). Cortical protein expression, ezrin cap formation, centrosome number, disposition and spindle architecture were imaged by confocal microscopy and quantified by ImageJ in triplicate, for each experimental condition.

Example 17: Assessment of cortical architecture and ezrin cap formation

Labelled cells were visualized using a Leica SP5 laser scanning confocal microscope. Ezrin cap formation was defined as a local accumulation of ezrin signal within the cortex at one cell pole, as previously defined (Hebert *et al. Genes Dev* 2012; **26**: 2709-2723). We assessed ezrin p-T567 as readout of active ezrin (Hebert *et al. Genes Dev* 2012; **26**: 2709-2723). Signal intensity within the cortex was assessed using ImageJ.

Cortical localization assays of NHERF1 after nontargeting or PKCz SiRNA transfection were conducted using the cell cycle S phase marker 5-ethynyl-2'-deoxyuridine (EdU) (Filippi-Chiela *et al. PLoS One* 2012; **7**: e42522), using the Click IT protocol (ThermoFisher, Dublin, Ireland) as per manufacturer instructions. Briefly, cells were transfected with either NT or PKCz SiRNA. Cells were then incubated with EdU reagent for 24 h and fixed with 3.7% formaldehyde in PBS for 15 mins at room temperature. Cells were washed, permeabilised with 0.1% Triton® X-100 in PBS. Freshly prepared Click IT reaction mix was added to cells and incubated in the dark for 30 min at room temperature. Cells washed with Click-iT® rinse buffer and incubated with NHERF1 antibody. DAPI was used as the nuclear counterstain.

Example 18: Assessment of centrosome clustering and spindle architecture

In cultured cells, centrosomes and microtubules (MTs) were identified by confocal microscopy using anti-pericentrin and anti α -tubulin antibodies, respectively while chromosomal DNA was imaged using 4969-diamidino-2-phenylindole (DAPI) staining. A bipolar mitotic spindle was defined by convergence of MTs towards each of 2 spindle poles while a multipolar spindle was defined by >2 poles combined with abnormal DNA separation patterns. Centrosomes in excess of 2 per cell were defined as supernumerary. Mitotic cells

were scored for clustered (> 1 centrosome at spindle poles) or unclustered centrosomes and for multipolar spindles using a Leica SP5 confocal microscope.

Example 19: Chromosome Spreads and FISH Analysis

- 5 Control, transfected and treated cells were incubated in 10µg/mL Karyomax Colcemid solution (0.05µg/ml; ThermoFisher Scientific, Dublin, Ireland) for 18h. Following this, cells were harvested and metaphases were collected by resuspending in hypotonic 75 mM KCl for 20 min at 37 °C, followed by fixation for 20 min at 4 °C in freshly prepared Carnoy solution (3 : 1 v/v methanol/acetic acid). After two more washes in Carnoy solution, cells
- 10 were dropped onto pre-warmed wet slides and air-dried at room temperature and aged at room temperature for 7 days. Aged slides were hybridised with whole chromosome fluorescence-labelled DNA probes (XCP, Whole-Chromosome Probe, MetaSystems) directed to chromosomes 1 (fluorochrome FITC) and chromosome 2 (fluorochrome Texas Red) and chromosome 19 ((fluorochrome Texas Red) as per manufacturer's instructions.
- 15 DNA denaturation (72 °C for 3 min) and hybridisation (37 °C for 8h) were performed using the HYBrite chamber system (Vysis). Slides were washed with 0.4×SCC at 72°C for 2 min and then with 2×SCC and 0.05% Tween-20 at room temperature for 30 s, and mounted in Prolong Gold containing DAPI for chromosome counterstaining (Molecular Probes, Oregon, USA). Slides were imaged using a Nikon Eclipse TiNS fluorescence microscope, with a 60x
- 20 objective.

Example 20: Intestinal organoid cultures

- C57B/6 wild-type mice (<10 weeks old) were used for experiments. All animal procedures were conducted in accordance with local and national regulations. Organoids were isolated
- 25 by modifications (Dominguez *Mol Cell Biol* 1992; **12**: 3776-3783) of previously described methods (see Example 8 for details)(Standaert *et al. Biochemistry* 2001; **40**: 249-255). Briefly, murine intestines were opened longitudinally, cut into 0.5 cm fragments, washed 7-10 times in 1x HBSS (low calcium, low magnesium (Gibco-BRL), 2%D-glucose, 0.035% NaHCO₃) to remove all luminal contents. The fragments were then finely chopped with a
 - 30 scalpel and digested in HBSS solution containing collagenase and dispase I neutral proteases (Sigma-Aldrich, Dorset, UK) at 1 mg/ml for 20 min at room temperature on a shaking platform. Digestion was stopped by the addition of 30 ml DMEM/F12 culture medium (Life Technologies, Renfrew, UK) supplemented with 5% FCS containing penicillin and streptomycin. Large fragments and muscle sheets were allowed to settle to the bottom of the
 - 35 flask. Supernatant containing the organoids was centrifuged for 3 min at 250 rpm, to pellet the organoids. The supernatant was removed and the organoid pellet was gently resuspended in 20 ml of the DMEM/F12 solution. The centrifugation step was repeated 5-6

times until the pellet contained a homogeneously sized organoid preparation. Organoids thus prepared were resuspended in a 2x volume of Matrigel (growth factor reduced, phenol red free; BD Biosciences, Oxford UK) supplemented with 50 ng/ml murine EGF, murine Noggin 100 ng/ml and 1µg/ml human R-Spondin (PeproTech, NJ, USA), as indicated for organoid culture (Dominguez *Mol Cell Biol* 1992; **12**: 3776-3783). Eight well multichambers were coated with a thin layer of undiluted Matrigel and allowed to solidify. Organoid preparations in Matrigel (100 µl suspension) were placed into each well, then overlaid with 250 µL/well culture medium (Dulbecco's modified Eagle medium/F12) supplemented with penicillin/streptomycin, 10 mmol/L HEPES, Glutamax supplements 1× N2, 1× B27 [ThermoFisher, Dublin, Ireland], 1 mmol/L *N*-acetylcysteine [Sigma-Aldrich, Dorset, UK], 50 ng/ml murine EGF, Noggin 100 ng/ml (ThermoFisher, Dublin, Ireland) and 1µg/ml human R-Spondin (Dominguez *Mol Cell Biol* 1992; **12**: 3776-3783). Organoids were cultured for 4 days with peptide treatments as defined.

Example 21: Confocal immunofluorescence microscopy of 3D cultures and organoids

Embedded glands were fixed in 2% PFA for 20 min and processed for immunofluorescence, as previously described (Jagan *et al. Neoplasia* 2013; **15**: 1218-1230; Chambers & Bretscher *Biochemistry* 2005; **44**: 3926-3932). Briefly, Caco-2 cells and organoids in 3D culture were fixed in 2% PFA for 20 min at room temperature, washed in PBS and permeabilized for 10 min in 0.5% Triton X-100 in PBS. Primary antibodies were diluted in block buffer and incubated overnight at 4°C. Cells were incubated with secondary antibodies and/or FITC-phalloidin for 1 h. DNA was stained and chamber slides mounted using Vectashield mounting medium containing DAPI (Vector Scientific, Belfast, NI). Labelled cells were visualized using a Nikon 90i fluorescence microscope or a Leica SP5 laser scanning confocal microscope and images were processed using Elements (Nikon) or Leica LAS AF software. Sequential scan images were taken at the midsection of glands at room temperature using the Leica confocal on a HCX PL APO lambda blue 63 x 1.40 oil immersion objective at 1x or 2x zoom. Images were collected and scale bars added using LAS AF confocal software (Leica).

Example 22: Image processing Fluorescence microscopy images were processed using Leica Fw4000 Imaging software and cropped using Adobe Photoshop (CS2). Confocal images were processed, merged and mean area quantified using LAS AF Leica Imaging Software.

Example 23: Analyses of cell nuclear morphology and DNA content with ImageJ

In 3D organotypic and organoid cultures, spindle orientation, epithelial configuration and lumen formation were assessed as previously defined (Jagan *et al. Neoplasia* 2013; **15**: 1218-1230)(Tait *et al. Cell Transplant* 1994; 3: 33-40). Spindle architecture was defined as bipolar or multipolar as outlined above. Bipolar spindle orientation was assessed by the angle between the spindle plane and a line drawn from the spindle midpoint to the gland centre using ImageJ, as previously described (Jaffe *et al. J Cell Biol* 2008; **183**: 625-633). Nuclear size, DNA content and “roundness” were assessed in 3D glandular structures using ImageJ, as previously described (Wohlke *et al. Histopathology* 2011; **59**: 857-866). Roundness was computed using the border perimeter determined by the formula $\text{perimeter}^2 / (4 \times \pi \times \text{nuclear area})$ (Wohlke *et al. Histopathology* 2011; **59**: 857-866).

Example 24: Human tumour samples NHERF1 immunohistochemistry (IHC) was conducted in two CRC sample collections (A and B) while Aurora A immunofluorescence (IF) was assayed in sample (A) only. Sample (A) comprised 35 whole tumour sections and 5 matched normal mucosa specimens while sample (B) was a tissue microarray (TMA) comprising 309 tumour CRC specimen cores across two TMA sections, derived from 92 CRCs, as previously described (Tait *et al. Cell Transplant* 1994; 3: 33-40). Sample B cores from rectal tumours that had received neoadjuvant radiotherapy (n=28 cores) were excluded from analyses.

NHERF1 IHC Assays were conducted in the Northern Ireland Molecular Pathology Laboratory. Sections were cut at 4 µm thickness on a rotary microtome, dried overnight then stained in an automated immunostainer in accordance with manufacturer’s instructions (Leica Bond-Max, Milton Keynes, UK). Sections were incubated with mouse monoclonal anti-NHERF-1 antibody (Lifespan Biosciences, Seattle, WALSA-B1873;1/200), after pretreatment with Bond-Max epitope retrieval solution 2 for 20 mins. Primary antibody binding was detected using a polymer-based detection system (Bond Polymer Refine Detection, Newcastle Upon Tyne, UK, catalogue No DS9800) containing a peroxide block, post-primary polymer reagent 3,3'-diaminobenzidine tetrahydrochloride (DAB) chromogen solution, and haematoxylin counter-stain. Stained sections were mounted in DPX. Apical localization of NHERF1 was defined as enrichment along the apical membrane domain of glandular structures (Parmentier *et al. BMC Cell Biol* 2004; **5**: 4).

Sample A - Whole tumour/tissue sections were scored for apical NHERF1 intensity from 0-3 representing 0 = absent, 1 = weak, 2 = moderate and 3 = strong apical staining. Scoring was conducted at 10X magnification over 40 fields per tumour/tissue section by 2 independent assessors. The highest intensity staining present in each field and the percentage of the field demonstrating this staining intensity within 20% increments were multiplied to provide a final

score. A mean score per tumour section was derived from the average apical staining score across the 40 fields.

5 *Sample B* - Because of the small size of sample (A) we conducted additional NHERF1 IHC assays in both cytoplasm and at the apical domain in TMA samples, prepared as outlined above. NHERF1 localization at apical regions of discernible glandular structures was scored 0-2 where 0 = absent, 1= moderate and 2 = strong.

10 Cytoplasmic staining was scored 0-3 as outlined above for sample A. Scores were multiplied by the percentage tumour cells with the particular staining intensity. Scoring was conducted by 2 independent assessors (JM and ML). Any discordant scores were resolved by consensus decision. As previously reported, PKC ζ p-T560 IHC was unsuccessful in formalin fixed tissue sections (Jagan *et al. Neoplasia* 2013; **15**: 1218-1230).

15 *Aurora A immunofluorescence* - Aurora A IF was conducted in Sample A tumours only. Freshly cut 6 μ m sections of tumour tissue (n = 34) or normal mucosa (n = 5) were used in assays of mitotic spindle architecture by Aurora A immunofluorescence, as previously described (Saito *et al. J Pathol Inform* 2016; **7**: 36). Slides were deparaffinised in xylene washes x3, washed in 100% and 95% ethanol and dH₂O. Nonspecific endogenous
20 peroxidase activity was blocked in 3% hydrogen peroxide in 99% alcohol/methylated spirit, then washed in water and Tris/HCl-buffered saline (TBS)/Tween 20 and placed on a staining rack. Sections were then subjected to antigen retrieval using heat pre-treatment and placed in metal racks in a pressure cooker at 15 psi, filled with 2000mls of Tris-EDTA pH 9.0 buffer (200mls x10 stock solution and 1800ml distilled water), for 20 min. Slides were incubated in
25 mouse anti-Aurora A/IAK1 antibody (clone 4, 1 h) (BD Biosciences, Oxford, UK) at 1:200 concentration in TBS, overnight at 4°C. Slides were washed twice in TBS/Tween 20 (TBST) before addition of goat anti-Mouse Alexa Fluor® 568 secondary antibody (IgG (H+L), (Invitrogen Molecular Probes, Paisley, UK 1:1000). Slides were incubated at 37°C for 2 hours in the dark (to prevent photobleaching of the fluorophores), before two five-minute
30 washes in TBST. Slides were then mounted with DAPI nuclear counterstain and VECTASHIELD® HardSet Antifade mounting medium (Vector Laboratories, Peterborough, UK) was applied. Slides were stored in the dark at 4°C until viewed.

