GPC3 CHIMERIC ANTIGEN RECEPTOR SECRETING

TECHNICAL FIELD

The present disclosure relates broadly to antigen binding proteins, chimeric antigen receptors, and modified cells specific to Epithelial Cellular Adhesion Molecule (EpCAM). In particular, the present disclosure relates to cell that expresses a chimeric antigen receptor targeting a tumour antigen, and a multi-specific antigen binding protein, variant or binding fragment thereof that binds one or more target, comprising a first antigen binding protein, variant or binding fragment thereof that binds to EpCAM (epithelial cell adhesion molecule) and a second antigen binding protein, variant or binding fragment thereof that binds to an immune cell marker.

BACKGROUND

Cancer incidence worldwide is expected to rise to 28.4 million cases in 2040. Among all adult human cancer cases, approximately 90% are solid tumour. Solid tumour such as liver cancer, lung cancer, stomach cancer and female breast cancer are among the leading causes of cancer-related death. Since the epithelial tissues are most abundantly found in the body, the malignancies of epithelial tissues, also known as “carcinomas”, account for 80 to 90 percent of all solid tumour cases. For example, hepatocellular carcinoma (HCC) is the most common form of liver cancer and accounts for ~90% of all liver cancer cases. It has a high mortality rate, with only 18% 5-year overall survival. As chemotherapy has fallen out of favour in HCC, the current standard of care in metastatic or unresectable patients with sufficient liver function is a combination of atezolizumab (anti-PD-L1) and bevacizumab (anti-VEGF). Thereafter, second line treatments are less well defined, and certainly there is a tremendous unmet need in post Atezo-Bev patients. The female breast cancers are another example, in which the epithelial tissue originated ductal carcinoma and lobular carcinoma are the most common primary breast cancer types and in total account for 90% of all breast cancer cases. Other examples include tumours found in skin, ovary (>90%), kidney (>85%), lung (>85%), stomach (~89%), pancreas (>90%), head and neck (>90%) and prostate gland (>95%), among which the majorities are carcinomas. Although the chance of cure for “carcinoma in situ” is high, the five-year survival rate for most invasive or metastatic carcinomas are

around 20-30%, even if multiple treatment modalities, such as surgery, chemotherapy, radiotherapy, and targeted therapies, are available.

Compared to conventional treatment strategies, antibody and cell-based therapies are gaining increasing attention and investments for its development as a potential more promising standard of care for the treatment of cancer. Cell-based therapy such as chimeric antigen receptor (CAR) T achieves great success in haematological malignancy but showed limited efficacy in treating solid tumours. Among the complex difficulties each approach has encountered, one intrinsic problem lies in the heterogeneous distribution of the tumour markers in the cell populations of solid tumours and antigen escape upon treatment.

Immune escape of target antigen-negative cells emerges as a major mechanism of cancer relapse. As a result, unsatisfactory efficacy of CAR-T cells targeting a single antigen has attributed to resistance to CAR T cell therapy. As tumour cells express diverse surface antigen repertoire, a potential approach to overcome this challenge is to engineer the CAR T cells with dual or multiple targeting to combat tumour escape, as exemplified by the bispecific anti-CD19/CD20 CAR T-cell therapy showing early promise to treat the relapsed or refractory B-cell Lymphoma. Nevertheless, it is extremely difficult to find multiple tumour specific targets on solid tumours. In fact, most solid tumour associated antigens (TAAs) are also expressed at low levels on normal tissues and hence the highly potent property of the CARs may induce intolerable on-target off-tumour toxicity.

Therefore, there is a need to provide an alternative method of preventing cancer relapse. There is a need to provide an alternative cell-based therapy. There is an urgent need to provide an alternative chimeric antigen receptor cell.

SUMMARY

Glypican 3 (GPC3) is an oncofoetal protein that can be used as a highly tumour-specific target. It is widely expressed during embryonic development; however, its expression is strictly suppressed in most adult tissues. Elevated GPC3 expression has been reported in a wide variety of tumour types such as liver, lung, gastric, ovarian, oesophageal and many others. Immunotherapies such as CAR T cell therapies targeting GPC3 has clinically proven to have some treatment efficacy and be generally safe. However, only a very limited number of patients achieved complete remission (CR). The metastasis and recurrence of the diseases following

therapies targeting these markers can be largely attributed to a fact that not all tumour cells express GPC3, though found at high levels in these cancers upon diagnosis. The re-appeared tumour masses are often derived from antigen-negative cancer cell growth or escaped cancer cells which subsequently down-regulated or even lost the antigen expression following CAR T cell therapies. Although with very high reported CR (complete remission), CAR T therapies using the FDA approved anti-CD19 directed "Kymriah" or "Yescarta" in treating various B cell originated leukaemia and lymphomas also faced the same problem of antigen escape. To overcome this challenge, the idea of CAR T cells with dual targeting was often applied and the bispecific anti-CD19/CD20 CAR T-cell therapy had shown early promise to treat the relapsed or refractory B-cell Lymphomas. However, on solid tumour, it is extremely hard to find two or more targetable tumour specific antigens that are highly specific to tumours, as most tumour markers are also expressed on normal tissues, albeit at relatively low levels. As CAR T cells have intrinsic properties of high potency in killing cancer cells, a design of dual CAR with multiple targeting of two or more antigens may induce intolerable on-target, off-tumour toxicity and hence are considered extremely dangerous to patients with solid tumours.

EpCAM (Epithelial Cellular Adhesion Molecule) represents another category of biomarkers that is widely expressed on almost all carcinomas. Cancer therapy targeting EpCAM has undergone more than 10 years of clinical development. However, since EpCAM is also expressed at low levels on normal epithelia, anti-EpCAM CAR T cells have been demonstrated to be highly toxic to normal tissues. Due to the same fact, both Solitomab and Catumaxomab (anti-EpCAM BiTE) showed dose-limiting toxicities and failed to achieve FDA approval.

Interestingly, as a broad-spectrum pan marker for epithelial cells, EpCAM is found to be co-expressed with GPC3 in many cancers. Instead of using dual CAR setting, a treatment that GPC3- targeting CAR T cells secreting anti-EpCAM bi-specific T cell engager (BiTE) was developed in this present disclosure so that the presence of anti-EpCAM BiTE is restricted at tumour site to contribute to tumour eradication and tumour escape prevention to the maximal extent and at the same time avoid systemic toxicity. These CAR T cells were named as "GE CAR-BiTE T" (GPC3 targeting CAR T cells with anti-EpCAM BiTE secretion.

CAR-BiTE T cells secreting anti-EpCAM BiTE first infiltrate and retain at the tumour site via CAR recognition of the specific tumour antigen expressed on the tumour cells. Upon CAR target engagement, the CAR T cells will be activated and

begin to expand and at the same time secrete anti-EpCAM BiTEs. While the CAR T cells can kill the target cells directly, the anti-EpCAM BiTEs exert their cytotoxicity by recruiting the T cells nearby. By this approach, bystander T cells are physically directed to the close proximity of tumours and at the same time being activated and subsequently help the clearance of tumour cells. Since GPC3 overexpression is tumour-specific, the expansion of the anti-EpCAM BiTEs is presumably restricted at or near the tumour sites. In addition, as the recombinant anti-EpCAM BiTE antibody secretion by the anti-GPC3 CAR T cells is not highly efficient, further diffusion to the normal tissues along with the requirement of just-in-time presence of the T cells would likely diminish its on-target, off-tissue toxicity. Moreover, EpCAM is also defined as a cancer stem cell marker as it is found to be expressed on cancer progenitor cells and cancer stem cells as well. Secretion of anti-EpCAM BiTE by CAR T cells will concomitantly contribute to the prevention of cancer relapse and recurrence by eliminating cancer stem cells and progenitor cells.

As shown by the experimental data of the present disclosure, the GE CAR-BiTE cells equipped with the new BiTEs (Nb01-013A or Nb01-013B) displayed superior properties in killing HepG2 and HT-29 cells at a comparable potency as MT110 BiTE. While T cells secreting MT110 also induced strong cytotoxicity against HeyA8 cells, GE CAR-BiTE T cells using the new BiTEs spared these target negative cells (FIGURE 5C, 5D and 5E). Furthermore, anti-EpCAM BiTE could be injected into tumour sites for targeted treatment. The anti-EpCAM BiTE of the present disclosure which is secreted by the CAR T cells only locally at the tumour site (and not administered systematically) restricts its potential toxicity. This turns an undruggable (or a difficult) target into a druggable target.

In one aspect, there is provided a modified cell expressing

(a) a chimeric antigen receptor targeting GPC3 and/or CD19, and

(b) a multi-specific antigen binding protein, variant or binding fragment thereof that binds one or more target, comprising a first antigen binding protein, variant or binding fragment thereof that binds to EpCAM (epithelial cell adhesion molecule) and a second antigen binding protein, variant or binding fragment thereof that binds to an immune cell marker,

wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region and/or a light chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising: (2C4, hu2C4, 1A5, 1B8, 2B7, and 2D10)

a CDR-H1 comprising:

- GSIFSGND (**SEQ ID NO: 25** - 2C4, hu2C4, 1A5, 2B7, and 2D10), or
- GSSERFTS (**SEQ ID NO: 29** – 1B8)

a CDR-H2 comprising:

- ITSGGST (**SEQ ID NO: 26** –2C4, hu2C4, 1A5, 2B7, and 2D10), or
- ITNGGST (**SEQ ID NO: 30** – 1B8) and

a CDR-H3 comprising:

- TNGRWSGDTYYAHH (**SEQ ID NO: 27** –2C4, hu2C4, 1A5, 2D10),
- MAGTS (**SEQ ID NO: 31** – 1B8), or
- TNGRWSGDTYYAHL (**SEQ ID NO: 33** – 2B7)

(ii) a heavy chain variable region comprising: (1B6, 1C1, 1C11, 1D4 and 1H6)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2)**, and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**)

(iii) a heavy chain variable region comprising: (1E4)

a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5)**,

a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6)**, and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7)**

(iv) a light chain variable region comprising: (1B6, 1C1, 1C11, 1D4, 1E4 and 1H6)

a CDR-L1 comprising:

- QSLLHSNGYNY (**SEQ ID NO: 9** - 1B6, 1C1, 1C11 and 1H6),
- QSLLHSNRYNY (**SEQ ID NO: 17** - 1D4), or
- QSISDF (**SEQ ID NO: 19** – 1E4)

a CDR-L2 comprising:

- LGS (**SEQ ID NO: 10** -1B6, 1C1, 1C11, 1D4 and 1H6), or
- AAS (**SEQ ID NO: 20** - 1E4), and

a CDR-L3 comprising:

- MQALQTPYT (**SEQ ID NO: 11** - 1B6, 1C1 and 1D4),
- MQGLQSPWT (**SEQ ID NO: 15** - 1C11),

- QQSYIMPDT (**SEQ ID NO: 21** – 1E4), or
- MQGLQTPYT (**SEQ ID NO: 23** – 1H6); and

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region comprising: (1A5, 1B8, 2B7, 2C4, 2D10 and hu2C4)

a CDR-H1 comprising:

- GSIFSGND (**SEQ ID NO: 25** - 1A5, 2B7, 2C4, 2D10 and hu2C4), or
- GSSERFTS (**SEQ ID NO: 29** – 1B8)

a CDR-H2 comprising:

- ITSGGST (**SEQ ID NO: 26** – 1A5, 2B7, 2C4, 2D10 and hu2C4), or
- ITNGGST (**SEQ ID NO: 30** – 1B8); and

a CDR-H3 comprising:

- TNGRWSGDTYYAHH (**SEQ ID NO: 27** – 1A5, 2C4, 2D10 and hu2C4),
- MAGTS (**SEQ ID NO: 31** – 1B8), or
- TNGRWSGDTYYAHL (**SEQ ID NO: 33** – 2B7)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising: (2C4-VHH, hu2C4-VHH, 1A5-VHH, and 2D10-VHH)

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**),

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**), and

a CDR-H3 comprising TNGRWSGDTYYAHH (**SEQ ID NO: 27**)

(ii) a heavy chain variable region comprising: (1B8-VHH)

a CDR-H1 comprising GSSERFTS (**SEQ ID NO: 29**),

a CDR-H2 comprising ITNGGST (**SEQ ID NO: 30**), and

a CDR-H3 comprising MAGTS (**SEQ ID NO: 31**); and

(iii) a heavy chain variable region comprising: (2B7-VHH)

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**),

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**), and

a CDR-H3 comprising TNGRWSGDTYYAHL (**SEQ ID NO: 33**)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region and/or a light chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising: (2C4, hu2C4, 1A5, and 2D10)

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**),

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**), and

a CDR-H3 comprising TNGRWSGDTYYAHH (**SEQ ID NO: 27**)

(ii) a heavy chain variable region comprising: (1B6, 1C1, 1C11, 1D4, 1H6)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**)

(iii) a heavy chain variable region comprising: (1E4)

a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**),

a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**), and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**)

(iv) a light chain variable region comprising: (1B6 and 1C1)

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**)

(v) a light chain variable region comprising: (1C11)

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**)

(vi) a light chain variable region comprising: (1D4)

a CDR-L1 comprising QSLLHSNRYNY (**SEQ ID NO: 17**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**)