Stained sections were analysed with a Nikon Eclipse Ti inverted two-channel fluorescence microscope (Nikon UK, Ltd, Surrey, England). Sections were screened using a 63/1.3 oil
35 objective lens. In mitotic cells, normal bipolar mitotic spindle architecture was characterized by two distinct polar Aurora A positive signals while multipolar mitoses were defined as >2 Aurora A positive signals (Saito *et al. J Pathol Inform* 2016; **7**: 36). Aurora A localizes to the

centrosome during late S phase at the time of centrosome replication (Godinho *et al. Nature* 2014; **510**: 167-171). Cells with dispersed Aurora A were thought to represent early S phase before relocalization to the centrosome and those with only one Aurora A signal were thought to contain unseparated centrosomes during interphase. Both of these categories
5 were excluded from the analysis. A total of 84.7 ± 4.7 high power field images containing 180 ± 72 mitotic cells per section were captured as TIFF files using Nikon Elements Imaging Software Viewer (Nikon, version 4.20) for review and analysis.

By the above methods, each CRC specimen had a mean quantitative score for apical NHERF1 intensity and a mean value for multipolar spindle frequency, as a percentage of the
10 total mitotic figures. We conducted correlation studies between these two quantitative endpoints.

Clinicopathological demographics Clinicopathological information was collected from Belfast Health and Social Services Trust Labcentre. Collated data included patient age, sex, site,
15 size, histological type and grade of primary tumour, local depth of invasion.

Example 25: Data analysis Descriptive statistics were expressed as the mean $\pm$ sem. Apical NHERF1 IHC data were skewed and were log-transformed to provide a normal distribution. Statistical analyses were by one- or two-way ANOVA, Student's t test or
20 Pearson's test of correlation using SPSS for Windows release 24.0 (IBM Corp, NY, USA) or Graphpad Prism software (v4.02; Graphpad CA 92037 USA). Scatterplots, bar charts and boxplots were used for display of quantitative numerical or categorical data.

Claims

1. A modified Caco-2 cell, the modified Caco-2 cell comprising a heterologous polynucleotide capable of hybridising to a polynucleotide encoding PKCz or PTEN or NHERF1 such that expression of said heterologous polynucleotide results in a reduction in, or elimination of, expression or level or activity of native PKCz or PTEN or NHERF1 polypeptide in the modified Caco-2 cell compared to an equivalent, unmodified cell.
2. A modified Caco-2 cell as claimed in claim 1, the cell comprising a polynucleotide expression construct, wherein the construct comprises:
 - a. a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1; and/or
 - b. a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2; and/or
 - c. a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3;
 and which upon expression results in a reduction in the expression level or activity of PKCz and/or PTEN and/or NHERF1 polypeptide respectively in the modified Caco-2 cell, relative to an equivalent, unmodified Caco-2 cell.
3. A modified Caco-2 cell as claimed in claim 2, wherein the expression of the polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1; and/or the polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2, and/or the polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3; is under the control of an inducer molecule.
4. A modified Caco-2 cell as claimed in any preceding claim, wherein the cell further comprises a polynucleotide expression construct, and wherein the polynucleotide expression construct comprises a polynucleotide sequence of SEQ ID NO: 4, or a sequence of at least 90% identity thereto, and/or a polynucleotide sequence of SEQ ID NO: 5 or a sequence of at least 90% identity thereto; and which upon

expression results in an increase in the expression level or activity of Polo Like Kinase 4 (PLK4) and/or Aurora A polypeptide respectively, in the modified Caco-2 cell, relative to an equivalent, unmodified Caco-2 cell.

5. A modified Caco-2 cell as claimed in claim 4, wherein the expression of the polynucleotide sequence of SEQ ID NO: 4, or a sequence of at least 90% identity thereto, and/or the polynucleotide sequence of SEQ ID NO: 5 or a sequence of at least 90% identity thereto; is under the control of an inducer molecule.
6. A modified caco-2 cell, having reduced levels or activity of PKCz and/or PTEN and/or NHERF1 polypeptide relative to an equivalent, unmodified Caco-2 cell.
7. A modified caco-2 cell as claimed in any preceding claim, wherein the modified Caco-2 cell has reduced levels or activity of PKCz polypeptide relative to an equivalent, unmodified Caco-2 cell.
8. A modified Caco-2 cell as claimed in any preceding claim, wherein the modified Caco-2 cell has increased levels or activity of PLK4 and/or Aurora A polypeptide relative to an equivalent, unmodified Caco-2 cell.
9. A modified Caco-2 cell as claimed in any preceding claim, wherein the modified Caco-2 cell has reduced levels or activity of PKCz polypeptide relative to an equivalent, unmodified Caco-2 cell; and increased levels or activity of PLK4 polypeptide relative to an equivalent, unmodified Caco-2 cell.
10. A cancer model system comprising more than one modified Caco-2 cell as claimed in any preceding claim.
11. A cancer model system as claimed in claim 10, wherein the cancer model system is a 3D cell culture.
12. A cancer model system as claimed in claim 11, wherein the cancer model system is an organotypic 3D cell culture.
13. A cancer model system as claimed in any of claims 10 to 12, wherein the resulting cancer model system is a low-grade cancer model system.
14. A cancer model system as claimed in any of claims 10 to 12, wherein the resulting cancer model system is a high-grade cancer model system.
15. A method of generating a cancer cell, comprising the steps of:
 - a. providing a Caco-2 cell; and

- b. reducing the expression or activity of PKC ζ and/or PTEN and/or NHERF1 polypeptide in said Caco-2 cell, relative to that of an equivalent, unmodified Caco-2 cell.
- 5 16. A method as claimed in claim 15, wherein the Caco-2 cell is transformed with a polynucleotide capable of hybridising to a polynucleotide encoding PKC ζ or PTEN or NHERF1 such that there is reduction in, or elimination of, expression or level of native PKC ζ or PTEN or NHERF1 polypeptide in such cells compared to an untransformed control Caco-2 cell.
- 10 17. A method as claimed in claim 16, wherein the Caco-2 cell is transformed with;
- a. a polynucleotide capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1; and/or
- b. a polynucleotide capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2; and/or
- 15 c. a polynucleotide capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3.
18. A method as claimed in any of claims 15 to 17, further comprising increasing the expression or activity of PLK4 and/or Aurora A polypeptide in the Caco-2 cell, relative to that of an equivalent, unmodified Caco-2 cell.
- 20 19. A method as claimed in claim 18, wherein the Caco-2 cell is transformed with;
- a. a polynucleotide of SEQ ID NO: 4 or a sequence of at least 90% identity therewith; and/or
- b. a polynucleotide of SEQ ID NO: 5 or a sequence of at least 90% identity therewith.
- 25 20. A method as claimed in claims 15 to 17, wherein the resulting cell is a low-grade cancer cell.
21. A method as claimed in claims 18 to 19, wherein the resulting cell is a high-grade cancer cell.
22. A method of making a cancer model system, comprising the steps of:
- 30 a. providing more than one modified Caco-2 cell as claimed in any of claims 1 to 9; and
- b. culturing said modified Caco-2 cells in a culture medium.

23. A method of making a cancer model system, as claimed in claim 22, wherein the cancer model system is a 3D cell culture, and wherein the method further comprises inoculating said modified Caco-2 cells onto a three-dimensional framework.
- 5 24. A method making a cancer model system, as claimed in claim 23, wherein the method further comprises culturing said modified Caco-2 cells so that the modified Caco-2 cells attach to and substantially envelope the framework.
25. A method of screening for a cancer-therapeutic compound, comprising;
- 10 a. providing a modified Caco-2 cell as claimed in any of claims 1 to 9, or a cancer model system as claimed in any of claims 10 to 14; and
- b. contacting the modified Caco-2 cell, cancer model system with a candidate cancer-therapeutic compound; and
- c. assaying the effect of the candidate cancer-therapeutic compound on a biological activity of the modified Caco-2 cell, or cancer model system.
- 15 26. A method of determining a candidate target for cancer-therapy, wherein the method comprises the steps of:
- a. providing a modified Caco-2 cell as claimed in any of claims 1 to 9; and
- b. determining the level or activity of one or more polynucleotides and/or polypeptides in the modified Caco-2 cell;
- 20 c. determining the level or activity of the corresponding one or more polynucleotides and/or polypeptides in an equivalent unmodified Caco-2 cell; and
- d. comparing the level or activity of the one or more polynucleotides and/or polypeptides in the modified Caco-2 cell with that in the equivalent, unmodified Caco-2 cell;
- 25 wherein observed changes in the level or activity of the one or more polynucleotides and/or polypeptides in the modified compared to the unmodified cell is indicative of a candidate target for cancer-therapy.
27. A method of determining a candidate diagnostic or prognostic marker of cancer, wherein the method comprises the steps of:
- 30 a. providing a modified Caco-2 cell as claimed in any of claims 1 to 9; and

- b. determining the level or activity of one or more polynucleotides and/or polypeptides in the modified Caco-2 cell;
- c. determining the level or activity of the corresponding one or more polynucleotides and/or polypeptides in an equivalent unmodified Caco-2 cell; and
- d. comparing the level or activity of the one or more polynucleotides and/or polypeptides in the modified Caco-2 cell with that in the equivalent, unmodified Caco-2 cell;

wherein observed changes in the level or activity of the one or more polynucleotides and/or polypeptides in the modified compared to the unmodified cell is indicative of a diagnostic or prognostic marker of cancer.

28. A cancer model system as claimed in any of claims 10 to 14, or a method as claimed in any of claims 15 to 27, wherein the cancer is an adenocarcinoma, preferably colorectal adenocarcinoma.

29. An expression vector for expression in a Caco-2 cell comprising;

- a. a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1 or a fragment thereof and/or a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 2 or a fragment thereof and/or a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 3 or a fragment thereof; and

- b. a polynucleotide sequence having a sequence of SEQ ID NO: 4 or a sequence of at least 90% thereto; and/or a sequence of SEQ ID NO: 5 or a sequence of at least 90% thereto;

- c. wherein said expression vector is capable of expression in a Caco-2 cell, and wherein upon expression in a host Caco-2 cell results in a reduction in, or elimination of, expression or level or activity of PKC ζ and/or PTEN and/or NHERF1 and an increase in the expression or level or activity of PLK4 and/or Aurora A in said Caco-2 cell relative to that of an unmodified control Caco-2 cell.

30. An expression vector as claimed in claim 29, wherein the expression of any of the polynucleotide sequences contained therein is under the control of an inducer molecule.
31. An expression vector as claimed in claim 29 or claim 30, comprising;
- 5 a. a polynucleotide sequence which is capable of hybridising to a polynucleotide having an mRNA sequence corresponding to a cDNA sequence of SEQ ID NO: 1 or a fragment thereof which upon expression in a host Caco-2 cell results in a reduction in, or elimination of, expression or level or activity of native PKCz polypeptide in the host Caco-2 cell, relative to an equivalent, unmodified Caco-2 cell; and
- 10 b. a polynucleotide sequence having a sequence of SEQ ID NO: 4 or a sequence of at least 90% thereto and which upon expression in a host Caco-2 cell results in an increase in the expression level or activity of PLK4 polypeptide in the host Caco-2 cell, relative to an equivalent, unmodified
- 15 Caco-2 cell.