(vii) a light chain variable region comprising: (1E4)

a CDR-L1 comprising QSISDF (**SEQ ID NO: 19**),

a CDR-L2 comprising AAS (**SEQ ID NO: 20**), and

a CDR-L3 comprising QQSYIMPDT (**SEQ ID NO: 21**)

(viii) a light chain variable region comprising: (1H6)

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**)

(ix) a heavy chain variable region comprising: (1B8-VHH)

a CDR-H1 comprising GSSERFTS (**SEQ ID NO: 29**),

a CDR-H2 comprising ITNGGST (**SEQ ID NO: 30**), and

a CDR-H3 comprising MAGTS (**SEQ ID NO: 31**); and

(x) a heavy chain variable region comprising: (2B7-VHH)

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**),

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**), and

a CDR-H3 comprising TNGRWSGDTYYAHL (**SEQ ID NO: 33**)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain and/or a light chain variable domain selected from the group consisting of:

(i) a heavy chain variable domain comprising QVQLQESGGGLVQAGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE LVAV**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 35** - 2C4-VHH)

(ii) a heavy chain variable domain comprising QVQLVESGGGLVQAGGSLRLSCAAS**GSIFSGND**MSWYRQAPGKGLE LVAV**ITSGGST**YYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYY C**TNGRWSGDTYYAHH**WGQGTL (**SEQ ID NO: 37** - hu2C4-VHH)

(iii) a heavy chain variable domain comprising EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLE WMGG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVY YC**ARSLGGRFRY**WGQGTL (**SEQ ID NO: 4** - 1B6, 1C1, 1C11, 1D4 and 1H6)

(iv) a heavy chain variable domain comprising QVQLQQSGPGLVKPSQTLSLTCAIS**GDSISSNSVA**WNWIRQSPSRGL EWLGR**TYYRSKWYS**DYAISVKGRLDINPDTSKNQFSLQLNSVTPEDT AVYYC**AREVEGSSYDAFDI**WGQGTM (**SEQ ID NO: 8** - 1E4),

(v) a light chain variable domain comprising
DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQ
SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**M**
QALQTPYTFGQGTK (**SEQ ID NO: 12** - 1B6 and 1C1)

(vi) a light chain variable domain comprising
EIVLTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQS
PQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQ**
GLQSPWTFGQGTK (**SEQ ID NO: 16** - 1C11)

(vii) a light chain variable domain comprising
DVVMTQSPLSLPVTPGESASISCRSS**QSLLHSNRYNY**LDWYLQKPGQ
SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**M**
QALQTPYTFGQGTK (**SEQ ID NO: 18** - 1D4)

(viii) a light chain variable domain comprising
DIQLTQSPSSLSASVGDRVTITCRAS**QSISDF**LNWYQQKPGKAPKLLIY
AASSLQTGVPSRFGGSGSGTEFTLTISSLQPEDLGTYYC**QQSYIMPDT**
FGQGTK (**SEQ ID NO: 22** - 1E4)

(ix) a light chain variable domain comprising
DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQ
SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYC**M**
QGLQTPYTFGQGTK (**SEQ ID NO: 24** - 1H6)

(x) a heavy chain variable domain comprising
QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVAV**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 28-** 1A5-VHH)

(xi) a heavy chain variable domain comprising
QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERE
LVAF**ITNGGST**RYTDPVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYY
C**MAGTS**WGQGTQ (**SEQ ID NO: 32** - 1B8-VHH)

(xii) a heavy chain variable domain comprising
QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVAV**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHL**WGQGTQ (**SEQ ID NO: 34** - 2B7-VHH)

(xiii) a heavy chain variable domain comprising
QVQLQESGGGLVQAGDSLRLSCADS**GSIFSGND**MAWYRRAPGVERE

LVA**VITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 36** - 2D10-VHH); and

or fragment or variant or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions.

In some embodiments, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a single domain heavy chain variable domain having a sequence:

(i) QVQLQESGGGLVQAGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVA**VITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 35** - clone 2C4-VHH), or

(ii) QVQLVESGGGLVQAGGSLRLSCAAS**GSIFSGND**MSWYRQAPGKGLE
LVA**VITSGGST**YYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYY
C**TNGRWSGDTYYAHH**WGQGTL (**SEQ ID NO: 37** - hu2C4-VHH), or

(iii) QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVA**VITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 28** - clone 1A5-VHH), or

(iv) QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERE
LVAF**ITNGGST**RYTDPVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYY
C**MAGTS**WGQGTQ (**SEQ ID NO: 32** - clone 1B8-VHH), or

(v) QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVA**VITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHL**WGQGTQ (**SEQ ID NO: 34** - clone 2B7-VHH), or

(vi) QVQLQESGGGLVQAGDSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVA**VITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 36** - clone 2D10-VHH),

or fragment or variant or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions.

In some embodiments, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising: (1B6, 1C1, 1C11, 1D4 and 1H6)
a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2)**, and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**)

(ii) a heavy chain variable region comprising: (1E4)

a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5)**,

a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6)**, and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7)**; and

comprises a light chain variable region comprising: (1B6, 1C1, 1C11, 1D4, 1E4 and 1H6)

a CDR-L1 comprising:

- QSLLHSNGYNY (**SEQ ID NO: 9** - 1B6, 1C1, 1C11 and 1H6),
- QSLLHSNRYNY (**SEQ ID NO: 17** - 1D4), or
- QSISDF (**SEQ ID NO: 19 –** 1E4)

a CDR-L2 comprising:

- LGS (**SEQ ID NO: 10** -1B6, 1C1, 1C11, 1D4 and 1H6), or
- AAS (**SEQ ID NO: 20** - 1E4), and

a CDR-L3 comprising:

- MQALQTPYT (**SEQ ID NO: 11** - 1B6, 1C1 and 1D4),
- MQGLQSPWT (**SEQ ID NO: 15** - 1C11),
- QQSYIMPDT (**SEQ ID NO: 21 –** 1E4), or
- MQGLQTPYT (**SEQ ID NO: 23 –** 1H6)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region and a light chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising: (1B6 and 1C1)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**); and

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**)

(ii) a heavy chain variable region comprising: (1C11)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**); and

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**)

(iii) a heavy chain variable region comprising: (1D4)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**); and

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNRYNY (**SEQ ID NO: 17**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**)

(iv) a heavy chain variable region comprising: (1H6)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**); and

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**); and

(v) a heavy chain variable region comprising: (1E4)

a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**),

a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**), and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**); and

a light chain variable region comprising: (1E4)

a CDR-L1 comprising QSISDF (**SEQ ID NO: 19**),

a CDR-L2 comprising AAS (**SEQ ID NO: 20**), and

a CDR-L3 comprising QQSYIMPDT (**SEQ ID NO: 21**)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some embodiments, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain and a light chain variable domain selected from the group consisting of:

(i) a heavy chain variable domain comprising: (1B6 and 1C1)
EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLEWM
GG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYC**ARS**
LGGRFRYWGQGTL (**SEQ ID NO: 4**), and
a light chain variable domain comprising
DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQSP
QLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQALQT**
PYTFGQGTK (**SEQ ID NO: 12**)

(ii) a heavy chain variable domain comprising: (1C11)
EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLEWM
GG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYC**ARS**
LGGRFRYWGQGTL (**SEQ ID NO: 4**), and
a light chain variable domain comprising
EIVLTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQSPQ
LLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQGLQSP**
WTFGQGTK (**SEQ ID NO: 16**)

(iii) a heavy chain variable domain comprising: (1D4)
EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLEWM
GG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYC**ARS**
LGGRFRYWGQGTL (**SEQ ID NO: 4**), and
a light chain variable domain comprising
DVVMTQSPLSLPVTPGESASISCRSS**QSLLHSNRYNY**LDWYLQKPGQSP
QLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQALQT**
PYTFGQGTK (**SEQ ID NO: 18**)

(iv) a heavy chain variable domain comprising: (1E4)
QVQLQQSGPGLVKPSQTLSLTCAIS**GDSISSNSVA**WNWIRQSPSR
GLEWLGR**TYYRSKWYS**DYAISVKGRLDINPDTSKNQFSLQLNSVT
PEDTAVYYC**AREVEGSSYDAFDI**WGQGTM (**SEQ ID NO: 8**), and
a light chain variable domain comprising
DIQLTQSPSSLSASVGDRVTITCRAS**QSISDF**LNWYQQKPGKAPKLLIY**AA**
SSLQTGVPSRFGGSGSGTEFTLTISSLQPEDLGTYYC**QQSYIMPDT**FGQG
TK (**SEQ ID NO: 22**), or

(v) a heavy chain variable domain comprising: (1H6)

EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLEWM GG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYC**ARS LGGRFRY**WGQGTL (**SEQ ID NO: 4**), and

a light chain variable domain

DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQSP QLLIY**LGS**NRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYC**MQGLQT PYT**FGQGTK (**SEQ ID NO: 24**), or

or fragment or variant or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions.

In some embodiments, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain encoded by a nucleotide sequence comprising:

(i) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG GGGCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG TGGCAATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT ACTATGCCCATCAC**TGGGGCCAGGGGACCCAG (**SEQ ID NO: 72** - clone 2C4), or

(ii) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGG GGGCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAG TGGCAATGAC**ATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACT CGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**TACTATGC AGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCaAG AAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGG CCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACT ATGCCCATCAC**TGGGGCCAGGGGACCCTG (**SEQ ID NO: 74 -** clone hu2C4-VHH); or

(iii) GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG CAGCTATGCT**ATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT

TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO:** 41 - for clones 1B6, 1C1, 1C11, 1D4, 1H6), or

(iv) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC**
TAGTAACAGTGTTGCTTGGAACTGGATCAGGCAGTCCCCATCGAG
AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
ACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC
AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG
(**SEQ ID NO: 45** - clone 1E4); or

(v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG
CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC
GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA**
CTATGCCCATCACTGGGGCCAGGGGACCCAG (**SEQ ID NO: 65** - clone 1A5), or

(vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
GGGGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAG**
ATTCACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGC
GCGAGTTGGTCGCATTT**ATTACTAATGGTGGTAGCACA**AGATATAC
AGACCCCGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAA
GAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACAC
GGCCGTCTATTATTGT**ATGGCGGGTACGTCC**TGGGGCCAGGGGAC
CCAG (**SEQ ID NO: 69** - clone 1B8), or

(vii)CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC

GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG
CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCATCTCTGGGGCCAGGGGACCCAG (**SEQ ID NO: 71** -
clone 2B7), or

(viii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGG
CTGGGGACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCT**
TCAGTGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTG
GAGCGCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CAC
TATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAAT
GCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAG
GACACGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGA**
TACTTACTATGCCCATCACTGGGGCCAGGGGACCCAG (**SEQ ID**
NO: 73 - clone 2D10), and/or

the light chain variable domain is encoded by a nucleotide sequence comprising:

(i) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 49** - clone 1B6 and 1C1), or

(ii) GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGTACAGATTT
TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGGTCTACAAAGTCCCTGGACG**TTCGGCCAAGGG
ACCAAG (**SEQ ID NO: 53** - clone 1C11), or

(iii) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATAGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG

GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 55** - clone 1D4), or

(iv) GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT
CCTGATCTAT**GCTGCATCG**AGTTTACAAACTGGGGTCCCCTCAAGA
TTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCACCATAAGCA
GTCTACAACCTGAAGATTTGGGAACTTATTACTGT**CAACAGAGTTA**
CATTATGCCCGACACTTTTGGCCAGGGGACGAAA (**SEQ ID NO: 59** - clone 1E4), or

(v) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
TTACTGC**ATGCAAGGTCTACAGACTCCGTACACT**TTTGGCCAGGG
GACCAAG (**SEQ ID NO: 61** - clone 1H6),

or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions.