FIGURE 1

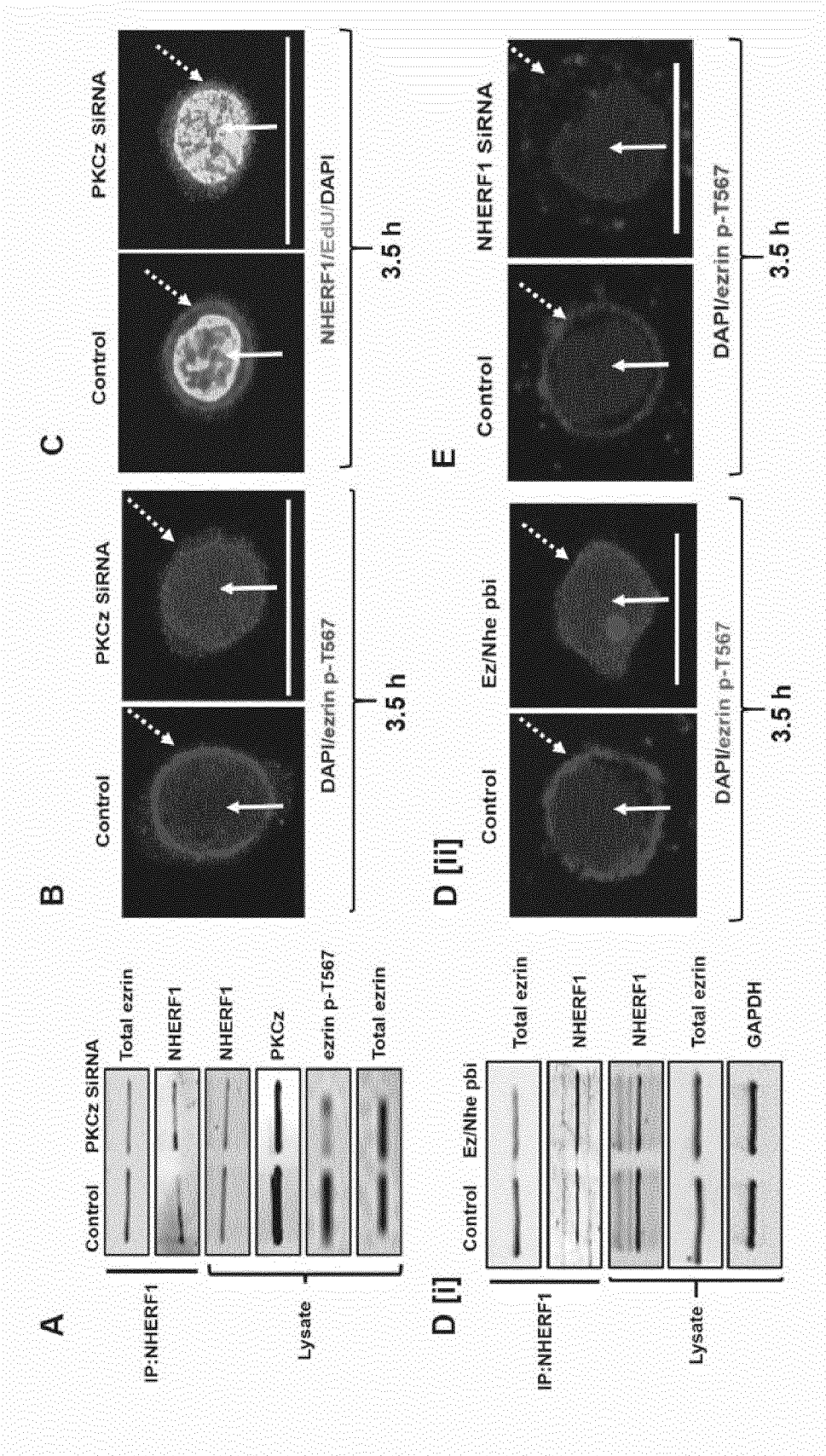


FIGURE 1

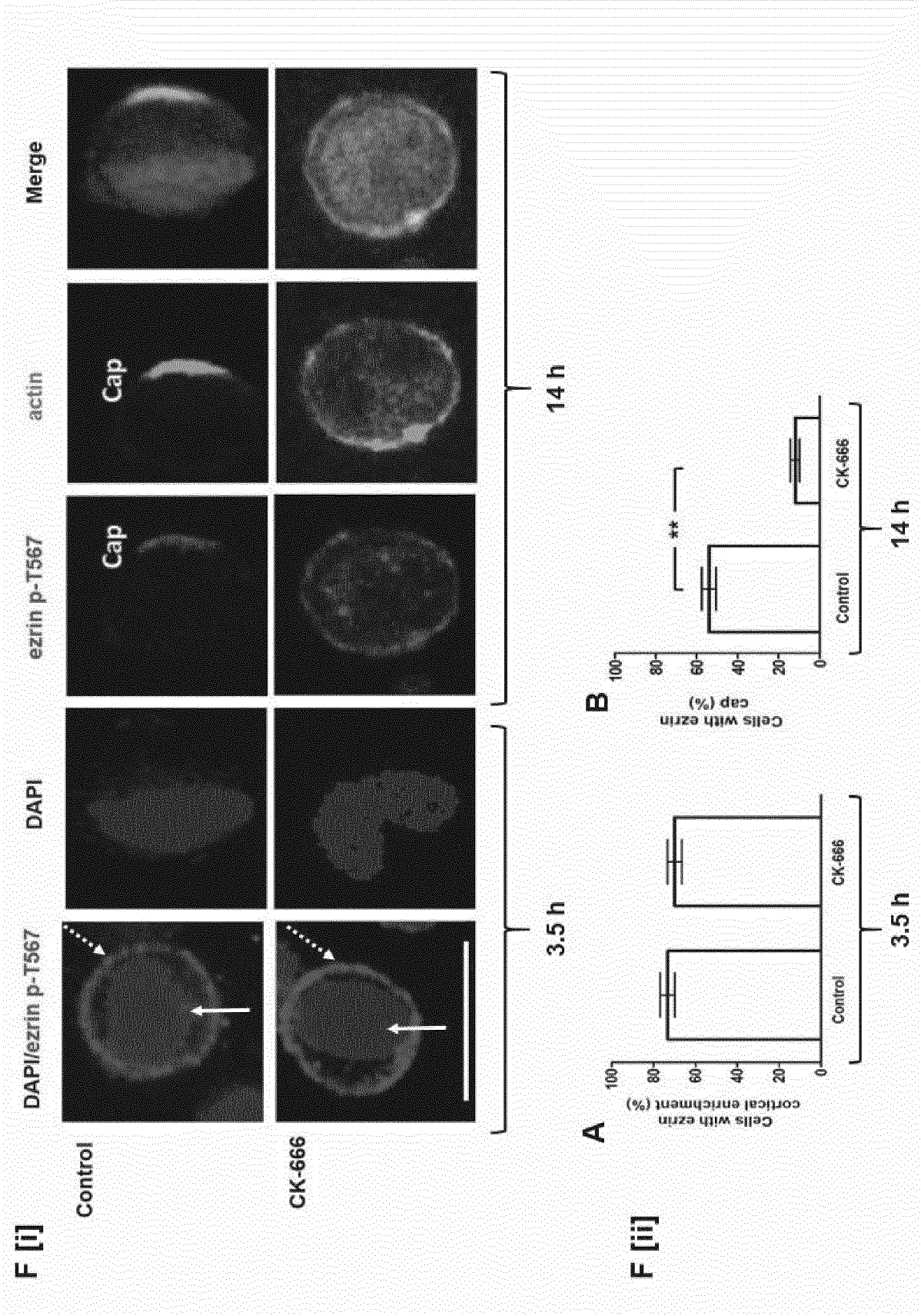


FIGURE 2

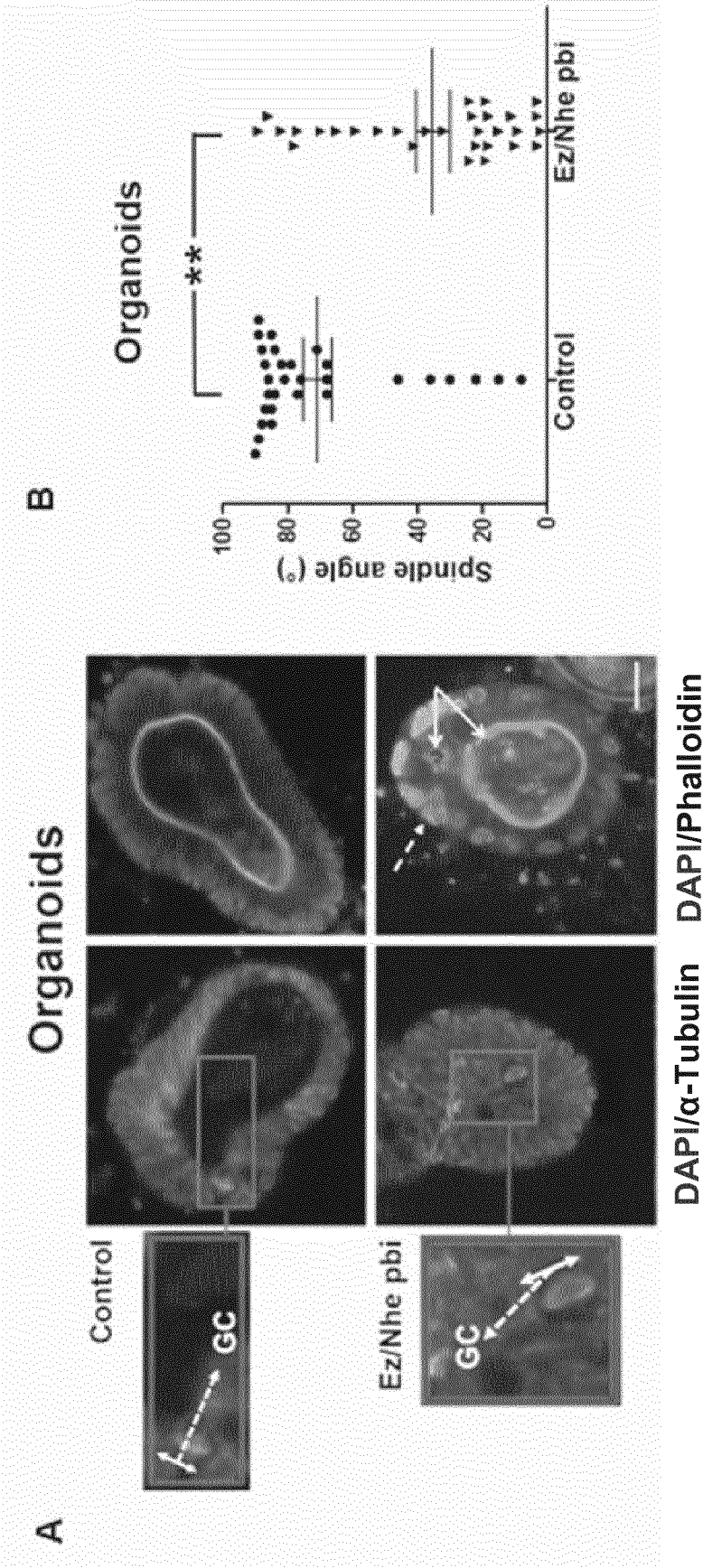


FIGURE 2

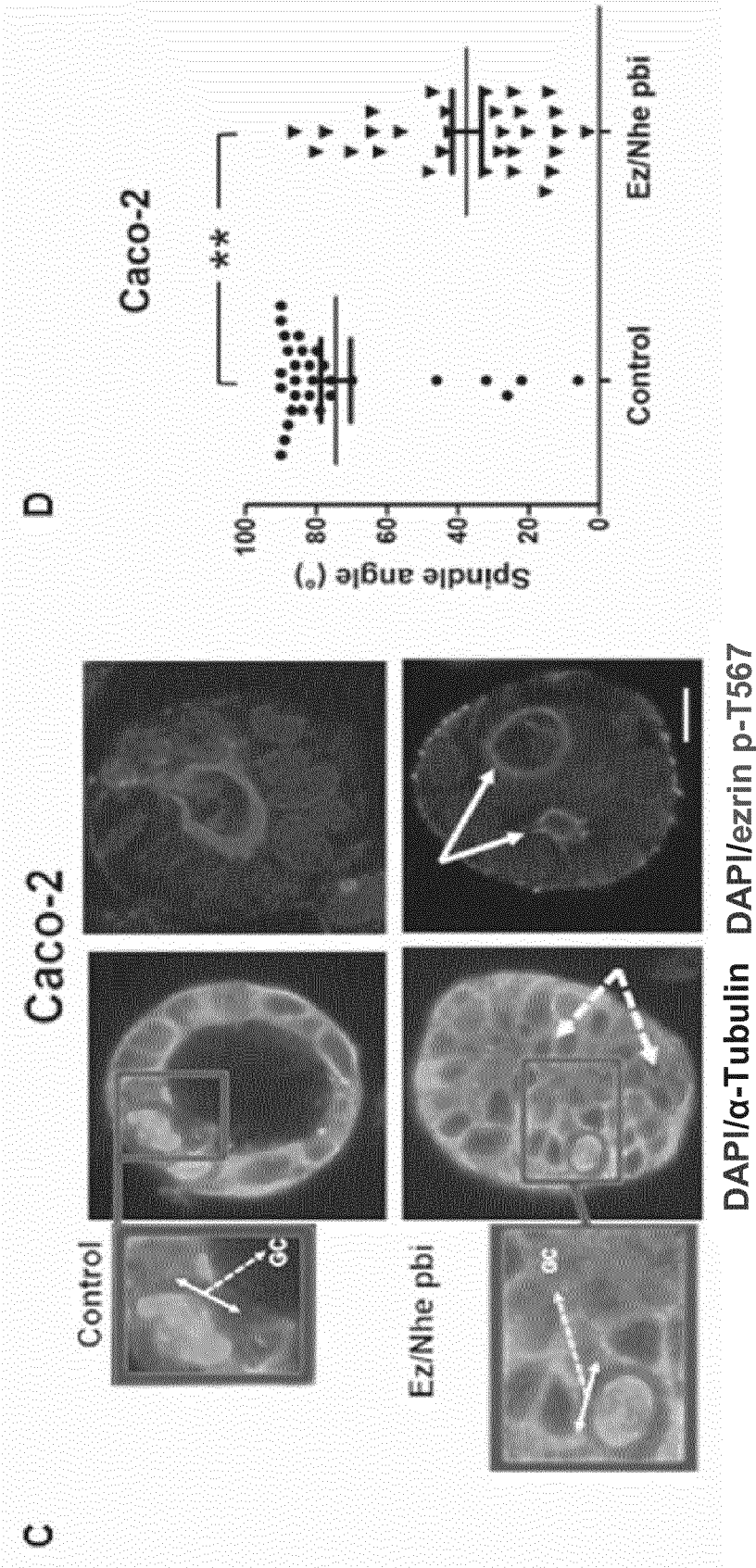


FIGURE 3

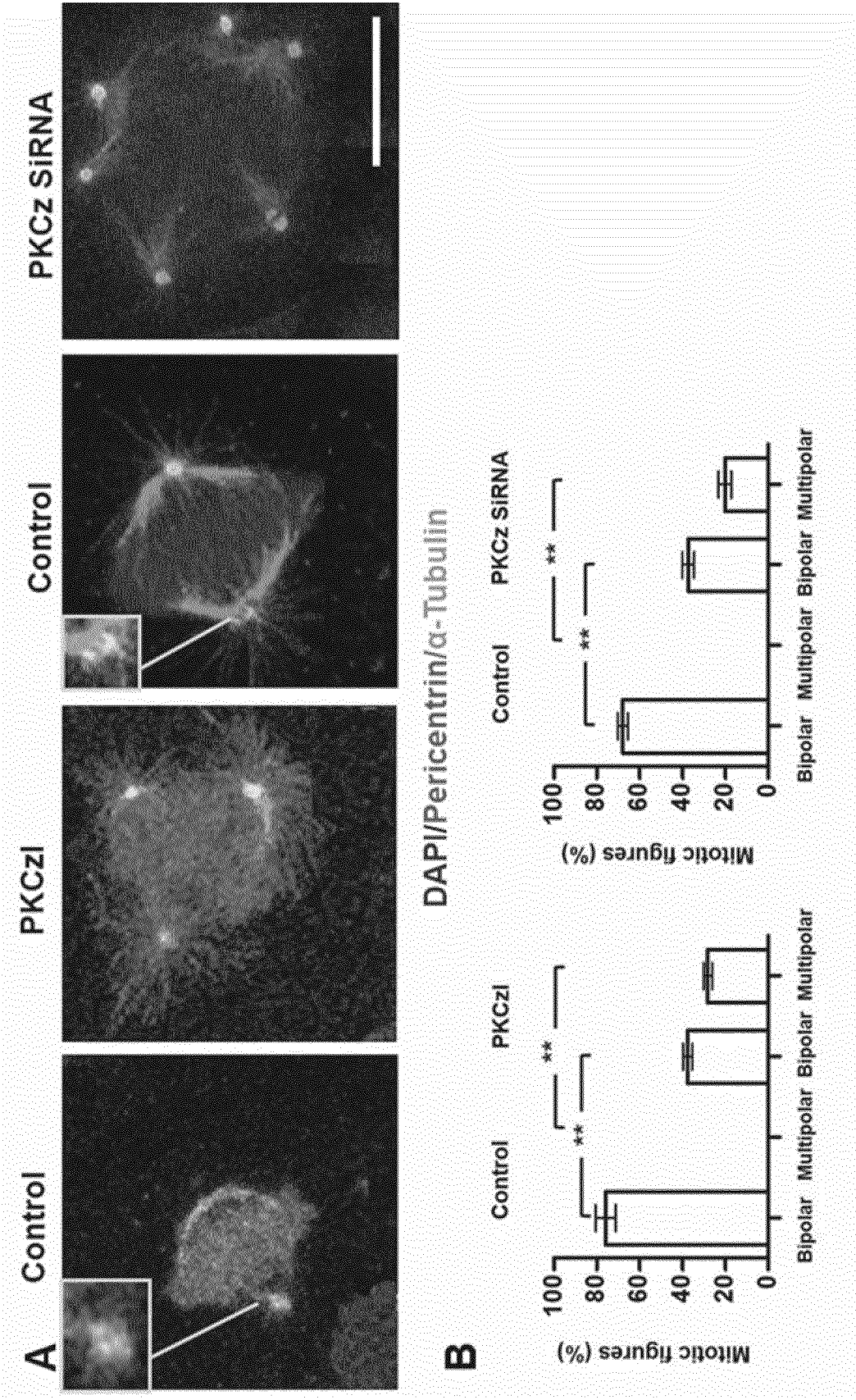


FIGURE 3

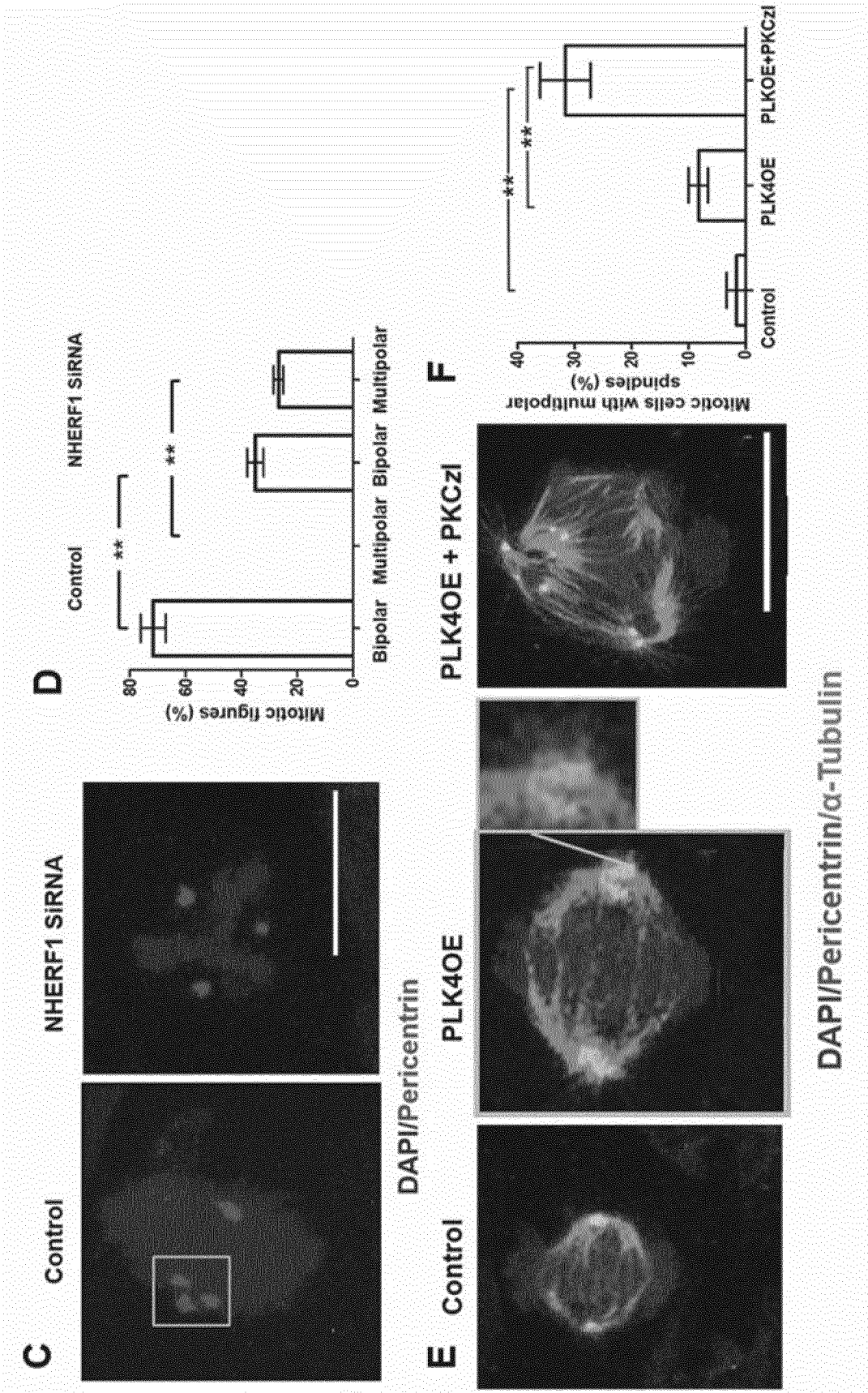


FIGURE 4

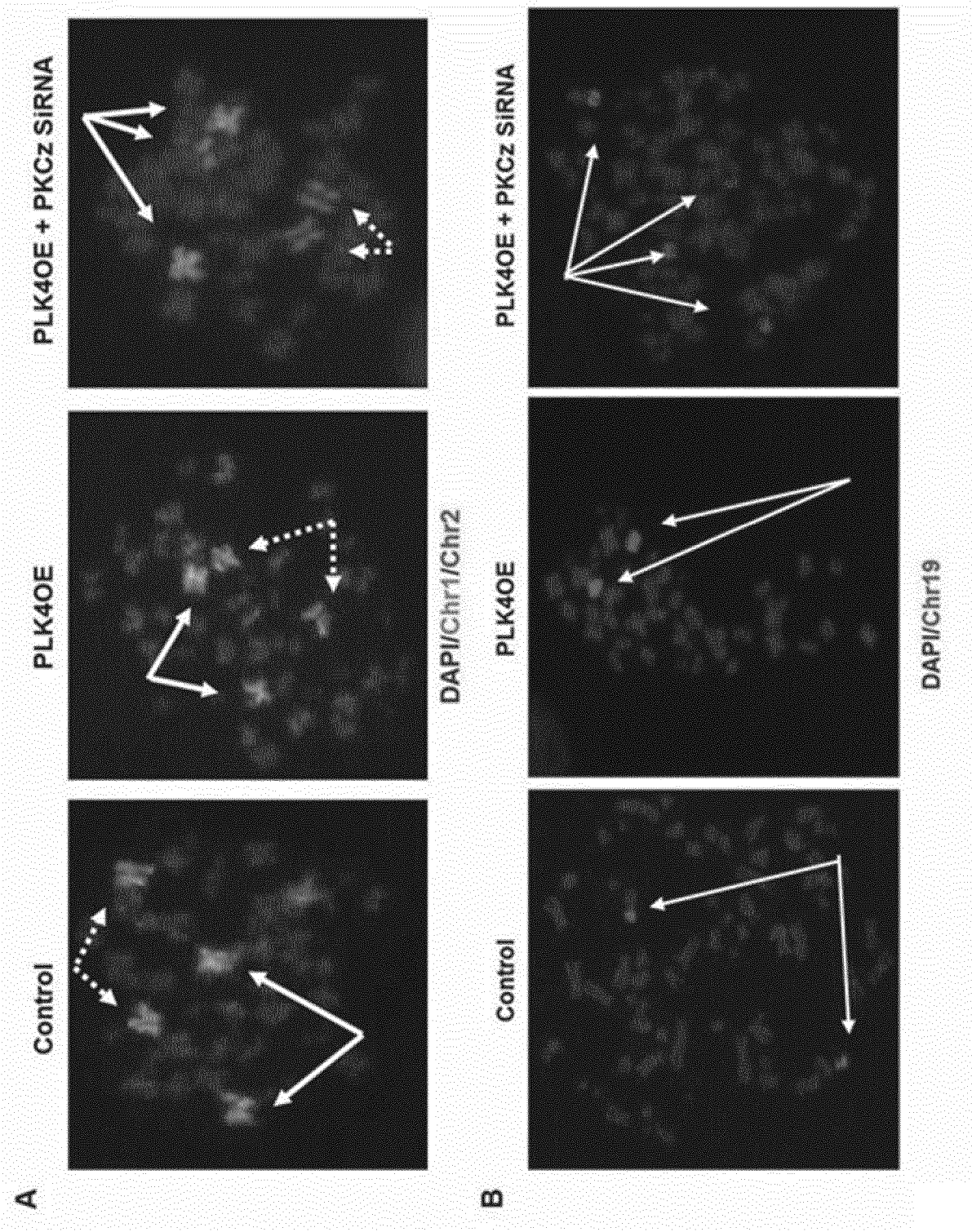


FIGURE 4

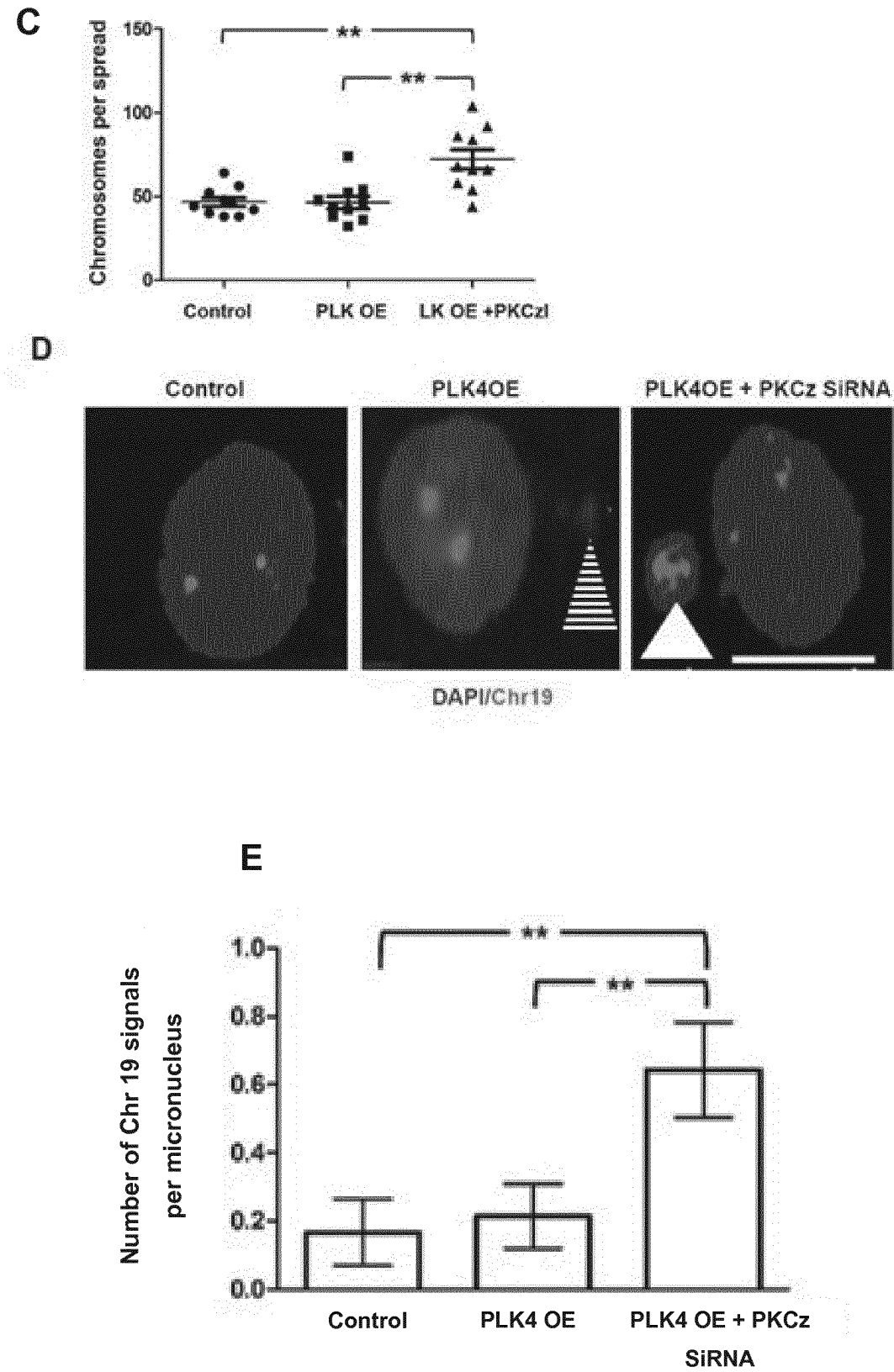


FIGURE 5

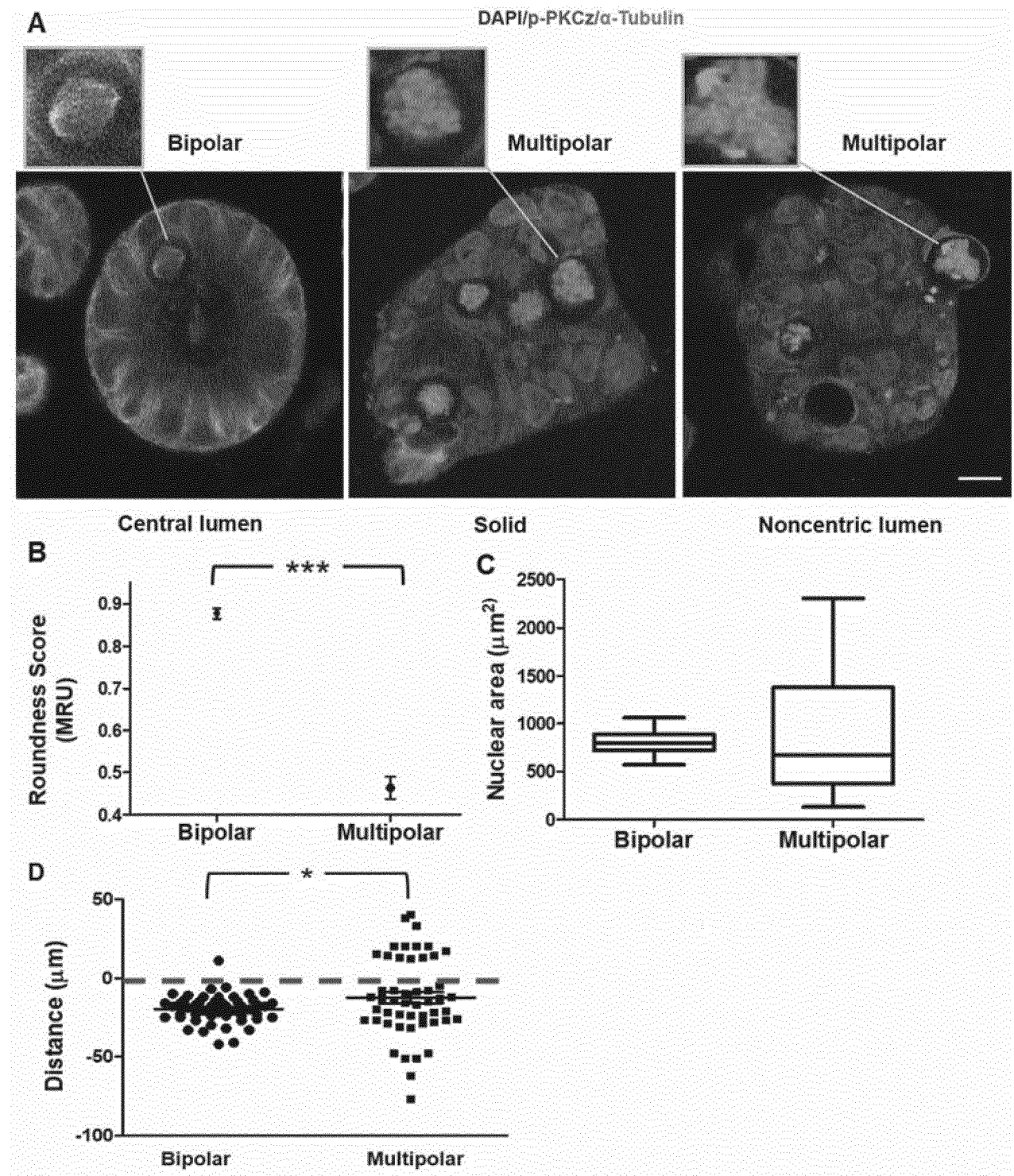


FIGURE 6

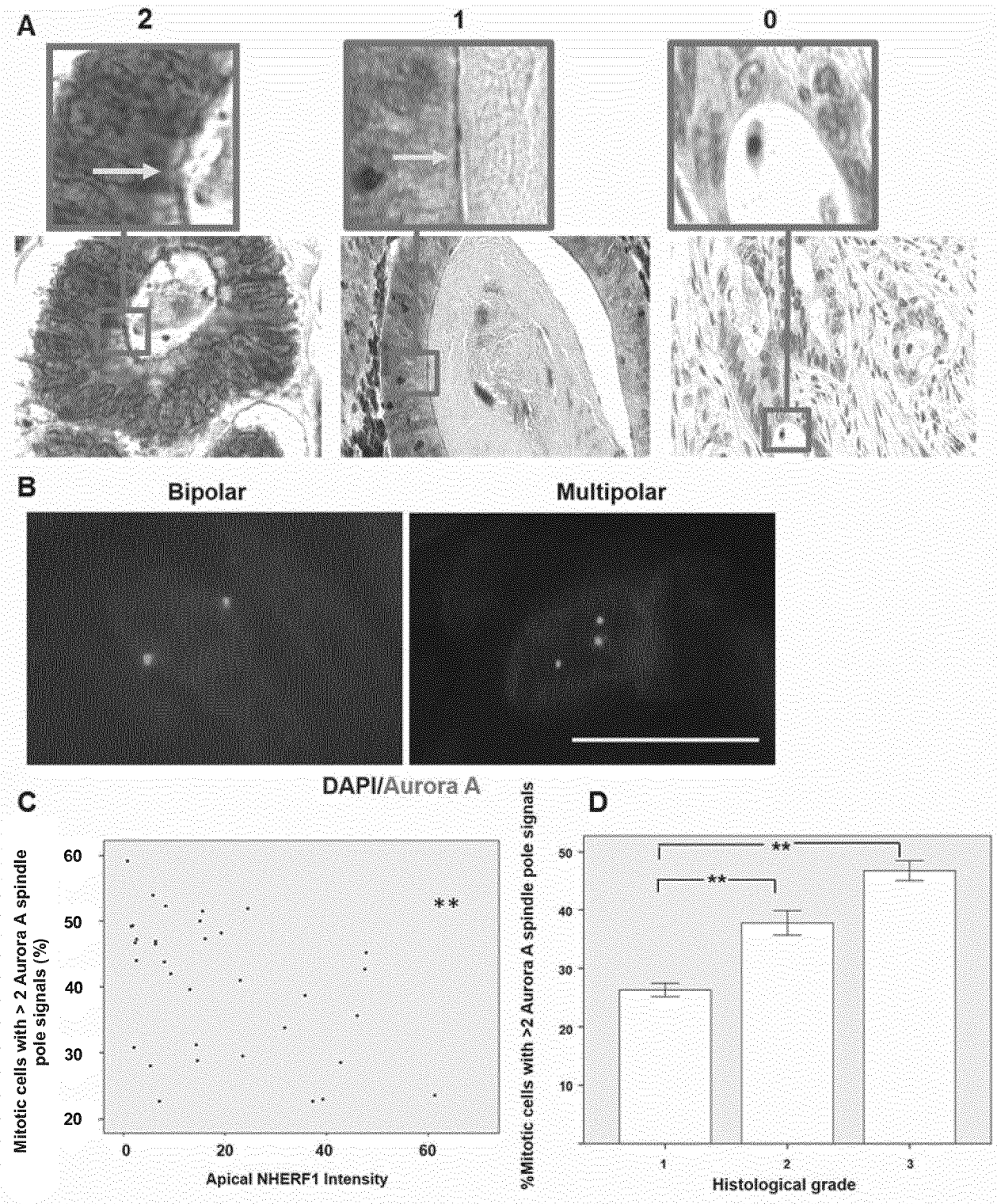


FIGURE 7

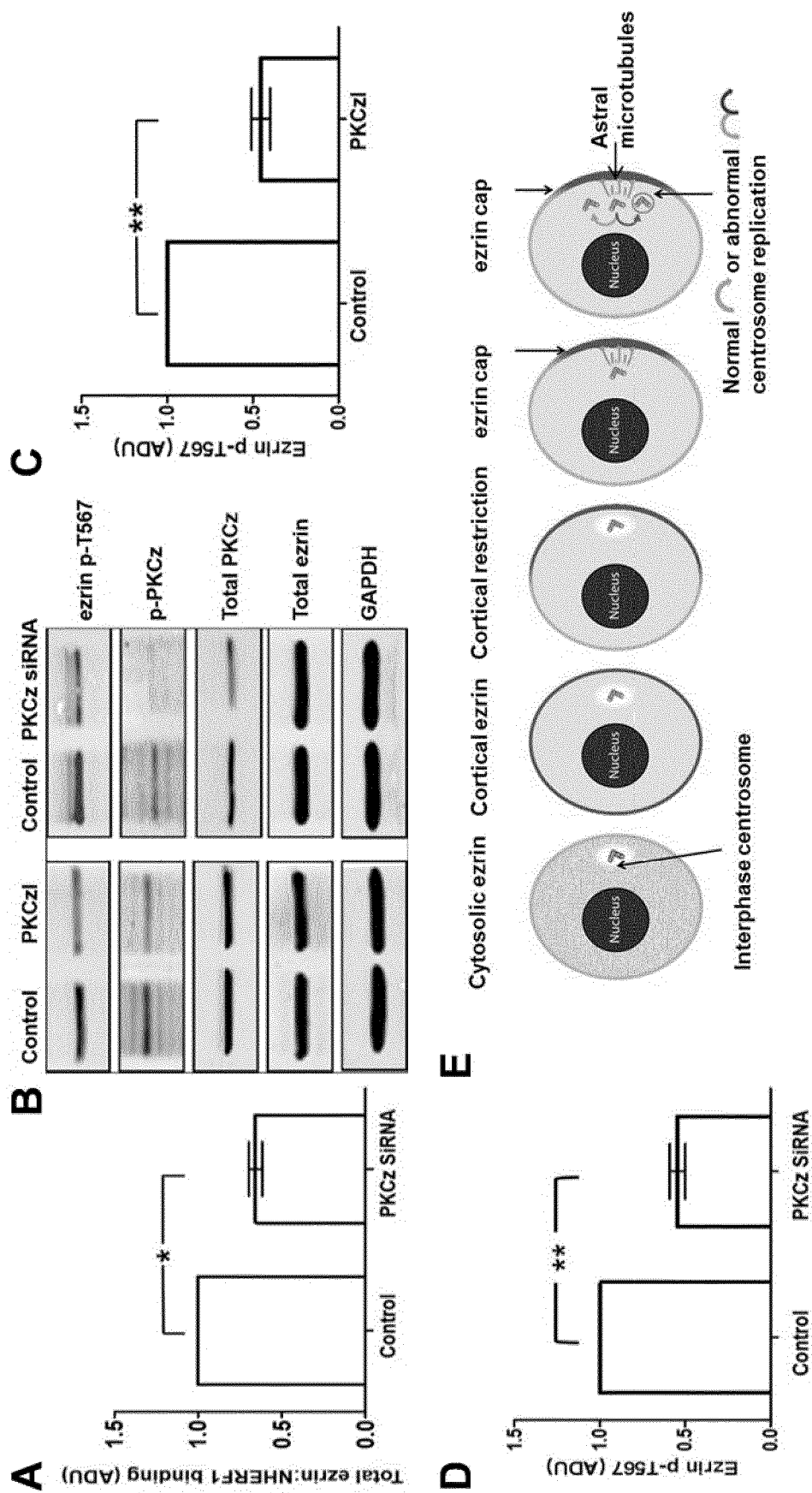