In some embodiments, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises heavy chain and/or light chain variable domains encoded by nucleotide sequences selected from the group consisting of:

(i) a heavy chain variable domain encoded by the nucleotide sequence comprising: (1B6 and 1C1)

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA

CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**), and/or
a light chain variable domain encoded by the nucleotide sequence comprising:
GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 49**)

(ii) a heavy chain variable domain encoded by the nucleotide sequence comprising: (1C11)
GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**), and/or
a light chain variable domain encoded by the nucleotide sequence comprising:
GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGTACAGATTT
TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGGTCTACAAAGTCCCTGGACG**TTCGGCCAAGGG
ACCAAG (**SEQ ID NO: 53**)

(iii) a heavy chain variable domain encoded by the nucleotide sequence comprising: (1D4)

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**), and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATAGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 55**)

(iv) a heavy chain variable domain encoded by the nucleotide sequence comprising: (1H6)

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**), and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT

TACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
TTACTGC**ATGCAAGGTCTACAGACTCCGTACACT**TTTGGCCAGGG
GACCAAG (**SEQ ID NO: 61**); and

(v) a heavy chain variable domain encoded by the nucleotide sequence comprising: (1E4)

CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC**
TAGTAACAGTGTTGCTTGGAACTGGATCAGGCAGTCCCCATCGAG
AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
ACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC
AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG
(**SEQ ID NO: 45**), and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT
CCTGATCTAT**GCTGCATCG**AGTTTACAAACTGGGGTCCCCTCAAGA
TTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCACCATAAGCA
GTCTACAACCTGAAGATTTGGGAACTTATTACTGT**CAACAGAGTTA**
CATTATGCCCGACACTTTTGGCCAGGGGACGAAA (**SEQ ID NO: 59**)

or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain encoded by the nucleotide sequence comprising:

(i) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGG
GCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG
TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC**TGG

GGCCAGGGGACCCAG (**SEQ ID NO: 72**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(ii) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGG GCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAGTGGCA ATGAC**ATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACTCGAGTTG GTCGCGGTT**ATTACTAGCGGTGGTAGTACA**TACTATGCAGACTCCGTG AAGGGCCGATTCACCATCTCCAGAGACAATTCCaAGAAcACCcTATATC TGCAAATGAACAGCCTGAGAGCTGAGGACACGGCCGTGTATTACTGC**A CAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC**TGGGGC CAGGGGACCCTG (**SEQ ID NO: 74**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions; or

(iii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAG GGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA ATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT ATCTGCAAACGAACGACTTGAAACCTGAGGACACGGCCGTGTATTACT GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC**TGG GGCCAGGGGACCCAG (**SEQ ID NO: 65**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(iv) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGGG GGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAGATTC ACATCA**GTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGCGCGAGTT GGTCGCATTT**ATTACTAATGGTGGTAGCACA**AGATATACAGACCCCGT GAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGAACACGGTGT ATCTGCAAATGAACAGCCTGAAAGCTGAGGACACGGCCGTCTATTATT GT**ATGGCGGGTACGTCC**TGGGGCCAGGGGACCCAG (**SEQ ID NO: 69**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAG GGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA ATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT

GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG
TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCTC**TGG
GGCCAGGGGACCCAG (**SEQ ID NO: 71**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGG
ACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG
TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC**TGG
GGCCAGGGGACCCAG (**SEQ ID NO: 73**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions,
GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGTC
CTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAGCAGCT**
ATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCTTGAGTGG
ATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTACGCACAGAAC
TTCCAGGGCAGAGTCACCATGACCGCAGACACCTCCATAAGCACAGC
CTACATGGAGCTGAGCAGCCTGAGATCTGAGGACACGGCCGTGTATTA
CTGT**GCGAGATCGTTGGGTGGGAGATTTCGCTAC**TGGGGCCAGGGAA
CCCTG (**SEQ ID NO:** 41) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(vii) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTCGCA
GACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTCTAGTAA**
CAGTGTTGCTTGGAACTGGATCAGGCAGTCCCCATCGAGAGGCCTTG
AGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGTACAGT**GATTAT
GCAATATCTGTGAAAGGTCGATTAGACATCAACCCAGACACATCCAAG
AACCAGTTCTCCCTGCAGCTGAACTCTGTGACTCCCGAGGACACGGCT
GTGTATTATTGT**GCAAGAGAAGTTGAGGGCAGCAGCTATGATGCTTTT**
GATATCTGGGGCCAAGGGACAATG (**SEQ ID NO: 45**) or a sequence at

least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions.

In some embodiments, the multi-specific antigen binding protein, variant or fragment thereof is a bispecific antibody.

In some embodiments, the multi-specific antigen binding protein, variant or fragment thereof is an immune cell engager selected from the group comprising a T cell engager, an NK cell engager, a monocyte engager and a macrophage engager.

In some embodiments, the multi-specific antigen binding protein, variant or fragment thereof is a bispecific T cell engager (BiTE), such as an inducible BiTE, a non-inducible BiTE or a constitutive expression BiTE.

In some embodiments, the second antigen binding protein, variant or binding fragment thereof of the immune cell engager binds to the immune marker selected from the group consisting of CD3, NKG2D, CD4, CD8, CD16, and CD64.

In some embodiments, the multi-specific antigen binding protein is an inducible bispecific T cell engager comprising a Heavy chain antibody variable region (i.e. VHH) and/or a single chain variable fragment (scFv).

In some embodiments, the cell for example selected from the group consisting of a T cell, a macrophage, a monocyte, and an NK cell.

In some embodiments, the cell is a T cell, optionally a CAR T-cell.

In some embodiments, the cell

i. binds to GPC3 and secretes an inducible bispecific T cell engager that targets EpCAM and CD3 (GE CAR-BiTE T); or

ii. binds to CD19 and secretes an inducible bispecific T cell engager that targets EpCAM and CD3 (CD19 CAR-BiTE T).

In another aspect, there is provided a polynucleotide encoding the cell as described herein.

In yet another aspect, there is provided a vector expressing the polynucleotide as described herein.

In yet another aspect, there is provided a host cell comprising the vector as described herein.

In yet another aspect, there is provided a method of producing / generating the cell as described herein, comprising introducing the polynucleotide as described herein into the cell.

In yet another aspect, there is provided a composition comprising the cell as described herein.

In yet another aspect, there is provided a method of treating a disease in a subject in need thereof, the method comprises administering to the subject the cell or composition as described herein, optionally the disease is a proliferative disease, optionally a cancer.

DEFINITIONS

The term "antigen binding protein" herein is used in the broadest sense and encompasses various antibody structures, including but not limited to monoclonal antibodies, polyclonal antibodies, multi-specific antibodies as long as they exhibit the desired antigen-binding activity.

The term "antibody" as used herein relates to whole (i.e., full length) antibodies (i.e., comprising the elements of two heavy chains and two light chains) and functionally active fragments thereof (i.e., molecules that contain an antigen binding domain that specifically binds an antigen, also termed antibody fragments or antigen-binding fragments). Features described herein with respect to antibodies also apply to antibody fragments unless context dictates otherwise. The term "antibody" encompasses monovalent, i.e., antibodies comprising only one antigen binding domain (e.g., one-armed antibodies comprising a full-length heavy chain and a full-length light chain interconnected, also termed "half-antibody"), and multivalent antibodies, i.e., antibodies comprising more than one antigen binding domain, e.g., bivalent.

The term "antigen binding fragment" as employed herein refers to functionally active antibody binding fragments including but not limited to Fab, modified Fab, Fab', modified Fab', F(ab')2, Fv, single domain antibodies, scFv, Fv, bi, tri or tetra-valent antibodies, Bis-scFv, diabodies, triabodies, tetrabodies and epitope-binding fragments of any of the above.

A "binding fragment" as employed herein refers to a fragment capable of binding a target peptide or antigen with sufficient affinity to characterize the fragment as specific for the peptide or antigen.

The term "monoclonal antibody" (or "mAb") refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e. each individual of a monoclonal antibody preparation is identical except for possible mutations (e.g., naturally occurring mutations), that may be present in minor amounts. Certain differences in the protein sequences linked to post-translational modifications (for example, cleavage of the heavy chain C-terminal lysine, deamidation of asparagine residues and/or isomerization of aspartate residues) may nevertheless exist between the various different antibody

molecules present in the composition. Contrary to polyclonal antibody preparations, each monoclonal antibody of a monoclonal antibody preparation is directed against a single determinant on an antigen.

The term "diabody" as employed herein refers to two Fv pairs, a first VH/VL pair and a further VH/VL pair which have two inter-Fv linkers, such that the VH of a first Fv is linked to the VL of the second Fv and the VL of the first Fv is linked to the VH of the second Fv.

The term "tribody" (also referred to a Fab(scFv)2) as employed herein refers to a Fab fragment with a first scFv appended to the C-terminal of the light chain and a second scFv appended to the C-terminal of the heavy chain. The term "tetrabody" as employed herein refers to a format similar to the diabody comprising fours Fvs and four inter-Fv linkers.

The term "multivalent antibody" refers to an antibody comprising more than one antigen binding domain e.g., bivalent.

The term "Fv" refers to two variable domains of full-length antibodies, for example co-operative variable domains, such as a cognate pair or affinity matured variable domains, i.e., a VH and VL pair. The term "scFv" refers to single chain variable fragment which is a fusion protein of the variable regions of the heavy and light chains of the immunoglobulins, connected with a short linker peptide of ten to about 25 amino acids. The term "bis-scFv" as described herein refers to a bispecific scFv.

The term "dsscFv" or "disulphide-stabilised single chain variable fragment" as employed herein refers to a single chain variable fragment which is stabilised by a peptide linker between the VH and VL variable domain and also includes an inter-domain disulphide bond between VH and VL.

The term "DVD-Ig" (also known as dual V domain IgG) refers to a full-length antibody with 4 additional variable domains, one on the N-terminus of each heavy and each light chain.

The term "Fab" refers to as used herein refers to an antibody fragment comprising a light chain fragment comprising a VL (variable light) domain and a constant domain of a light chain (CL), and a VH (variable heavy) domain and a first constant domain (CHI) of a heavy chain. Dimers of a Fab' according to the present disclosure create a F(ab')2 where, for example, dimerization may be through the hinge. The term "F(ab')" refers to a monovalent fragment of a single light chain homodimer, which is obtained by pepsin digestion of IgG, followed by reduction of the light chain disulfide bond. The term "F(ab')2" as described herein refers to a fragment of IgG that is prepared by pepsin

digestion of IgG. The F(ab')2 fragment is a disulfide-linked homodimer of the two light chain dimers, so it retains bivalent epitope binding like whole IgG, but as it lacks the heavy chains, it is smaller in size compared to a whole IgG. F(ab')2 and F(ab') fragments do not bind to immunoglobulin receptors on cells, which can be useful for achieving specific staining of the primary antibody target.

The terms "constant domain(s)" or "constant region", as used herein are used interchangeably to refer to the domain(s) of an antibody which is outside the variable regions. The constant domains are identical in all antibodies of the same isotype but are different from one isotype to another. Typically, the constant region of a heavy chain is formed, from N to C terminal, by CH1 -hinge -CH2-CH3-optionally CH4, comprising three or four constant domains.

The term "DiFab" as employed herein refers to two Fab molecules linked via their C-terminus of the heavy chains or two Fab' molecules linked via one or more disulfide bonds in the hinge region thereof.

The term "antigen binding variant" refers to a polypeptide, for example, an antibody possessing the desired characteristics described herein and comprising a VH and/or a VL that has at least about 80% amino acid sequence identity with a VH and/or a VL of the reference antibody. Such antibody variants include, for instance, antibodies wherein one or more amino acid residues are added to or deleted from the VH and/or a VL domain. Ordinarily, an antibody variant will have at least about 80% amino acid sequence identity, alternatively at least about 85%, 90%, 95%, 96%, 97%, 98%, or 99% amino acid sequence identity, to an antibody described herein. Optionally, variant antibodies will have no more than one conservative amino acid substitution as compared to an antibody sequence provided herein, alternatively no more than about any of 2, 3, 4, 5, 6, 7, 8, 9, or 10 conservative amino acid substitutions as compared to an antibody sequence provided herein.

The term "immune cell" refers to a type of specialized cell that plays a crucial role in the body's defense against infections and foreign substances. They are a part of the immune system, which is responsible for identifying and eliminating harmful pathogens, such as bacteria, viruses, and parasites, as well as abnormal or cancerous cells. As used herein, an 'immune cell' refers to any cell of the immune system, including but not limited to T-cells, helper T-cells, B-cells, natural killer (NK) cells, dendritic cells (DC), granulocytes (such as basophils, eosinophils, neutrophils), mast cells, monocytes, and macrophages.

The term “specifically” as employed herein in the context of antibodies is intended to refer to an antibody that only recognizes the antigen to which it is specific or an antibody that has significantly higher binding affinity to the antigen to which it is specific compared to binding to antigens to which it is non-specific, for example at least 5, 6, 7, 8, 9, 10 times higher binding affinity.

The term "epitope" or “binding site” in the context of antibodies refers to a site (or a part) on an antigen to which the paratope of an antibody binds or recognizes. Epitopes can be formed both from contiguous amino acids (also often called “linear epitopes”) or non-contiguous amino acids formed by tertiary folding of a protein (often called “conformational epitopes”). Epitopes formed from contiguous amino acids are typically retained on exposure to denaturing solvents whereas epitopes formed by folding are typically lost on treatment with denaturing solvents. An epitope typically includes at least 3, and more usually, at least 5-10 amino acids in a unique spatial conformation. Epitopes usually consist of chemically active surface groups of molecules such as amino acids, sugar side chains and usually have specific 3D structural and charge characteristics.

The "class" of an antibody refers to the type of constant domain or constant region possessed by its heavy chain. There are five major classes of antibodies: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA1, and IgA2. The heavy chain constant domains that correspond to the different classes of immunoglobulins are called a, d, e, g, and m, respectively.

The term "chimeric antibody" (or antigen-binding fragment thereof) is an antibody molecule (or antigen-binding fragment thereof) in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, e. g., an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity. For example, a mouse antibody can be modified by replacing its constant region with the constant region from a human immunoglobulin. Due to the replacement with a human constant region, the chimeric antibody can retain its specificity in recognizing the antigen while having reduced antigenicity in human as compared to the original mouse antibody.

The term “chimeric antigen receptors” refers to receptor protein that has been engineered to give T cells the new ability to target a specific antigen. The receptors are

chimeric in that they combine both antigen-binding and T cell activating functions into a single receptor. CAR T cell therapy uses T cells engineered with CARs to treat cancer. T cells in CAR T immunotherapy are modified to recognize cancer cells in order to more effectively target and destroy them. CAR T cells can be derived either from T cells in a patient's own blood (autologously) or from the T cells of another, healthy, donor (allogeneically). Once isolated from a person, these T cells are genetically engineered to express a specific CAR, which programs them to target an antigen that is present on the surface of tumours. For safety, CAR T cells are engineered to be specific to an antigen that is expressed on a tumour but is not expressed on healthy cells. CAR T cells destroy cells through extensive stimulated cell proliferation, increasing the degree to which they are toxic to other living cells (cytotoxicity) and by causing the increased secretion of factors that can affect other cells such as cytokines, interleukins and growth factors. The surface of CAR T cells can bear either two types of co-receptors, CD4 and CD8, each with different and interacting cytotoxic effects.

The term "human antibody" or "humanized antibody" (or antigen-binding fragment thereof), as used herein, is intended to include antibodies (and antigen-binding fragments thereof) having variable regions in which both the framework and CDR regions are derived from sequences of human origin. Depending on the amino acid sequence of the constant region of their heavy chains, antibodies or immunoglobulins are divided into the classes: IgA, IgD, IgE, IgG and IgM, and several of these may be further divided into subclasses (subtypes), e.g. IgG1, IgG2, IgG3, and IgG4, IgA1, and IgA2. Therefore, human IgG constant region domains may be used, especially of the IgG1 and IgG3 isotypes when the antibody molecule is intended for therapeutic uses and antibody effector functions are required. Alternatively, IgG2 and IgG4 isotypes may be used when the antibody molecule is intended for therapeutic purposes and antibody effector functions are not required. Furthermore, if the antibody contains a constant region, the constant region also is derived from such human sequences. A humanized antibody (or antigen-binding fragment thereof) retains the reactivity of a non-human antibody while being less immunogenic in humans. This can be achieved, for instance, by retaining the non-human CDR regions and replacing the remaining parts of the antibody with their human counterparts (i.e., the constant region as well as the framework portions of the variable region). Additional framework region modifications may be made within the human framework sequences as well as within the CDR sequences derived from the germline of another mammalian species. The humanized antibodies of the present disclosure may include amino acid residues not encoded by human sequences (e. g.,

mutations introduced by random or site-specific mutagenesis *in vitro* or by somatic mutation *in vivo*, or a conservative substitution to promote stability or manufacturing). This definition of a humanized antibody specifically excludes a humanized antibody comprising non-human antigen-binding residues. Human antibodies can be produced using various techniques known in the art, including phage-display libraries, administering the antigen to a transgenic animal that has been modified to produce such antibodies in response to antigenic challenge, but whose endogenous loci have been disabled, e. g., immunized Xeno mice via a human B-cell hybridoma technology.

The term "recombinant humanized antibody" as used herein, includes all human antibodies that are prepared, expressed, created or isolated by recombinant means, such as antibodies isolated from a host cell transformed to express the humanized antibody, e. g., from a transfectoma, and antibodies prepared, expressed, created or isolated by any other means that involve splicing of all or a portion of a human immunoglobulin gene, sequences to other DNA sequences.

The term "isolated" means, throughout this specification, that the antibody, or polynucleotide, as the case may be, exists in a physical milieu distinct from that in which it may occur in nature. The term "isolated" nucleic acid refers to a nucleic acid molecule that has been isolated from its natural environment or that has been synthetically created. An isolated nucleic acid may comprise synthetic DNA, for instance produced by chemical processing, cDNA, genomic DNA or any combination thereof. An isolated antibody refers to an antibody that is substantially free of other cellular material and/or chemicals.

The term "Complementarity Determining Regions" ("CDRs") refers to amino acid sequences with boundaries determined using any of a number of well-known schemes, including those described by Kabat (i.e., "Kabat" numbering scheme); Al-Lazikani ("Chothia" numbering scheme); ImMunoGenTics (IMGT) numbering ("IMGT" numbering scheme); and the like. The term "Complementarity Determining Regions" ("CDRs") refers to regions of hypervariability that contain the binding domain that interacts with an antigen. Antibodies typically comprise six CDRs: three in the VH (H1, H2, H3), and three in the VL (L1, L2, L3).

As used herein, the term "sequence identity" refers to the percentage sequence identities that are determined with antibody sequences maximally aligned by the Kabat numbering convention. After alignment, if a subject antibody region (e.g., the entire mature variable region of a heavy or light chain) is being compared with the same region of a reference antibody, the percentage sequence identity between the subject and reference antibody regions is the number of positions occupied by the same amino acid

in both the subject and reference antibody region divided by the total number of aligned positions of the two regions, with gaps not counted, multiplied by 100 to convert to percentage. In some examples, the antigen binding protein comprises a sequence that is at least 60% identical to any one of the sequences disclosed herein. For example, the antigen binding protein may comprise a sequence that is at least about 60%, at least about 61%, at least about 62%, at least about 63%, at least about 64%, at least about 65%, at least about 66%, at least about 67%, at least about 68%, at least about 69%, at least about 70%, at least about 71%, at least about 72%, at least about 73%, at least about 74%, at least about 75%, at least about 76%, at least about 77%, at least about 78%, at least about 79%, at least about 80%, at least about 81%, at least about 82%, at least about 83%, at least about 84%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% identical to any of the sequences disclosed herein (e.g. about 60%, or about 61%, or about 62%, or about 63%, or about 64%, or about 65%, or about 66%, or about 67%, or about 68%, or about 69%, or about 70%, or about 71%, or about 72%, or about 73%, or about 74%, or about 75%, or about 76%, or about 77%, or about 78%, or about 79%, or about 80%, or about 81%, or about 82%, or about 83%, or about 84%, or about 85%, or about 86%, or about 87%, or about 88%, or about 89%, or about 90%, or about 91%, or about 92%, or about 93%, or about 94%, or about 95%, or about 96%, or about 97%, or about 98%, about 99% or about 100% sequence identity to any one of the sequences disclosed herein. In some examples, the antigen binding protein comprises a sequence or an amino acid region or is encoded by a nucleotide region that differs by about one, about two, about three, about four, about five, about six, about seven, about eight, about nine, about ten or more amino acids or nucleobase with the sequence as disclosed herein. In some examples, the antigen binding protein comprises an amino acid sequence having one or more amino acid mutations with respect to any one of the sequences disclosed herein. In some examples, the antigen binding protein comprises an amino acid sequence having one, or two, or three, or four, or five, or six, or seen, or eight, or nine, or ten, or fifteen, or twenty amino acid mutations with respect to any one of the sequences disclosed herein. In some examples, the one or more amino acid mutations may be independently selected from substitutions, insertions, deletions, and truncations.

In some examples, the amino acid mutations are amino acid substitutions, and may include conservative and/or non-conservative substitutions.

"Conservative substitutions" may be made, for instance, on the basis of similarity in polarity, charge, size, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the amino acid residues involved. The 20 naturally occurring amino acids can be grouped into the following six standard amino acid groups: (1) hydrophobic: Met, Ala, Val, Leu, Ile; (2) neutral hydrophilic: Cys, Ser, Thr; Asn, Gln; (3) acidic: Asp, Glu; (4) basic: His, Lys, Arg; (5) residues that influence chain orientation: Gly, Pro; and (6) aromatic: Trp, Tyr, Phe.

As used herein, "conservative substitutions" are defined as exchanges of an amino acid by another amino acid listed within the same group of the six standard amino acid groups shown above. For example, the exchange of Asp by Glu retains one negative charge in the so modified polypeptide. In addition, glycine and proline may be substituted for one another based on their ability to disrupt α-helices.

As used herein, "non-conservative substitutions" are defined as exchanges of an amino acid by another amino acid listed in a different group of the six standard amino acid groups (1) to (6) shown above.

In some examples, the substitutions may also include non-classical amino acids. Illustrative non-classical amino acids include, but are not limited to, selenocysteine, pyrrolysine, N-formylmethionine β-alanine, GABA and δ-Aminolevulinic acid, 4-aminobenzoic acid (PABA), D-isomers of the common amino acids, 2,4-diaminobutyric acid, α-amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, γ-Abu, ε-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, ornithine, norleucine, norvaline, hydroxyproline, sarcosme, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, β-alanine, fluoro-amino acids, designer amino acids such as β methyl amino acids, C α-methyl amino acids, N α-methyl amino acids, and amino acid analogs in general.

In some examples, the amino acid mutation may be in the CDRs of the antigen binding protein (e.g., the CDR1, CDR2 or CDR3 regions). In another example, amino acid alteration may be in the framework regions (FRs) of the antigen binding protein (e.g., the FR1, FR2, FR3, or FR4 regions).

In some examples, the mutations do not substantially reduce the antigen binding protein's capability to specifically bind to a target. In some examples, the mutations do not substantially reduce the antigen binding protein's capability to specifically bind to a target and without functionally modulating (e.g., partially or fully neutralizing) the target.

Modification of the amino acid sequences may be achieved using any known technique in the art e.g., site-directed mutagenesis or PCR based mutagenesis.

The term "polynucleotide" refers to a linear polymer whose molecule is composed of many nucleotide units, constituting a section of a nucleic acid molecule. Polynucleotides are made up of long chains of nucleotides like deoxyribonucleic acid (DNA) and ribonucleic acid (RNA).

The term "affinity" refers to the strength of all noncovalent interactions between an antibody thereof and the target protein. Unless indicated otherwise, as used herein, the term "binding affinity" refers to intrinsic binding affinity which reflects a 1: 1 interaction between members of a binding pair (e.g., antibody and antigen). The affinity of a molecule for its binding partner can be generally represented by the dissociation constant (KD). Affinity can be measured by common methods known in the art, including those described herein.

The term "K_D" as used herein refers to the constant of dissociation which is obtained from the ratio of K_d to K_a (i.e. K_d / K_a) and is expressed as a molar concentration (M). K_d and K_a refer to the dissociation rate and association rate, respectively, of a particular antigen-antibody interaction. K_D values for antibodies can be determined using methods well established in the art. As used herein, the term "low affinity" refers to K_D of 100 nM or more.

As used herein, the term" moderate affinity" refers to K_D ranging from 10nM to 100nM.

As used herein, the term "high affinity" refers to K_D of 1 to 10nM.

As used herein, the term "very high affinity" refers to K_D of 1nM or less.

The term "EC50," as used herein, refers to the concentration of an antibody or an antigen-binding protein/portion thereof, which induces a response, either in an *in vivo* or an *in vitro* assay, which is 50% of the maximal response (i.e., halfway between the maximal response and the baseline).

The term "multi-specific" or "multi-specific antibody" as employed herein refers to an antibody as described herein which has at least two binding domains, i.e. two or more binding domains, for example two or three binding domains, wherein the at least two binding domains independently bind two different antigens or two different epitopes on the same antigen. Multi-specific antibodies are generally monovalent for each specificity (antigen). Multi-specific antibodies described herein encompass monovalent and multivalent, e.g. bivalent, trivalent, tetravalent multi-specific antibodies.

The term "bispecific" or "bispecific antibody" as employed herein refers to an antibody with two antigen specificities or an antibody that has the ability to simultaneously bind to two target antigens/sites.

As used herein, a "bispecific T cell engager (BiTE)" refers to a class of artificial bispecific monoclonal antibodies that direct a host's immune system, such as the T cells' cytotoxic activity against target cells (such as cancer cells). BiTEs are fusion proteins consisting of two single-chain variable fragments (scFvs) of different antibodies, or amino acid sequences from four different genes, on a single peptide chain of about 55 kDa. One of the scFvs binds to an immune cell (such as a T cell via the CD3 receptor), and the other to target of interest (e.g., a tumour cell via a tumour specific molecule). Like other bispecific antibodies, BiTEs form a link between an immune cell (e.g., a T cell) and a target cell (such as a tumour cell). This causes the immune cell (e.g., T cell) to exert cytotoxic activity on tumour cells. For example, if the immune cell is a T cell, the T cell would exert cytotoxic activity by producing proteins like perforin and granzymes that enter tumour cells and initiate the cell's apoptosis.

In some examples, BiTE may refer to the BiTE® immune-oncology platform that was developed by Amgen® Oncology. In some examples, BiTE may also refer to the bispecific T-cell engager platform as known in the art that refers to recombinant proteins that simultaneously bind two different antigens and have the ability to engage an immune cell (such as T cell).

As used herein, the term "nanobody" refers to a single domain antibody (sdAb), with an antibody fragment consisting of a single monomeric variable antibody domain. In some examples, the bispecific T cell engager (BiTE) is a nanobody with a heavy chain only (VHH). As used herein, a "nanobody with a heavy chain only" refers to nanobody-based heavy chain antibody. A heavy-chain antibody is an antibody which consists of two heavy chains and lacks the two light chains usually found in antibodies.

As described herein, a "vector" is any molecule or composition that has the ability to carry a nucleic acid sequence into a suitable host cell where e.g., synthesis of the encoded polypeptide can take place. Typically, and preferably, a vector is a nucleic acid that has been engineered, using recombinant DNA techniques that are known in the art, to incorporate a desired nucleic acid sequence (e.g., a nucleic acid of the present disclosure). Expression vectors typically contain one or more of the following components (if they are not already provided by the nucleic acid molecules): a promoter, one or more enhancer sequences, an origin of replication, a transcriptional termination sequence, a complete intron sequence containing a donor and acceptor splice site, a

leader sequence for secretion, a ribosome binding site, a polyadenylation sequence, a polylinker region for inserting the nucleic acid encoding the polypeptide to be expressed, and a selectable marker element.

Vectors are typically selected to be functional in the host cell in which the vector will be used (the vector is compatible with the host cell machinery such that amplification of the gene and/or expression of the gene can occur. The vector as described herein may be an expression vector and/or a cloning vector.