FIGURE 7

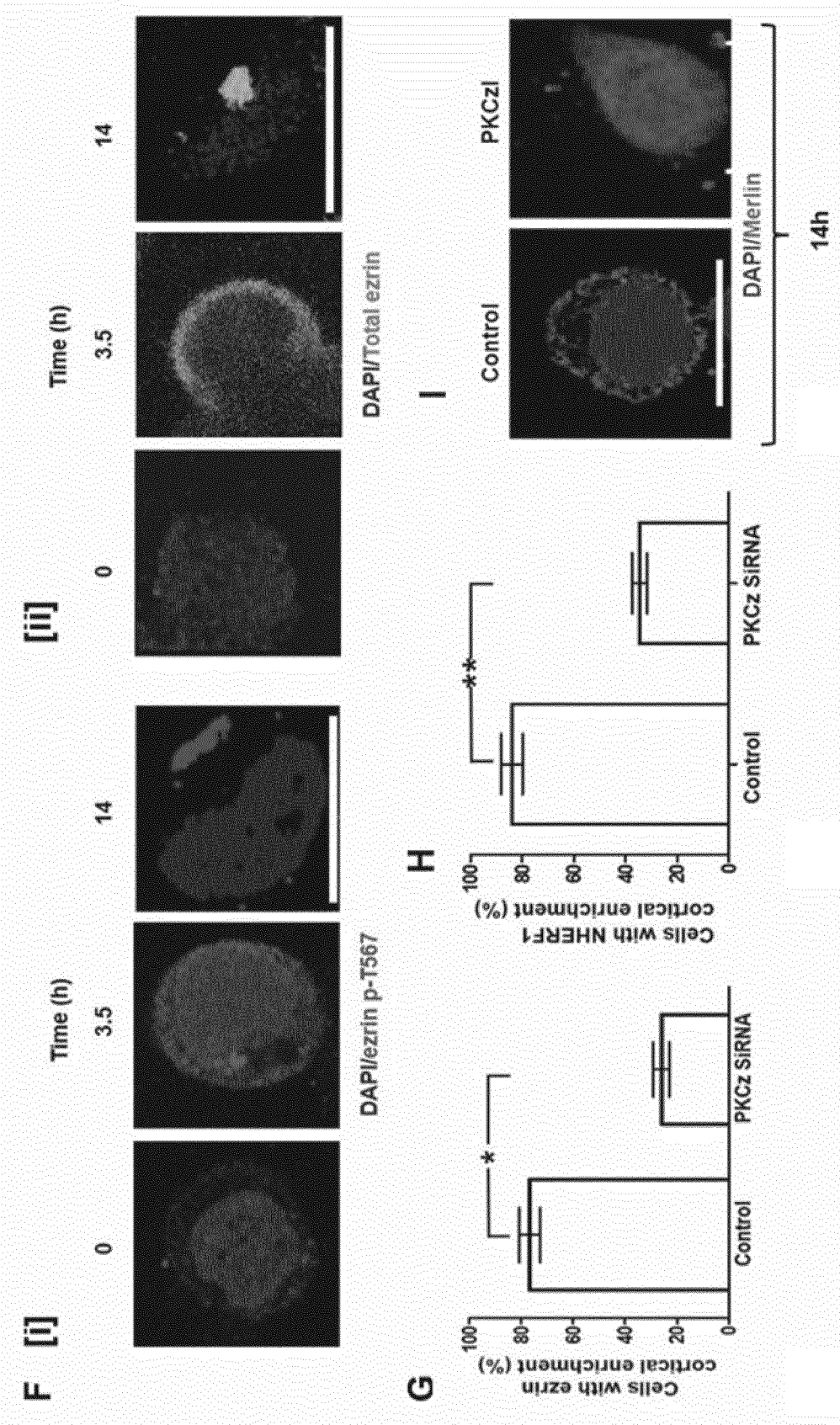


FIGURE 7

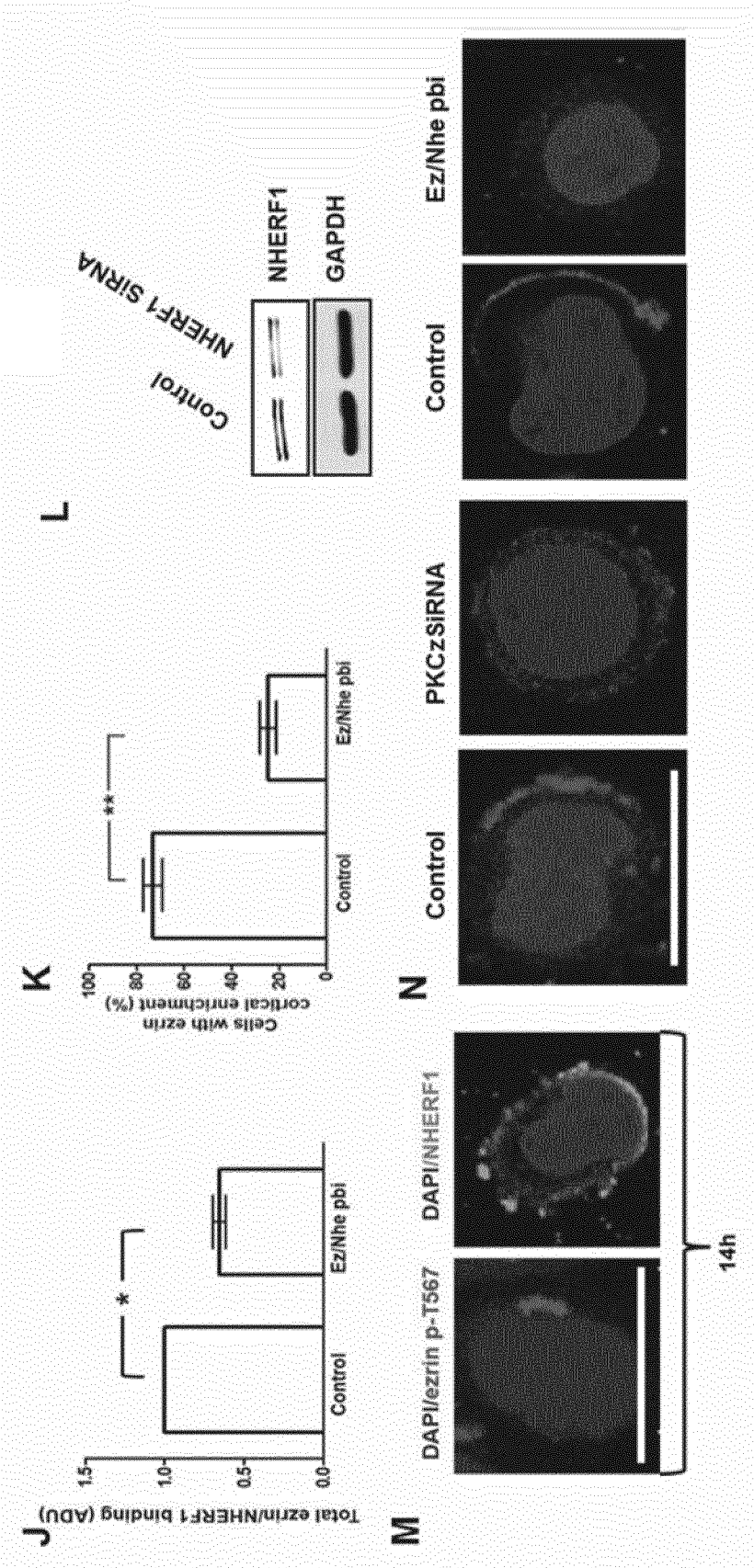


FIGURE 8

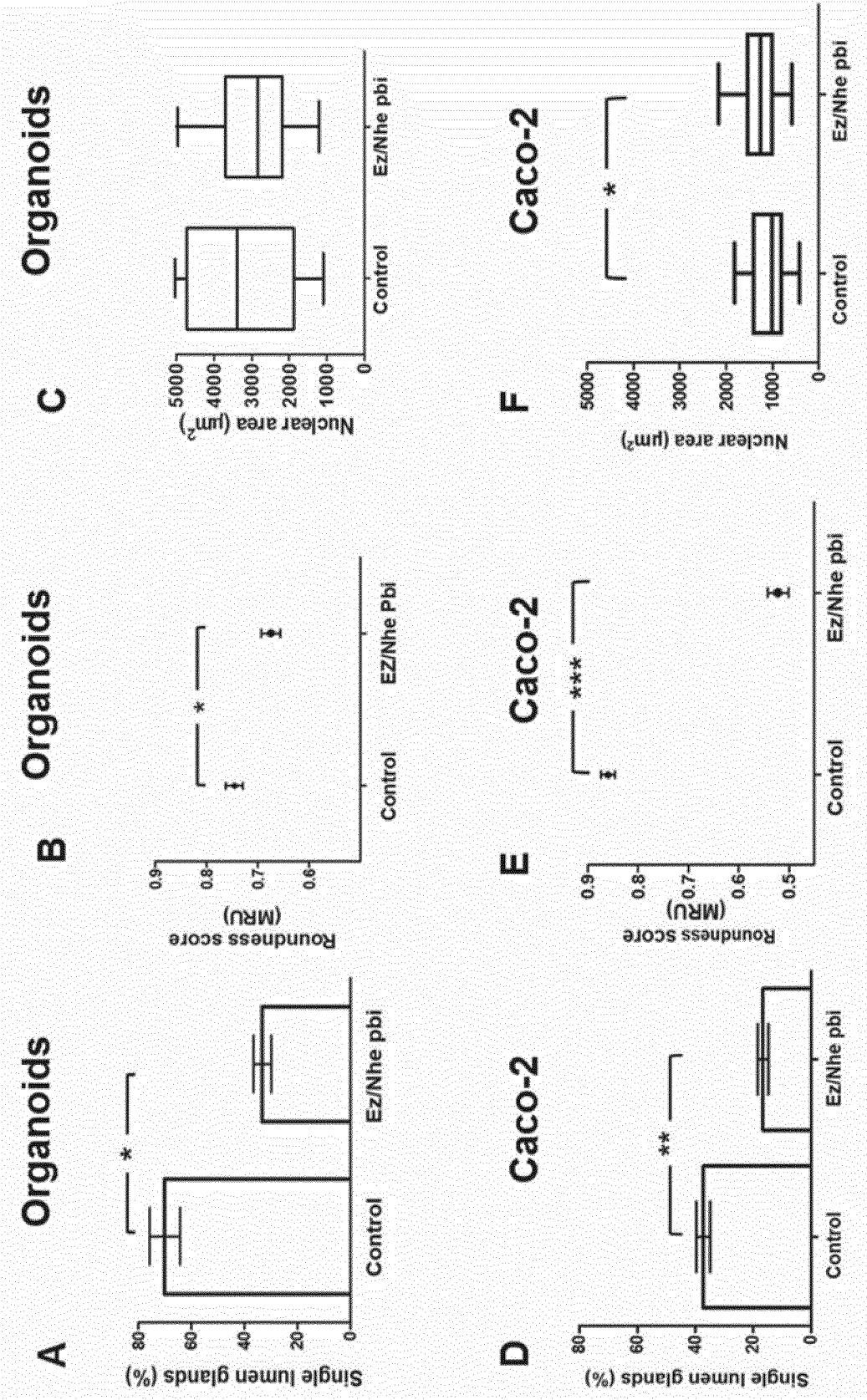


FIGURE 9

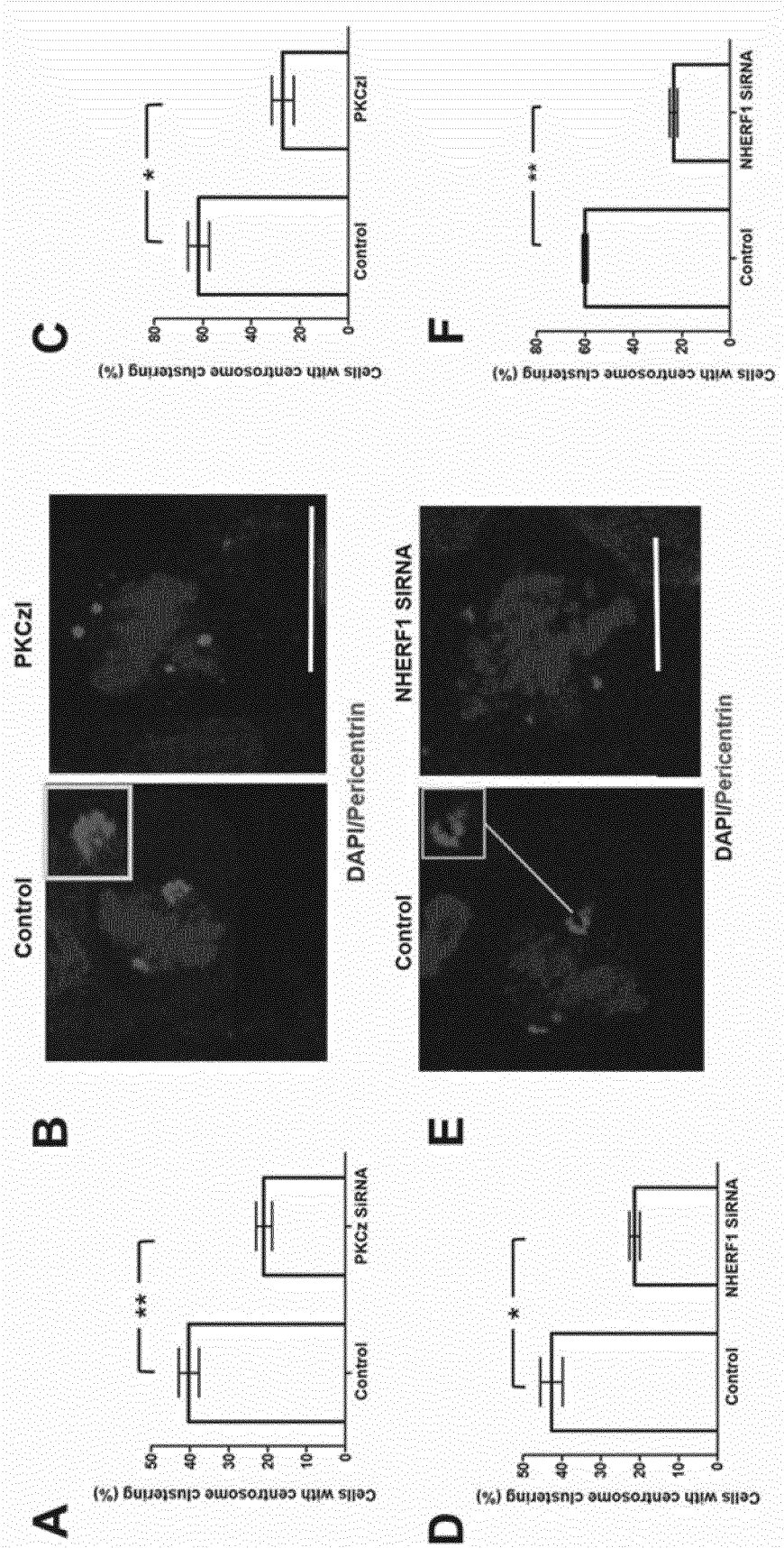


FIGURE 9

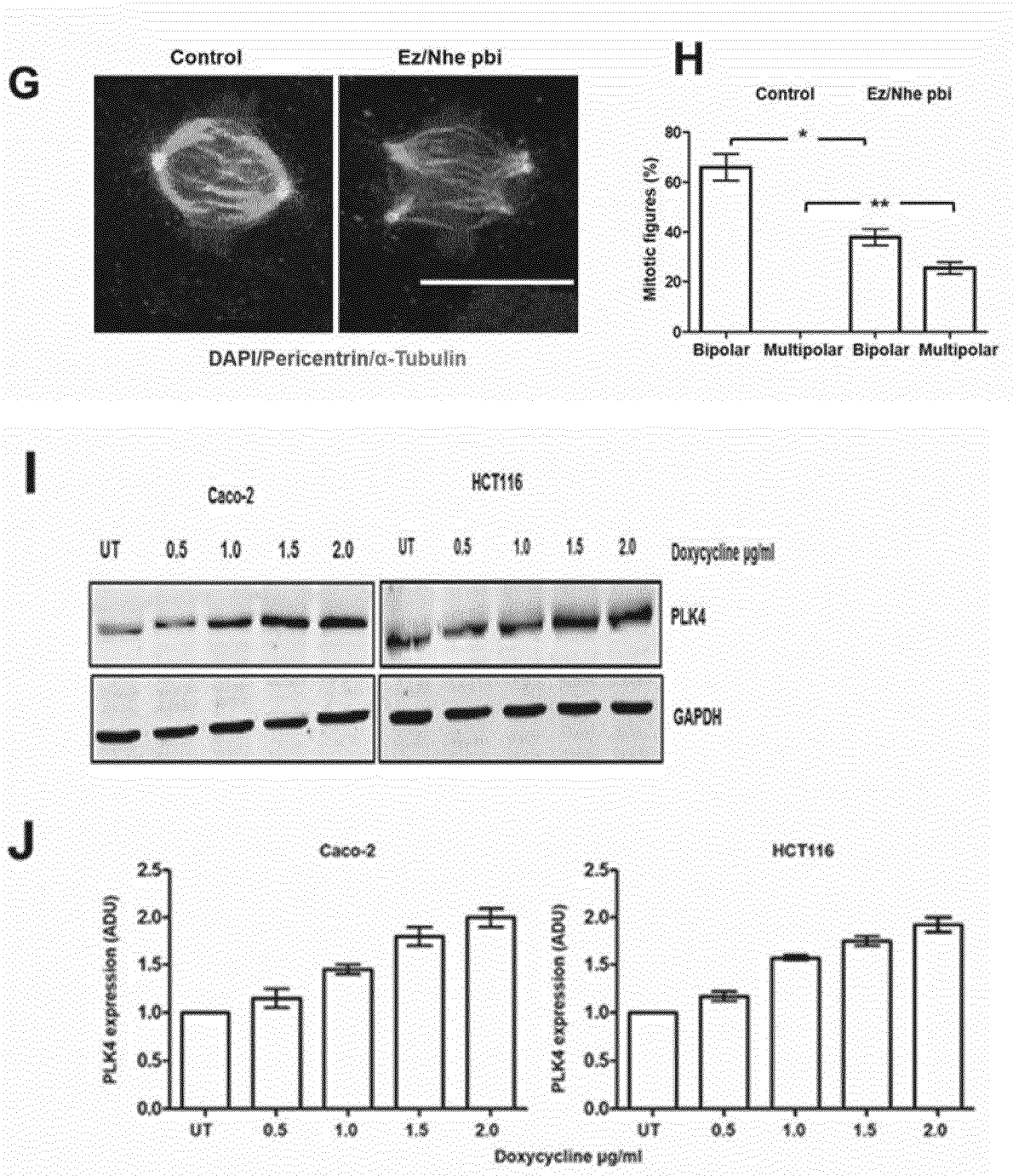


FIGURE 9

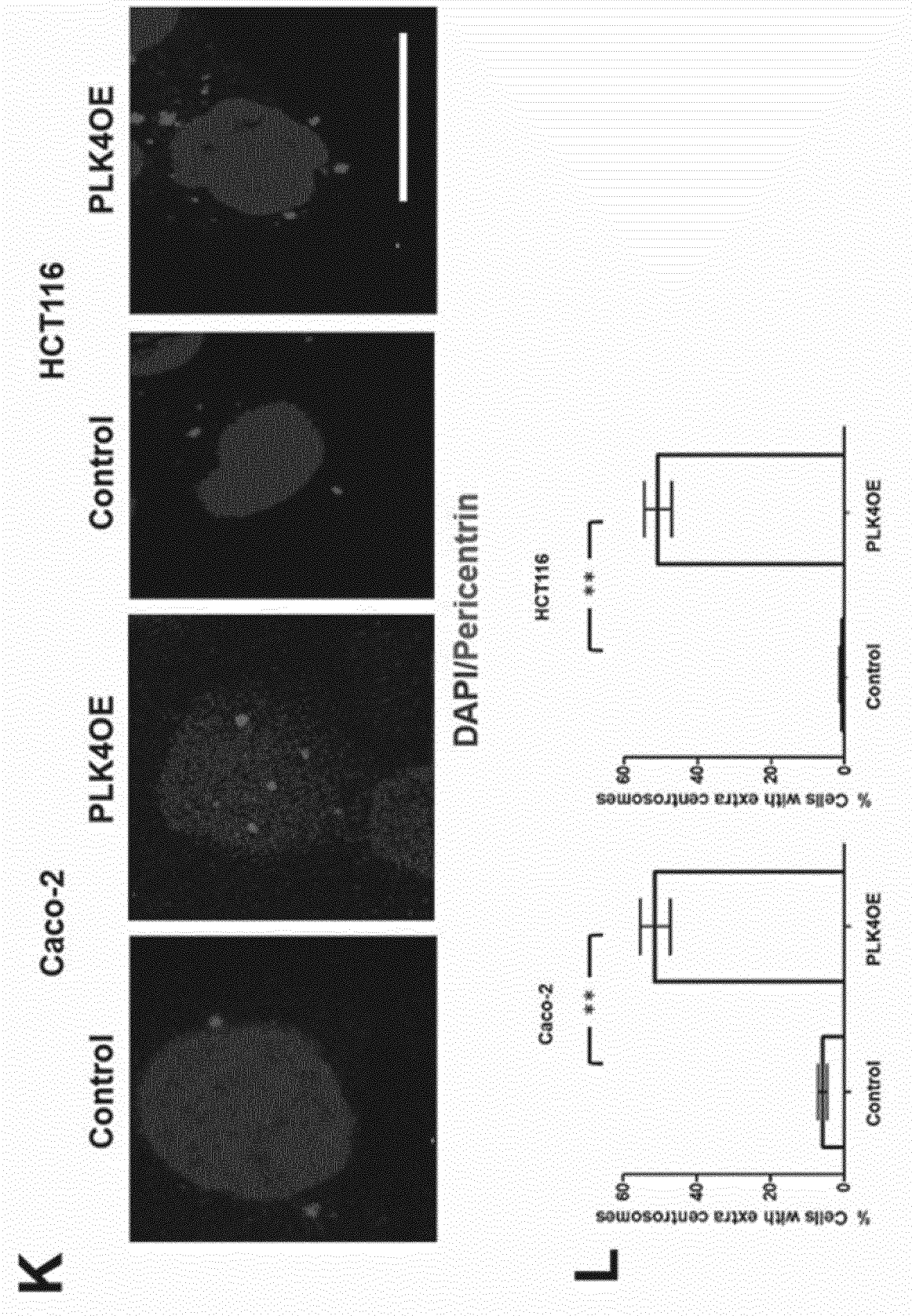


FIGURE 10

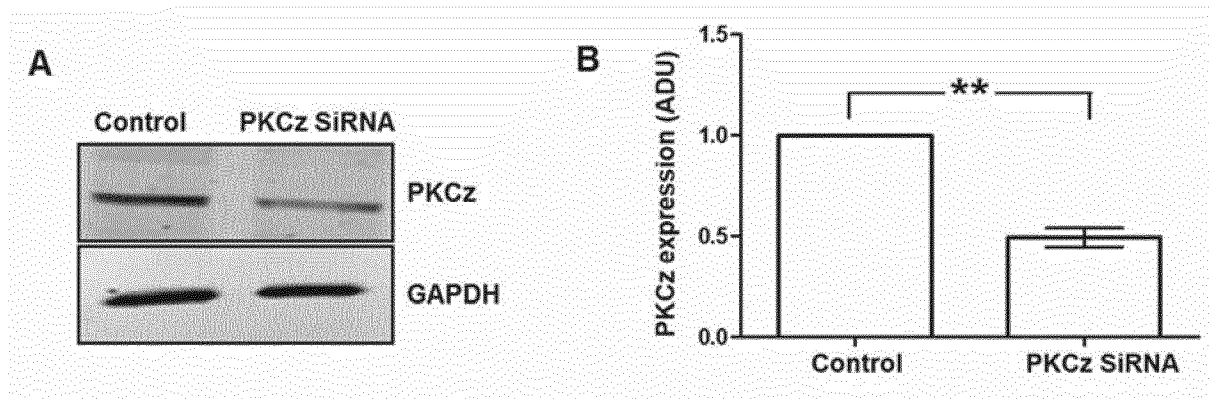


FIGURE 11

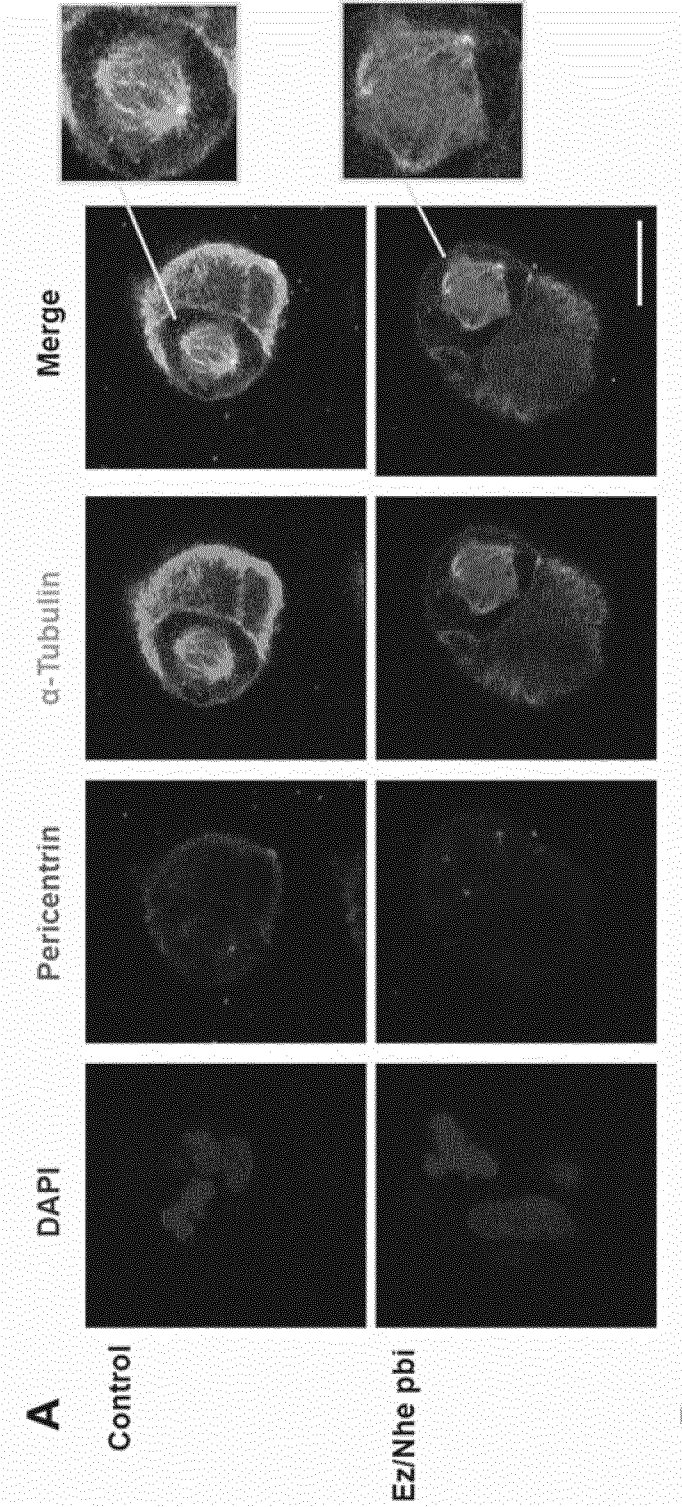


FIGURE 11

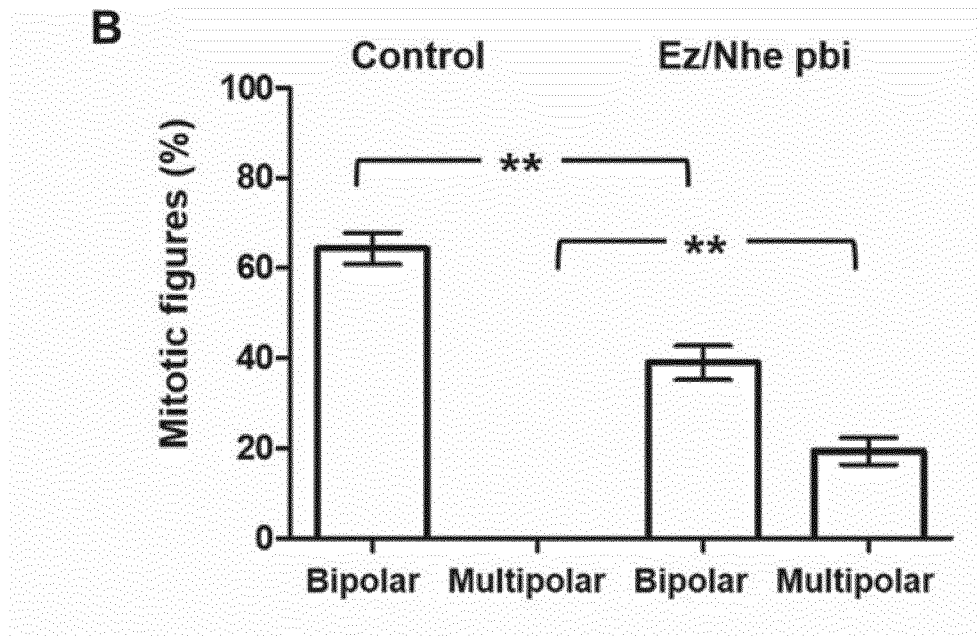
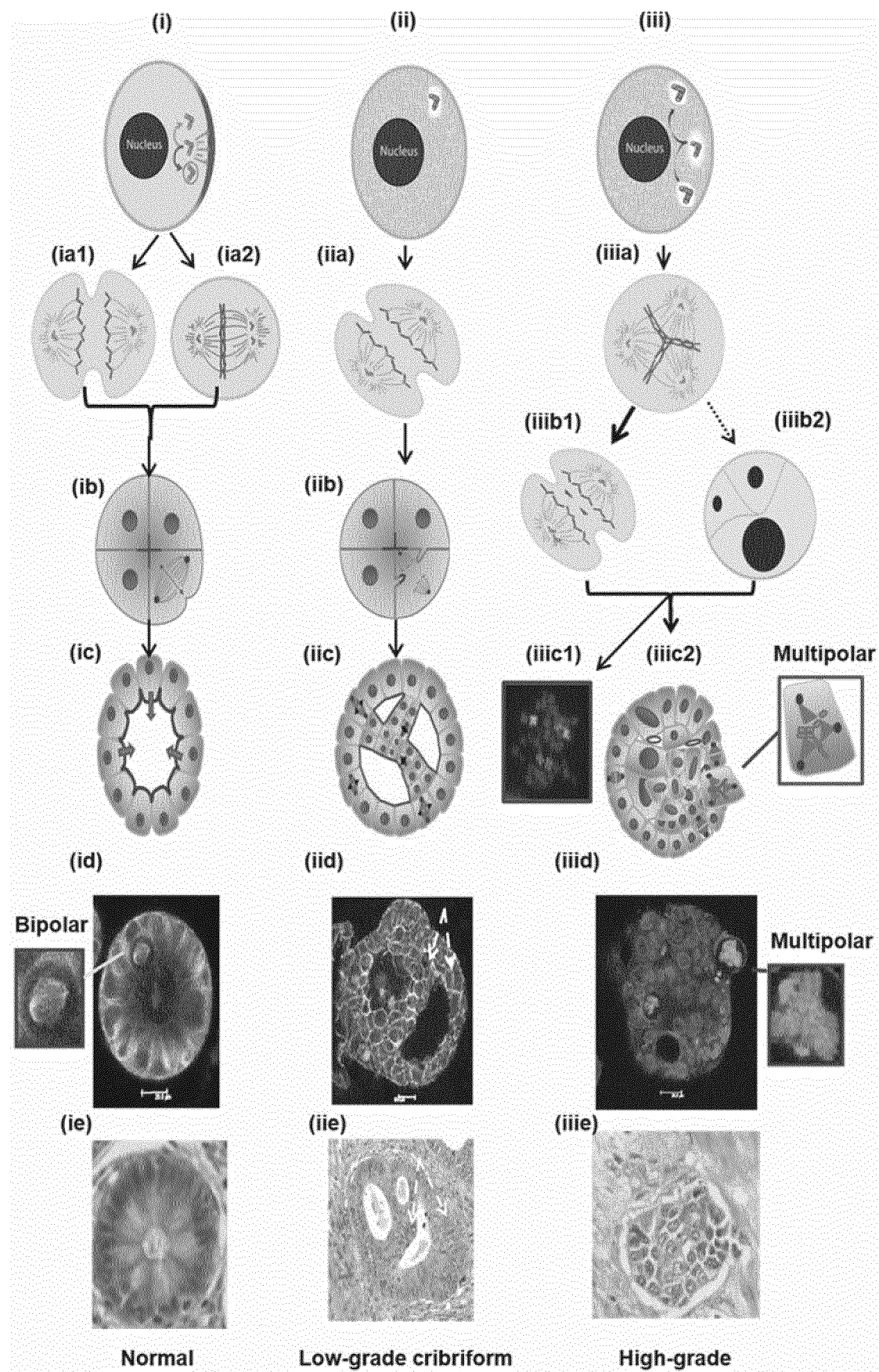


FIGURE 12



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2019/052149

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N5/07 G01N33/50 C12Q1/6886
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, BIOSIS, Sequence Search, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ARMAN JAVADI ET AL: "PTEN controls glandular morphogenesis through a juxtamembrane [beta]-Arrestin1/ARHGAP21 scaffolding complex", ELIFE, vol. 6, 27 July 2017 (2017-07-27), XP055572111, DOI: 10.7554/eLife.24578 the whole document ----- -/--	1-3,6, 10, 15-17,22



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

22 March 2019

Date of mailing of the international search report

05/04/2019

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Botz, Jürgen

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2019/052149

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	IMENE JAADANE ET AL: "The activation of the atypical PKC zeta in light-induced retinal degeneration and its involvement in L-DNase II control", JOURNAL OF CELLULAR AND MOLECULAR MEDICINE, vol. 19, no. 7, 1 July 2015 (2015-07-01), pages 1646-1655, XP055572125, RO ISSN: 1582-1838, DOI: 10.1111/jcmm.12539 the whole document -----	1-3,6, 10, 15-17,22
X	WO 2010/059400 A1 (TRANSLATIONAL GENOMICS RES INST [US]; KISLIN KERRI [US]; BERENS MICHAEL) 27 May 2010 (2010-05-27) the whole document -----	1-3,6, 10, 15-17,22
X	CHATTERJEE ISHITA ET AL: "Mo1765: 3D Cell Culture of CaCo2: A Better Model to Study the Functionality and Regulation of Intestinal Ion Transporters", GASTROENTEROLOGY, vol. 146, no. 5, Suppl. 1, May 2014 (2014-05), page S654, XP002789943, & 55TH ANNUAL MEETING OF THE SOCIETY-FOR-SURGERY-OF-THE-ALIMENTARY-TRACT (SSAT) / DIGESTIVE DISEASE WEEK; CHICAGO, IL, USA; MAY 03 -06, 2014 ISSN: 0016-5085, DOI: 10.1016/S0016-5085(14)62381-7 abstract -----	10-14
X	PIRMIN H. LAKNER ET AL: "Applying phasor approach analysis of multiphoton FLIM measurements to probe the metabolic activity of three-dimensional in vitro cell culture models", SCIENTIFIC REPORTS, vol. 7, no. 1, 13 February 2017 (2017-02-13), XP055572382, DOI: 10.1038/srep42730 the whole document -----	10-14

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2019/052149

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2010059400 A1	27-05-2010	EP 2350656 A1	03-08-2011
		US 2012245045 A1	27-09-2012
		US 2015176011 A1	25-06-2015
		WO 2010059400 A1	27-05-2010