The term "host cell," as used herein, is intended to refer to a cell into which an expression vector has been introduced. It should be understood that such terms are intended to refer not only to the particular subject cell but to the progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term "host cell" as used herein.

The terms "treating", "treat" and "therapy," and synonyms thereof refer to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) a medical condition, which includes but is not limited to diseases, symptoms and disorders. A medical condition also includes a body's response to a disease or disorder, e.g. inflammation. Those in need of such treatment include those already with a medical condition as well as those prone to getting the medical condition or those in whom a medical condition is to be prevented.

The term "subject" as used herein includes patients and non-patients. The term "patient" refers to individuals suffering or are likely to suffer from a medical condition, while "non-patients" refer to individuals not suffering and are likely to not suffer from the medical condition. "Non-patients" include healthy individuals, non-diseased individuals and/or an individual free from the medical condition. The term "subject" includes humans and animals. Animals may include, but is not limited to, mammals (for example non-human primates, canine, murine and the like), and the like. "Murine" refers to any mammal from the family Muridae and / or Leporidae, such as mouse, rat, rabbit, and the like.

The term "preventing" and/or "reducing the severity of symptoms" as used herein refers to process of delaying the onset, reducing the severity of symptoms, reducing and/or preventing weight loss, preventing death, inhibiting deterioration, inhibiting further deterioration, and/or ameliorating at least one sign or symptom of a disease.

The term "and/or", e.g., "X and/or Y" is understood to mean either "X and Y" or "X or Y" and should be taken to provide explicit support for both meanings or for either meaning.

Further, in the description herein, the word "substantially" whenever used is understood to include, but not restricted to, "entirely" or "completely" and the like. In addition, terms such as "comprising", "comprise", and the like whenever used, are intended to be non-restricting descriptive language in that they broadly include elements/components recited after such terms, in addition to other components not explicitly recited. For example, when "comprising" is used, reference to a "one" feature is also intended to be a reference to "at least one" of that feature. Terms such as "consisting", "consist", and the like, may in the appropriate context, be considered as a subset of terms such as "comprising", "comprise", and the like. Therefore, in embodiments disclosed herein using the terms such as "comprising", "comprise", and the like, it will be appreciated that these embodiments provide teaching for corresponding embodiments using terms such as "consisting", "consist", and the like. Further, terms such as "about", "approximately" and the like whenever used, typically means a reasonable variation, for example a variation of +/- 5% of the disclosed value, or a variance of 4% of the disclosed value, or a variance of 3% of the disclosed value, a variance of 2% of the disclosed value or a variance of 1% of the disclosed value.

Furthermore, in the description herein, certain values may be disclosed in a range. The values showing the end points of a range are intended to illustrate a preferred range. Whenever a range has been described, it is intended that the range covers and teaches all possible sub-ranges as well as individual numerical values within that range. That is, the end points of a range should not be interpreted as inflexible limitations. For example, a description of a range of 1% to 5% is intended to have specifically disclosed sub-ranges 1% to 2%, 1% to 3%, 1% to 4%, 2% to 3% etc., as well as individually, values within that range such as 1%, 2%, 3%, 4% and 5%. It is to be appreciated that the individual numerical values within the range also include integers, fractions and decimals. Furthermore, whenever a range has been described, it is also intended that the range covers and teaches values of up to 2 additional decimal places or significant figures (where appropriate) from the shown numerical end points. For example, a description of a range of 1% to 5% is intended to have specifically disclosed the ranges 1.00% to 5.00% and also 1.0% to 5.0% and all their intermediate values (such as 1.01%, 1.02% ... 4.98%, 4.99%, 5.00% and 1.1%, 1.2% ... 4.8%, 4.9%, 5.0% etc.,) spanning the

ranges. The intention of the above specific disclosure is applicable to any depth/breadth of a range.

"At least 95% identical" as employed herein is intended to refer to an amino acid sequence which over its full length is 95% identical or more to a reference sequence, such as 96, 97, 98 or 99% identical. Software programmes can be employed to calculate percentage identity.

Additionally, when describing some embodiments, the disclosure may have disclosed a method and/or process as a particular sequence of steps. However, unless otherwise required, it will be appreciated that the method or process should not be limited to the particular sequence of steps disclosed. Other sequences of steps may be possible. The particular order of the steps disclosed herein should not be construed as undue limitations. Unless otherwise required, a method and/or process disclosed herein should not be limited to the steps being carried out in the order written. The sequence of steps may be varied and still remain within the scope of the disclosure.

Furthermore, it will be appreciated that while the present disclosure provides embodiments having one or more of the features/characteristics discussed herein, one or more of these features/characteristics may also be disclaimed in other alternative embodiments and the present disclosure provides support for such disclaimers and these associated alternative embodiments.

DESCRIPTION OF EMBODIMENTS

Exemplary, non-limiting embodiments of cells that expresses chimeric antigen receptor targeting GPC3 (or GE3) and/or CD19, and a multi-specific antigen binding protein, variant or binding fragment that binds to EpCAM and an immune cell marker. Also disclosed is an engineered cell expressing the polynucleotide as described herein. Also disclosed is an engineered cell expressing (a) a chimeric antigen receptor targeting a first antigen, and (b) a multi-specific antigen binding protein that binds to EpCAM (epithelial cell adhesion molecule), optionally the first target antigen in (a) is an antigen associated with a disease. Without wishing to be bound by any theory, engineering an cell in order for the EpCAM BiTE to be secreted at the target site minimizes toxicity, reduces side effects, and/or reduces non-specific cytotoxicity.

For example, the target antigen is a molecule associated with a disease. For example, the molecule may be an extracellular molecule, an intracellular molecule,

and/or a transmembrane molecule. In some examples, the molecule may be a polypeptide, a polynucleotide, a carbohydrate, and the like.

In some examples, the chimeric antigen receptor targets a diseased cell. In some examples, the diseased cells may include but are not limited to a cell from a proliferative disease (such as tumour / cancer, an inflammatory disease), and the like. In some examples, the diseased cell may be a cell from cancer / tumour, optionally the cancer cell is a solid tumour.

In some examples, the solid tumour may include tumours with epithelial origins such as carcinoma. In some examples, carcinoma may include but is not limited to liver cancer (such as hepatocellular carcinoma), lung cancer (such as lung squamous cell carcinoma), stomach cancer (such as gastric adenocarcinoma), breast cancer, skin cancer (such as melanoma), ovary cancer (such as ovarian clear cell carcinoma), kidney cancer, pancreas cancer, head and neck cancer, prostate gland cancer, esophagus cancer, bladder cancer, colon cancer, childhood cancers (such as hepatoblastomas, nephroblastomas, yolk sac tumours, and the like), and the like.

In some examples, the tumour cell may include benign, premalignant, malignant tumours, and the like. In some examples, the tumour cell may include a stem cell, a progenitor cell, and the like.

In some examples, the tumour cell may include a human xenograft in an animal model. In some examples, the human xenograft may include Hep3B, HepG2, and the like. In some examples, the human xenograft is a Hep3B xenograft and / or a HepG2 xenograft in a mice model.

In some examples, the tumour cell may include tumour cell lines such as but is not limited to gastric adenocarcinoma cell line (such as AGS), breast cancer cell line (such as MDA-MB468), liver adenocarcinoma cell line (such as Hep3B ($GPC3^{High}EpCAM^{High}$), HepG2 ($GPC3^{High}EpCAM^{High}$), HT-29 ($GPC3^{low}EpCAM^{High}$), HeyA8 ($GPC3^{-ve}EpCAM^{-ve}$), Huh7 ($GPC3^{High}EpCAM^{High}$), and the like.

In some examples, the target antigen is an epithelial marker, which may include but is not limited to receptor tyrosine-protein kinase erbB-2 (HER2), glypican 3 (GPC3), Claudin 18.2, receptor tyrosine kinase like orphan receptor 1 (ROR1), delta like canonical notch ligand 3 (DLL3), carcinoembryonic antigen (CEA), mucin 1 (MUC1), mucin 16 (MUC16), CEA Cell Adhesion Molecule 7 (CEACAM7), prominin-1 (CD133), cluster of differentiation 147 (CD147), prostate stem cell antigen (PSCA), prostate-specific membrane antigen (PSMA), mesothelin (MSLN), mesenchymal-epithelial transition factor (c-Met),folate receptor alpha (FRα), and the like.

In some examples, the chimeric antigen receptor binds to GPC3.

Without wishing to be bound to theory, GPC3 is a well-known oncofoetal protein that can be used as a highly tumour-specific target. It is widely expressed during embryonic development but strictly suppressed in most adult tissues. Elevated GPC3 expression has been reported in a wide variety of tumour types such as liver, lung, gastric, ovarian, oesophageal and many others. At the same time, EpCAM is widely expressed on almost all carcinomas but is also expressed at low levels on normal epithelia. Anti-EpCAM CAR T cells have been demonstrated to be highly toxic to normal tissues, with anti-EpCAM BiTE (such as Solitomab and Catumaxomab) showing dose-limiting toxicities and failed to achieve FDA approval.

Hence, the inventors of the present disclosure engineered CAR-T cells to produce a second tumour-targeting molecule, namely the anti-EpCAM bispecific T cell engager (BiTEs), via secretion. EpCAM is a well-known tumour associated antigen that is frequently over-expressed in almost all solid tumours originating from the epithelium, however, neither CAR T cells nor systemically administered BiTEs targeting EpCAM has ever achieved success due to its high systematic toxicity as EpCAM is also widely expressed on normal epithelia, albeit at low levels. Secretion of anti-EpCAM BiTE by CAR-T cells localizes the anti-EpCAM BiTE at or near the tumour site, which not only reduces tumour escape but also greatly enhances safety. As anti-GPC3 CAR-T cells have previously shown clinically proven efficacy and safety profiles, the present disclosure applied GPC3 targeting CAR-T cells as the carriers of anti-EpCAM BiTEs. The present disclosure demonstrated that GPC3 targeting CAR T cells secreting anti-EpCAM BiTE (named as "GE CAR-BiTE T") exerted superior efficacy in eradicating human xenograft tumours originating from hepatocellular carcinomas, both *in vitro* and *in vivo*.

Therefore, in one aspect, there is provided an cell expressing

(a) a chimeric antigen receptor targeting GPC3 and/or CD19, and

(b) a multi-specific antigen binding protein, variant or binding fragment thereof that binds one or more target, comprising a first antigen binding protein, variant or binding fragment thereof that binds to EpCAM (epithelial cell adhesion molecule) and a second antigen binding protein, variant or binding fragment thereof that binds to an immune cell marker,

wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region and/or a light chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising: (2C4, hu2C4, 1A5, 1B8, 2B7, and 2D10)

a CDR-H1 comprising:

- GSIFSGND (**SEQ ID NO: 25** – 2C4, hu2C4, 1A5, 2B7, and 2D10), or
- GSSERFTS (**SEQ ID NO: 29** – 1B8)

a CDR-H2 comprising:

- ITSGGST (**SEQ ID NO: 26** – 2C4, hu2C4, 1A5, 2B7, and 2D10), or
- ITNGGST (**SEQ ID NO: 30** – 1B8) and

a CDR-H3 comprising:

- TNGRWSGDTYYAHH (**SEQ ID NO: 27** – 2C4, hu2C4, 1A5, 2C4, and 2D10),
- MAGTS (**SEQ ID NO: 31** – 1B8), or
- TNGRWSGDTYYAHL (**SEQ ID NO: 33** – 2B7)

(ii) a heavy chain variable region comprising: (1B6, 1C1, 1C11, 1D4 and 1H6)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2)**, and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**)

(iii) a heavy chain variable region comprising: (1E4)

a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5)**,

a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6)**, and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7)**

(iv) a light chain variable region comprising: (1B6, 1C1, 1C11, 1D4, 1E4 and 1H6)

a CDR-L1 comprising:

- QSLLHSNGYNY (**SEQ ID NO: 9** - 1B6, 1C1, 1C11 and 1H6),
- QSLLHSNRYNY (**SEQ ID NO: 17** - 1D4), or
- QSISDF (**SEQ ID NO: 19** – 1E4)

a CDR-L2 comprising:

- LGS (**SEQ ID NO: 10** -1B6, 1C1, 1C11, 1D4 and 1H6), or
- AAS (**SEQ ID NO: 20** - 1E4), and

a CDR-L3 comprising:

- MQALQTPYT (**SEQ ID NO: 11** - 1B6, 1C1 and 1D4),
- MQGLQSPWT (**SEQ ID NO: 15** - 1C11),
- QQSYIMPDT (**SEQ ID NO: 21** – 1E4), or
- MQGLQTPYT (**SEQ ID NO: 23** – 1H6); and

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region comprising: (2C4, hu2C4, 1A5, 1B8, 2B7, and 2D10)

a CDR-H1 comprising:

- GSIFSGND (**SEQ ID NO: 25** - 2C4, hu2C4, 1A5, 2B7, and 2D10) or
- GSSERFTS (**SEQ ID NO: 29** – 1B8)

a CDR-H2 comprising:

- ITSGGST (**SEQ ID NO: 26** – 2C4, hu2C4, 1A5, 2B7, and 2D10), or
- ITNGGST (**SEQ ID NO: 30** – 1B8); and

a CDR-H3 comprising:

- TNGRWSGDTYYAHH (**SEQ ID NO: 27** – 2C4, hu2C4, 1A5, 2D10),
- MAGTS (**SEQ ID NO: 31** – 1B8), or
- TNGRWSGDTYYAHL (**SEQ ID NO: 33** – 2B7)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising: (2C4-VHH, hu2C4-VHH, 1A5-VHH, and 2D10-VHH and)

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**),

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**), and

a CDR-H3 comprising TNGRWSGDTYYAHH (**SEQ ID NO: 27**)

(ii) a heavy chain variable region comprising: (1B8-VHH)

a CDR-H1 comprising GSSERFTS (**SEQ ID NO: 29**),

a CDR-H2 comprising ITNGGST (**SEQ ID NO: 30**), and

a CDR-H3 comprising MAGTS (**SEQ ID NO: 31**); and

(iii) a heavy chain variable region comprising: (2B7-VHH)

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**),

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**), and

a CDR-H3 comprising TNGRWSGDTYYAHL (**SEQ ID NO: 33**)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region and/or a light chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising: (2C4, hu2C4, 1A5, and 2D10)

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**),

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**), and

a CDR-H3 comprising TNGRWSGDTYYAHH (**SEQ ID NO: 27**)

(ii) a heavy chain variable region comprising: (1B6, 1C1, 1C11, 1D4, 1H6)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**)

(iii) a heavy chain variable region comprising: (1E4)

a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**),

a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**), and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**)

(iv) a light chain variable region comprising: (1B6 and 1C1)

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**)

(v) a light chain variable region comprising: (1C11)

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**)

(vi) a light chain variable region comprising: (1D4)

a CDR-L1 comprising QSLLHSNRYNY (**SEQ ID NO: 17**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**)

(vii) a light chain variable region comprising: (1E4)

a CDR-L1 comprising QSISDF (**SEQ ID NO: 19**),

a CDR-L2 comprising AAS (**SEQ ID NO: 20**), and

a CDR-L3 comprising QQSYIMPDT (**SEQ ID NO: 21**)

(viii) a light chain variable region comprising: (1H6)

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**)

(ix) a heavy chain variable region comprising: (1B8-VHH)

a CDR-H1 comprising GSSERFTS (**SEQ ID NO: 29**),

a CDR-H2 comprising ITNGGST (**SEQ ID NO: 30**), and

a CDR-H3 comprising MAGTS (**SEQ ID NO: 31**); and

(x) a heavy chain variable region comprising: (2B7-VHH)

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**),

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**), and

a CDR-H3 comprising TNGRWSGDTYYAHL (**SEQ ID NO: 33**)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain and/or a light chain variable domain selected from the group consisting of:

(i) a heavy chain variable domain comprising QVQLQESGGGLVQAGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE LVA**VITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 35** - 2C4-VHH)

(ii) a heavy chain variable domain comprising QVQLVESGGGLVQAGGSLRLSCAAS**GSIFSGND**MSWYRQAPGKGLE LVA**VITSGGST**YYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYY C**TNGRWSGDTYYAHH**WGQGTL (**SEQ ID NO: 37** - hu2C4-VHH)

(iii) a heavy chain variable domain comprising EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLE WMGG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVY YC**ARSLGGRFRY**WGQGTL (**SEQ ID NO: 4** - 1B6, 1C1, 1C11, 1D4 and 1H6)

(iv) a heavy chain variable domain comprising QVQLQQSGPGLVKPSQTLSLTCAIS**GDSISSNSVA**WNWIRQSPSRGL EWLGR**TYYRSKWYS**DYAISVKGRLDINPDTSKNQFSLQLNSVTPEDT AVYYC**AREVEGSSYDAFDI**WGQGTM (**SEQ ID NO: 8** - 1E4),

(v) a light chain variable domain comprising DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQ SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**M QALQTPYT**FGQGTK (**SEQ ID NO: 12** - 1B6 and 1C1)

(vi) a light chain variable domain comprising EIVLTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQS PQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQ GLQSPWT**FGQGTK (**SEQ ID NO: 16** - 1C11)

(vii) a light chain variable domain comprising DVVMTQSPLSLPVTPGESASISCRSS**QSLLHSNRYNY**LDWYLQKPGQ SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**M QALQTPYT**FGQGTK (**SEQ ID NO: 18** - 1D4)

(viii) a light chain variable domain comprising DIQLTQSPSSLSASVGDRVTITCRAS**QSISDF**LNWYQQKPGKAPKLLIY **AAS**SLQTGVPSRFGGSGSGTEFTLTISSLQPEDLGTYYC**QQSYIMPDT** FGQGTK (**SEQ ID NO: 22** - 1E4)

(ix) a light chain variable domain comprising DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQ SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYC**M QGLQTPYT**FGQGTK (**SEQ ID NO: 24** - 1H6)

(x) a heavy chain variable domain comprising QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE LVAV**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 28-** 1A5-VHH)

(xi) a heavy chain variable domain comprising QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERE LVAF**ITNGGST**RYTDPVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYY C**MAGTS**WGQGTQ (**SEQ ID NO: 32** - 1B8-VHH)

(xii) a heavy chain variable domain comprising QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE

LVA**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHL**WGQGTQ (**SEQ ID NO: 34** - 2B7-VHH)

(xiii) a heavy chain variable domain comprising
QVQLQESGGGLVQAGDSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVA**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 36** - 2D10-VHH); and

or fragment or variant or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a single domain heavy chain variable domain having a sequence:

(i) QVQLQESGGGLVQAGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVA**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 35** - clone 2C4-VHH), or

(ii) QVQLVESGGGLVQAGGSLRLSCAAS**GSIFSGND**MSWYRQAPGKGLE
LVA**ITSGGST**YYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYY
C**TNGRWSGDTYYAHH**WGQGTL (**SEQ ID NO: 37** - hu2C4-VHH), or

(iii) QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVA**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 28** - clone 1A5-VHH), or

(iv) QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERE
LVAF**ITNGGST**RYTDPVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYY
C**MAGTS**WGQGTQ (**SEQ ID NO: 32** - clone 1B8-VHH), or

(v) QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVA**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHL**WGQGTQ (**SEQ ID NO: 34** - clone 2B7-VHH), or

(vi) QVQLQESGGGLVQAGDSLRLSCADS**GSIFSGND**MAWYRRAPGVERE
LVA**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY
C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 36** - clone 2D10-VHH),

or fragment or variant or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising: (1B6, 1C1, 1C11, 1D4 and 1H6)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**)

(ii) a heavy chain variable region comprising: (1E4)

a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**),

a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**), and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**); and/or

comprises a light chain variable region comprising: (1B6, 1C1, 1C11, 1D4, 1E4 and 1H6)

a CDR-L1 comprising:

- QSLLHSNGYNY (**SEQ ID NO: 9** - 1B6, 1C1, 1C11 and 1H6),
- QSLLHSNRYNY (**SEQ ID NO: 17** - 1D4), or
- QSISDF (**SEQ ID NO: 19** – 1E4)

a CDR-L2 comprising:

- LGS (**SEQ ID NO: 10** -1B6, 1C1, 1C11, 1D4 and 1H6), or
- AAS (**SEQ ID NO: 20** - 1E4), and

a CDR-L3 comprising:

- MQALQTPYT (**SEQ ID NO: 11** - 1B6, 1C1 and 1D4),
- MQGLQSPWT (**SEQ ID NO: 15** - 1C11),
- QQSYIMPDT (**SEQ ID NO: 21** – 1E4), or
- MQGLQTPYT (**SEQ ID NO: 23** – 1H6)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region and/or a light chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising: (1B6 and 1C1)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**); and/or

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**)

(ii) a heavy chain variable region comprising: (1C11)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**); and/or

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**)

(iii) a heavy chain variable region comprising: (1D4)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**); and/or

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNRYNY (**SEQ ID NO: 17**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**)

(iv) a heavy chain variable region comprising: (1H6)

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**),

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**), and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**); and/or

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**),

a CDR-L2 comprising LGS (**SEQ ID NO: 10**), and

a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**); and

(v) a heavy chain variable region comprising: (1E4)

a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**),

a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**), and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**); and/or

a light chain variable region comprising: (1E4)

a CDR-L1 comprising QSISDF (**SEQ ID NO: 19**),

a CDR-L2 comprising AAS (**SEQ ID NO: 20**), and

a CDR-L3 comprising QQSYIMPDT (**SEQ ID NO: 21**)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain and/or a light chain variable domain selected from the group consisting of:

(i) a heavy chain variable domain comprising: (1B6 and 1C1)
EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLEWM
GG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYC**ARS**
LGGRFRYWGQGTL (**SEQ ID NO: 4**), and/or
a light chain variable domain comprising
DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQSP
QLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQALQT**
PYTFGQGTK (**SEQ ID NO: 12**)

(ii) a heavy chain variable domain comprising: (1C11)
EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLEWM
GG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYC**ARS**
LGGRFRYWGQGTL (**SEQ ID NO: 4**), and/or
a light chain variable domain comprising
EIVLTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQSPQ
LLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQGLQSP**
WTFGQGTK (**SEQ ID NO: 16**)

(iii) a heavy chain variable domain comprising: (1D4)
EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLEWM
GG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYC**ARS**
LGGRFRYWGQGTL (**SEQ ID NO: 4**), and/or
a light chain variable domain comprising
DVVMTQSPLSLPVTPGESASISCRSS**QSLLHSNRYNY**LDWYLQKPGQSP
QLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQALQT**
PYTFGQGTK (**SEQ ID NO: 18**)

(iv) a heavy chain variable domain comprising: (1E4)
QVQLQQSGPGLVKPSQTLSLTCAIS**GDSISSNSVA**WNWIRQSPSRGLEWL
GR**TYYRSKWYS**DYAISVKGRLDINPDTSKNQFSLQLNSVTPEDTAVYYC**A**
REVEGSSYDAFDIWGQGTM (**SEQ ID NO: 8**), and/or
a light chain variable domain comprising

DIQLTQSPSSLSASVGDRVTITCRAS**QSISDF**LNWYQQKPGKAPKLLIY**AA**
SSLQTGVPSRFGGSGSGTEFTLTISSLQPEDLGTYYC**QQSYIMPDT**FGQG
TK (**SEQ ID NO: 22**), or

(v) a heavy chain variable domain comprising: (1H6)
EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLEWM
GG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYC**ARS**
LGGRFRYWGQGTL (**SEQ ID NO: 4**), and/or
a light chain variable domain
DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQSP
QLLIY**LGS**NRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYC**MQGLQT**
PYTFGQGTK (**SEQ ID NO: 24**), or

or fragment or variant or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a light chain constant domain having a sequence:

(i) RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVAEQDSKDSTYSLSSTLTLSKADYEKHKLYACEVTHQGLSSPVTK
SFNRGEC (**SEQ ID NO: 13** - clone 1B6 – light chain constant domain), or

(ii) RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
NSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTK
SFSRGEC (**SEQ ID NO: 14** clone 1C1 – light chain constant domain),

or fragment or variant or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions.

In one embodiment, the antigen binding protein, variant or fragment thereof comprises a heavy chain variable region and/or a light chain variable region encoded by nucleic acid sequences selected from the group consisting of:

(i) a heavy chain variable region comprising: (1B6 and 1C1)
a CDR-H1 comprising GGAGGCACCTTCAGCAGCTATGCT (**SEQ ID NO: 38**),
a CDR-H2 comprising ATCATCCCTATCTTTGGTACAGCA (**SEQ ID NO: 39**), and

a CDR-H3 comprising GCGAGATCGTTGGGTGGGAGATTTCGCTAC (**SEQ ID NO: 40**); and/or

a light chain variable region comprising:

a CDR-L1 comprising CAGAGCCTCCTGCATAGTAATGGATACAACTAT (**SEQ ID NO: 46**),

a CDR-L2 comprising TTGGGTTCT (**SEQ ID NO: 47**), and

a CDR-L3 comprising ATGCAAGCTCTACAAACTCCGTACACT (**SEQ ID NO: 48**)

(ii) a heavy chain variable region comprising: (1C11)

a CDR-H1 comprising GGAGGCACCTTCAGCAGCTATGCT (**SEQ ID NO: 38**),

a CDR-H2 comprising ATCATCCCTATCTTTGGTACAGCA (**SEQ ID NO: 39**), and

a CDR-H3 comprising GCGAGATCGTTGGGTGGGAGATTTCGCTAC (**SEQ ID NO: 40**); and/or

a light chain variable region comprising:

a CDR-L1 comprising CAGAGCCTCCTGCATAGTAATGGATACAACTAT (**SEQ ID NO: 46**),

a CDR-L2 comprising TTGGGTTCT (**SEQ ID NO: 47**), and

a CDR-L3 comprising ATGCAAGGTCTACAAAGTCCCTGGACG (**SEQ ID NO: 52**)

(iii) a heavy chain variable region comprising: (1D4)

a CDR-H1 comprising GGAGGCACCTTCAGCAGCTATGCT (**SEQ ID NO: 38**),

a CDR-H2 comprising ATCATCCCTATCTTTGGTACAGCA (**SEQ ID NO: 39**), and

a CDR-H3 comprising GCGAGATCGTTGGGTGGGAGATTTCGCTAC (**SEQ ID NO: 40**); and/or

a light chain variable region comprising:

a CDR-L1 comprising CAGAGCCTCCTGCATAGTAATAGATACAACTAT (**SEQ ID NO: 54**),

a CDR-L2 comprising TTGGGTTCT (**SEQ ID NO: 47**), and

a CDR-L3 comprising ATGCAAGCTCTACAAACTCCGTACACT (**SEQ ID NO: 48**)

(iv) a heavy chain variable region comprising: (1H6)

a CDR-H1 comprising GGAGGCACCTTCAGCAGCTATGCT (**SEQ ID NO: 38**),

a CDR-H2 comprising ATCATCCCTATCTTTGGTACAGCA (**SEQ ID NO: 39**), and

a CDR-H3 comprising GCGAGATCGTTGGGTGGGAGATTTCGCTAC (**SEQ ID NO: 40**); and/or

a light chain variable region comprising:

a CDR-L1 comprising CAGAGCCTCCTGCATAGTAATGGATACAACTAT (**SEQ ID NO: 46**),

a CDR-L2 comprising TTGGGTTCT (**SEQ ID NO: 47**), and

a CDR-L3 comprising ATGCAAGGTCTACAGACTCCGTACACT (**SEQ ID NO: 60**); and

(v) a heavy chain variable region comprising: (1E4)

a CDR-H1 comprising GGGGACAGTATCTCTAGTAACAGTGTTGCT (**SEQ ID NO: 42**),

a CDR-H2 comprising ACATACTACAGGTCCAAGTGGTACAGT (**SEQ ID NO: 43**), and

a CDR-H3 comprising GCAAGAGAAGTTGAGGGCAGCAGCTATGATGCTTTTGATATC (**SEQ ID NO: 44**); and/or

a light chain variable region comprising: (1E4)

a CDR-L1 comprising CAGAGTATTAGCGACTTT (**SEQ ID NO: 56**),

a CDR-L2 comprising GCTGCATCG (**SEQ ID NO: 57**), and

a CDR-L3 comprising TTACATTATGCCCGACACT (**SEQ ID NO: 58**)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In one embodiment, the antigen binding protein, variant or fragment thereof comprises a heavy chain variable region encoded by nucleic acid sequences selected from the group consisting of:

(i) a heavy chain variable region comprising: (2C4-VHH, 1A5-VHH, 2D10-VHH and hu2C4-VHH)

a CDR-H1 comprising GGAAGCATCTTCAGTGGCAATGAC (**SEQ ID NO: 62**),

a CDR-H2 comprising ATTACTAGCGGTGGTAGTACA (**SEQ ID NO: 63**), and

a CDR-H3 comprising ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC (**SEQ ID NO: 64**)

(ii) a heavy chain variable region comprising: (1B8-VHH)

a CDR-H1 comprising GGAAGCTCCGAAAGATTCACATCA (**SEQ ID NO: 66**),

a CDR-H2 comprising ATTACTAATGGTGGTAGCACA (**SEQ ID NO: 67**), and

a CDR-H3 comprising ATGGCGGGTACGTCC (**SEQ ID NO: 68**); and

(iii) a heavy chain variable region comprising: (2B7-VHH)

a CDR-H1 comprising GGAAGCATCTTCAGTGGCAATGAC (**SEQ ID NO: 62**),

a CDR-H2 comprising ATTACTAGCGGTGGTAGTACA (**SEQ ID NO: 63**), and

a CDR-H3 comprising ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCTC (**SEQ ID NO: 70**)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain encoded by a nucleotide sequence comprising:

(i) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG GGGCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG TGGCAATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT ACTATGCCCATCAC**TGGGGCCAGGGGACCCAG (**SEQ ID NO: 72** - clone 2C4), or

(ii) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGG GGGCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAG TGGCAATGAC**ATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACT CGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**TACTATGC

AGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCaAG
AAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGG
CCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACT**
ATGCCCATCACTGGGGCCAGGGGACCCTG (**SEQ ID NO: 74** - clone
hu2C4-VHH); or

(iii) GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO:** 41 - for clones 1B6,
1C1, 1C11, 1D4, 1H6), or

(iv) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC**
TAGTAACAGTGTTGCTTGGAACTGGATCAGGCAGTCCCCATCGAG
AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
ACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC
AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG
(**SEQ ID NO: 45** - clone 1E4); or

(v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG
CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC
GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA**
CTATGCCCATCACTGGGGCCAGGGGACCCAG (**SEQ ID NO: 65** -
clone 1A5), or

(vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
GGGGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAG**
ATTCACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGC

GCGAGTTGGTCGCATTT**ATTACTAATGGTGGTAGCACA**AGATATAC
AGACCCCGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAA
GAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACAC
GGCCGTCTATTATTGT**ATGGCGGGTACGTCC**TGGGGCCAGGGGAC
CCAG (**SEQ ID NO: 69** - clone 1B8), or

(vii)CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG
CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCATCTCTGGGGCCAGGGGACCCAG (**SEQ ID NO: 71** - clone 2B7), or

(viii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGG
CTGGGGACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCT**
TCAGTGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTG
GAGCGCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CAC
TATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAAT
GCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAG
GACACGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGA**
TACTTACTATGCCCATCACTGGGGCCAGGGGACCCAG (**SEQ ID NO: 73** - clone 2D10), and/or

the light chain variable domain is encoded by a nucleotide sequence comprising:

(vi) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 49** - clone 1B6 and 1C1), or

(vii)GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG

GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGTACAGATTT
TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGGTCTACAAAGTCCCTGGACG**TTCGGCCAAGGG
ACCAAG (**SEQ ID NO: 53** - clone 1C11), or

(viii) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCAC
CCCTGGAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCT**
CCTGCATAGTAATAGATACAACTATTTGGATTGGTACCTGCAGAAG
CCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGG
CCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAG
ATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGT
TTATTACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAG
GGGACCAAG (**SEQ ID NO: 55** - clone 1D4), or

(ix) GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT
CCTGATCTAT**GCTGCATCG**AGTTTACAAACTGGGGTCCCCTCAAGA
TTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCACCATAAGCA
GTCTACAACCTGAAGATTTGGGAACTTATTACTGT**CAACAGAGTTA**
CATTATGCCCGACACTTTTGGCCAGGGGACGAAA (**SEQ ID NO: 59** - clone 1E4), or

(x) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
TTACTGC**ATGCAAGGTCTACAGACTCCGTACACT**TTTGGCCAGGG
GACCAAG (**SEQ ID NO: 61** - clone 1H6),

or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain encoded by a nucleotide sequence comprising:

(i) GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG CAGCTATGCT**ATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG CTAC**TGGGGCCAGGGAACCCTG (**SEQ ID NO:** 41 - for clones 1B6, 1C1, 1C11, 1D4, 1H6), or

(ii) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC TAGTAACAGTGTTGCT**TGGAACTGGATCAGGCAGTCCCCATCGAG AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT ACAGT**GATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC AGCAGCTATGATGCTTTTGATATC**TGGGGCCAAGGGACAATG (**SEQ ID NO: 45** - clone 1E4); or

(iii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG TGGCAATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA CTATGCCCATCAC**TGGGGCCAGGGGACCCAG (**SEQ ID NO: 65** - clone 1A5), or

(iv) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG GGGGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAG ATTCACATCA**GTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGC GCGAGTTGGTCGCATTT**ATTACTAATGGTGGTAGCACA**AGATATAC AGACCCCGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAA GAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACAC GGCCGTCTATTATTGT**ATGGCGGGTACGTCC**TGGGGCCAGGGGAC CCAG (**SEQ ID NO: 69** - clone 1B8), or

(v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG TGGCAATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT ACTATGCCCATCTC**TGGGGCCAGGGGACCCAG (**SEQ ID NO: 71 -** clone 2B7), or

(vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG GGGCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG TGGCAATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT ACTATGCCCATCAC**TGGGGCCAGGGGACCCAG (**SEQ ID NO: 72 -** clone 2C4), or

(vii)CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG GGACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGT GGCAATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCG CGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGC AGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCA GAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACAC GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA CTATGCCCATCAC**TGGGGCCAGGGGACCCAG (**SEQ ID NO: 73 -** clone 2D10), or

(viii) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGG CTGGGGGCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCT TCAGTGGCAATGAC**ATGTCCTGGTACCGCCAGGCTCCAGGGAAG GGACTCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**TAC TATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATT CCaAGAAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGA CACGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATA**

CTTACTATGCCCATCACTGGGGCCAGGGGACCCTG (**SEQ ID NO: 74** - clone hu2C4-VHH); and/or

the light chain variable domain is encoded by a nucleotide sequence comprising:

(xi) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC ATAGTAATGGATACAACTAT**TTGGATTGGTACCTGCAGAAGCCAG GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT TACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGG ACCAAG (**SEQ ID NO: 49** - clone 1B6 and 1C1), or

(xii)GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC ATAGTAATGGATACAACTAT**TTGGATTGGTACCTGCAGAAGCCAG GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGTACAGATTT TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT TACTGC**ATGCAAGGTCTACAAAGTCCCTGGACG**TTCGGCCAAGGG ACCAAG (**SEQ ID NO: 53** - clone 1C11), or

(xiii) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCAC CCCTGGAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCT CCTGCATAGTAATAGATACAACTAT**TTGGATTGGTACCTGCAGAAG CCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGG CCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAG ATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGT TTATTACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAG GGGACCAAG (**SEQ ID NO: 55** - clone 1D4), or

(xiv) GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATC TGTAGGAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTAT TAGCGACTTT**TTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCG AAGCTCCTGATCTAT**GCTGCATCG**AGTTTACAAACTGGGGTCCCCT CAAGATTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCACCAT AAGCAGTCTACAACCTGAAGATTTGGGAACTTATTACTGT**CAACAG AGTTACATTATGCCCGACACT**TTTGGCCAGGGGACGAAA (**SEQ ID NO: 59** - clone 1E4), or

(xv) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCAC CCCTGGAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCT CCTGCATAGTAATGGATACAACTAT**TTGGATTGGTACCTGCAGAAG CCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGG CCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAG ATTTTACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGG TTTATTACTGC**ATGCAAGGTCTACAGACTCCGTACACT**TTTGGCCA GGGGACCAAG (**SEQ ID NO: 61 -** clone 1H6),

or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises heavy chain and/or light chain variable domains encoded by nucleotide sequences selected from the group consisting of:

(i) a heavy chain variable domain encoded by the nucleotide sequence comprising: (1B6 and 1C1)

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG CAGCTATGCT**ATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG CTAC**TGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**), and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC ATAGTAATGGATACAACTAT**TTGGATTGGTACCTGCAGAAGCCAG GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT TACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGG ACCAAG (**SEQ ID NO: 49**)

(ii) a heavy chain variable domain encoded by the nucleotide sequence comprising: (1C11)

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**), and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGTACAGATTT
TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGGTCTACAAAGTCCCTGGACG**TTCGGCCAAGGG
ACCAAG (**SEQ ID NO: 53**)

(iii) a heavy chain variable domain encoded by the nucleotide sequence comprising: (1D4)

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**), and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATAGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG

GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 55**)

(iv) a heavy chain variable domain encoded by the nucleotide sequence comprising: (1H6)

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**), and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
TTACTGC**ATGCAAGGTCTACAGACTCCGTACACT**TTTGGCCAGGG
GACCAAG (**SEQ ID NO: 61**); and

(v) a heavy chain variable domain encoded by the nucleotide sequence comprising: (1E4)

CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC**
TAGTAACAGTGTTGCTTGGAACTGGATCAGGCAGTCCCCATCGAG
AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
ACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC
AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**

AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG (**SEQ ID NO: 45**), and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG ACTTTT**TTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT CCTGATCTAT**GCTGCATCG**AGTTTACAAACTGGGGTCCCCTCAAGA TTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCACCATAAGCA GTCTACAACCTGAAGATTTGGGAACTTATTACTGT**CAACAGAGTTA CATTATGCCCGACACT**TTTGGCCAGGGGACGAAA (**SEQ ID NO: 59**)

or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions.

In one embodiment, the antigen binding protein, variant or fragment thereof comprises comprises a light chain constant domain encoded by the nucleotide sequence comprising

CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCA GTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCC CAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTA ACTCCCAGGAGAGTGTCGCAGAGCAGGACAGCAAGGACAGCACCTACAG CCTCAGCAGCACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAC TCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCTCGCCCGTCACAAAG AGCTTCAACAGGGGAGAGTGT (**SEQ ID NO: 50 - clone 1B6 – light chain constant domain**), or

CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCA GTTGAAATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCC CAGAGAGGCCAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTA ACTCCCAGGAGAGTGTCACAGAGCAGGACAGCAAGGACAGCACCTACAG CCTCAGCAGCACCCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAG TCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCTCGCCCGTCACAAAG AGCTTCAGCAGGGGAGAGTGT (**SEQ ID NO: 51 -** clone 1C1 – light chain constant domain), or

a sequence at least 60% identical thereto and/or having 10-20 nucleic acid substitutions.

In some examples, the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain encoded by the nucleotide sequence comprising:

(i) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGG GCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA ATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC**TGG GGCCAGGGGACCCAG (**SEQ ID NO: 72**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(ii) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGG GCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAGTGGCA ATGAC**ATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACTCGAGTTG GTCGCGGTT**ATTACTAGCGGTGGTAGTACA**TACTATGCAGACTCCGTG AAGGGCCGATTCACCATCTCCAGAGACAATTCCaAGAAcACCcTATATC TGCAAATGAACAGCCTGAGAGCTGAGGACACGGCCGTGTATTACTGC**A CAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC**TGGGGC CAGGGGACCCTG (**SEQ ID NO: 74**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions; or

(iii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAG GGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA ATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT ATCTGCAAACGAACGACTTGAAACCTGAGGACACGGCCGTGTATTACT GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC**TGG GGCCAGGGGACCCAG (**SEQ ID NO: 65**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(iv) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGGG GGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAGATTC**

ACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGCGCGAGTT
GGTCGCATTT**ATTACTAATGGTGGTAGCACA**AGATATACAGACCCCGT
GAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGAACACGGTGT
ATCTGCAAATGAACAGCCTGAAAGCTGAGGACACGGCCGTCTATTATT
GT**ATGGCGGGTACGTCC**TGGGGCCAGGGGACCCAG (**SEQ ID NO: 69**)
or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
identical thereto and/or having 10-20 nucleic acid substitutions, or

(v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAG
GGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG
TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCTC**TGG
GGCCAGGGGACCCAG (**SEQ ID NO: 71**) or a sequence at least 80%, 85%,
90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having
10-20 nucleic acid substitutions, or

(vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGG
ACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG
TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC**TGG
GGCCAGGGGACCCAG (**SEQ ID NO: 73**) or a sequence at least 80%, 85%,
90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having
10-20 nucleic acid substitutions,
GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGTC
CTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAGCAGCT**
ATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCTTGAGTGG
ATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTACGCACAGAAC
TTCCAGGGCAGAGTCACCATGACCGCAGACACCTCCATAAGCACAGC
CTACATGGAGCTGAGCAGCCTGAGATCTGAGGACACGGCCGTGTATTA
CTGT**GCGAGATCGTTGGGTGGGAGATTTCGCTAC**TGGGGCCAGGGAA
CCCTG (**SEQ ID NO:** 41) or a sequence at least 80%, 85%, 90%, 95%, 96%,

97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(vii)CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTCGCA GACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTCTAGTAA CAGTGTTGCT**TGGAACTGGATCAGGCAGTCCCCATCGAGAGGCCTTG AGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGTACAGT**GATTAT GCAATATCTGTGAAAGGTCGATTAGACATCAACCCAGACACATCCAAG AACCAGTTCTCCCTGCAGCTGAACTCTGTGACTCCCGAGGACACGGCT GTGTATTATTGT**GCAAGAGAAGTTGAGGGCAGCAGCTATGATGCTTTT GATATC**TGGGGCCAAGGGACAATG (**SEQ ID NO: 45**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions.

In some examples, the multi-specific antigen binding protein may comprise an amino acid sequence having one or more amino acid mutations with respect to any one of the sequences disclosed herein. In some examples, the antigen binding protein comprises an amino acid sequence having one, or two, or three, or four, or five, or six, or seen, or eight, or nine, or ten, or fifteen, or twenty amino acid mutations with respect to any one of the sequences disclosed herein. In some examples, the one or more amino acid mutations may be independently selected from substitutions, insertions, deletions, and truncations. In some examples, the amino acid mutations are amino acid substitutions, and may include conservative and/or non-conservative substitutions. In some examples, the mutations do not substantially reduce the antigen binding protein's capability to specifically bind to a target. In some examples, the mutations do not substantially reduce the antigen binding protein's capability to specifically bind to a target and without functionally modulating (e.g., partially or fully neutralizing) the target.

In some examples, the multi-specific antigen binding protein may be a bispecific, a tri-specific, a tetra-specific antigen binding protein, and the like. In some examples, the multi-specific antigen binding protein is a bispecific antigen binding protein. In some examples, the multi-specific antigen binding protein, variant or fragment thereof is a bispecific antibody. In some examples, the multi-specific antigen binding protein may be provided as a nanobody. In some examples, the multi-specific antigen binding protein may be provided as a Fc region. In some examples, the multispecific antigen binding protein is an inducible bispecific T cell engager comprising a heavy chain antibody

variable region (i.e. VHH) and/or a single chain variable fragment (scFv). In some embodiments, the bispecific T cell engager (BiTE) is a nanobody with a heavy chain only (VHH).

In some examples, the multi-specific antigen binding protein, variant or fragment thereof is an inducible bispecific immune cell engager. In some examples, the multi-specific antigen binding protein is secreted by the cell, optionally an immune cell. In some examples, the multi-specific antigen binding protein, variant or fragment thereof is a bispecific T cell engager (BiTE).

As used herein, a bispecific T cell engager (BiTE) refers to a class of artificial bispecific monoclonal antibodies that direct a host's immune system, such as the T cells' cytotoxic activity against target cells (such as cancer cells). BiTEs are fusion proteins consisting of two single-chain variable fragments (scFvs) of different antibodies, or amino acid sequences from four different genes, on a single peptide chain of about 55 kDa. One of the scFvs binds to T cells via the CD3 receptor, and the other to a target of interest (e.g. a tumour cell via a tumour specific molecule). Like other bispecific antibodies, BiTEs form a link between T cells and target cells (such as tumour cells). This causes T cells to exert cytotoxic activity on the target cells (such as tumour cells) by producing proteins like perforin and granzymes that enter tumour cells and initiate the cell's apoptosis.

In some examples, BiTE may refer to the BiTE® immune-oncology platform that was developed by Amgen® Oncology. In some examples, BiTE may also refer to the bispecific T-cell engager platform as known in the art that refers to recombinant proteins that simultaneously bind two different antigens and have the ability to engage an immune cell (such as T cell).

In some examples, the bispecific T cell engager (BiTE) binds to two antigens; wherein the first antigen is EpCAM; and wherein the second antigen is an immune cell marker. In some examples, the second antigen binding protein, variant or binding fragment thereof binds to the immune marker selected from the group consisting of CD3, NKG2D, CD4, CD8, CD16, and CD64.

In some examples, wherein the multi-specific antigen binding protein is an inducible bispecific T cell engager comprising a Heavy chain antibody variable region (i.e. VHH) and/or a single chain variable fragment (scFv).

In some examples, the multi-specific antigen binding protein, variant or fragment thereof is an inducible bispecific T cell engager. In some examples, the multi-specific antigen binding protein is secreted by the cell.

In some examples, the bispecific T cell engager (BiTE) binds to two antigens; wherein the first antigen is EpCAM; and wherein the second antigen is an immune cell marker. In some examples, the second antigen binding protein, variant or binding fragment thereof binds to the immune marker selected from the group consisting of CD3, NKG2D, CD4, CD8, CD16, CD64, and the like.

In some examples, wherein the multi-specific antigen binding protein is an inducible bispecific T cell engager comprising a Heavy chain antibody variable region (i.e. VHH) and/or a single chain variable fragment (scFv).

In some examples, the bispecific T cell engager (BiTE) binds to two antigens; wherein the first antigen is an Epithelial Cellular Adhesion Molecule (EpCAM); and wherein the second antigen is CD3. In some examples, the CD3 may include CD3ζ, CD3ε, CD3γ, CD3δ, and the like.

In some examples, the modified/engineered immune cell may be but is not limited to a macrophage, a dendritic cell, a T cell, a B cell, an eosinophil, a basophil, a neutrophil, a mast cell, a natural killer T cell (NKT cell), a natural killer cell (NK cell), a macrophage, a monocyte, and the like. In some examples, the modified/engineered immune cell is a T cell, an NK cell, or a macrophage. In some examples, the modified/engineered immune cell is a T cell.

In some examples, the cell may bind to more than one host cell antigen. Therefore, in some examples, the cell may further bind to one host cell antigen, two host cell antigens, three host cell antigens, four host cell antigens, and the like.

In some examples, the engineered cell binds to a target antigen and secretes a BiTE that targets EpCAM and CD3 (CAR-BITE T). In some examples, the cell i. binds to GPC3 and secretes an inducible bispecific T cell engager that targets EpCAM and CD3 (GE CAR-BiTE T); or ii. binds to CD19 and secretes an inducible bispecific T cell engager that targets EpCAM and CD3 (CD19 CAR-BiTE T). In some examples, the cell binds to GPC3 and secretes a BiTE that targets EpCAM and CD3 (GE CAR-BiTE T). In some examples, the CAR T cell targets GPC3 alone (GE CAR-T).

Without wishing to be bound by theory, while the CAR T cells kill target cells directly, the anti-EpCAM BiTEs exert their cytotoxicity by recruiting the immune cells (such as T cells) nearby. By this approach, bystander T cells are physically directed to the close proximity of tumours and at the same time being activated and help the clearance of tumour cells. By applying local secretion of anti-EpCAM BiTE by CAR T cells, it converts a non-druggable target into a druggable target as it diminishes the on-target, off-tumour toxicity that might be brought by systematic delivery.

As EpCAM is also defined as a cancer stem cell marker expressed on cancer progenitor cells and cancer stem cells, secretion of anti-EpCAM BiTE by CAR T cells will concomitantly contribute to the prevention of cancer relapse and recurrence by eliminating cancer stem cells and progenitor cells.

As noted in the experimental data of the present disclosure, GPC3 targeting CAR T cells secreting anti-EpCAM BiTE (GE CAR BiTE T) exerted superior efficacy in eradicating human xenograft tumours originating from hepatocellular carcinomas, both *in vitro* and *in vivo*.

In terms of *in vitro* data, GE CAR-BiTE T cells killed the Hep3B ($GPC3^{High}$, $EpCAM^{High}$) target cells at a slightly faster rate than the anti-GPC3 CAR alone (GE CAR T)(FIGURE 2C), suggesting that the BiTE has a crucial role in facilitating target-cell killing. Culture supernatant harvested from the GE CAR-BiTE T cells were shown to direct efficient T cells mediated target cell killing (FIGURE 2D).

In terms of *in vivo* data, in Hep3B xenografts, while both the CAR T and GE CAR-BiTE T treatments reduced tumour growth, only the GE CAR-BiTE treated mice achieved complete tumour regression (FIGURE 3C, FIGURE 3D) and similarly exhibited good safety profile without causing weight loss of mice (FIGURE 3B). FIGURE 3F and FIGURE 3G further show the superior efficacy in tumour control and long-term mice survival when using GE CAR-BiTE T cells expanded in culture medium containing IL-7 and IL-15. In HepG2 tumour xenografts which is a more aggressive mouse model, different from the inefficiency of the CAR T alone treated group, tumours from all three GE CAR-BiTE T cell treated mice were successfully cured (FIGURE 4D, FIGURE 4E) and the mice all survived (FIGURE 4G) without any significant weight loss (FIGURE 4F).

Also disclosed is a polypeptide comprising a multispecific antigen binding protein. In some examples, the multispecific antigen binding protein is a bi-specific antibody. In some examples, the polypeptide comprises a multispecific antigen binding protein that binds to EpCAM (epithelial cell adhesion molecule) and an immune cell.

In some examples, the multispecific antigen binding protein is a bispecific immune cell engager that is capable of engaging both an antigen and an immune cell.

In some examples, the polypeptide comprises an anti-EpCAM antigen binding protein. In some examples, the polypeptide comprises a single domain anti-EpCAM antibody, optionally an anti-EpCAM H-chain antibody variable region (i.e., VHH).

In some examples, the polypeptide comprises an anti-immune cell antigen binding protein. In some examples, the polypeptide comprises an anti-immune cell antigen binding protein that binds to an immune cell activation marker. In some

examples, the immune cell activation marker is CD3, NKG2D, CD4, CD8, CD16, CD64, and the like. In some examples, the polypeptide binds to CD3. In some examples, the CD3 may include CD3ζ, CD3ε, CD3γ, CD3δ, and the like.

In some examples, the polypeptide comprises an anti-CD3 antigen binding protein. In some examples, the polypeptide comprises a single-chain variable fragment of an anti-CD3 antibody (anti-CD3 scFv).

In some examples, the immune cell is a T cell, an NK cell, a macrophage, or a monocyte. In some examples, the immune cell is a T cell.

In some examples, the polypeptide is a bispecific antibody/antigen binding protein. In some examples, the bispecific antibody/antigen binding protein is a bispecific T cell engager (BiTE), such as an inducible BiTE, a non-inducible BiTE or a constitutive expression BiTE. In some examples, the bispecific T cell engager (BiTE) binds to two antigens, wherein the first antigen is EpCAM and wherein the second antigen is an immune cell marker. In some examples, the second antigen is an immune cell marker that is involved in the activation of the immune cell. In some examples, the second antigen targeted by the bispecific T cell engager (BiTE) may include but is not limited to CD3, NKG2D, CD28, CD16, CD64, and the like. In some examples, the polypeptide comprises a BiTE that bi-specifically binds to EpCAM and a T cell. In some examples, the polypeptide comprises a BiTE that bi-specifically binds to EpCAM and CD3. In some examples, the EpCAM targeted by the BiTE is modified with an anti-EpCAM VHH paired with an anti-CD3 scFv.

In some examples, the anti-EpCAM VHH pairs with an anti-CD3 scFv that may include a clone Okt3 (Nb01-013A), or an anti-CD3 clone used by another anti-EpCAM in the art (MT110)(Nb01-013B).

In another aspect, there is provided a polynucleotide encoding the cell and/or the polypeptide and/or multispecific antigen binding protein as described herein.

Also disclosed is a polynucleotide comprising a sequence encoding an immune cell engager and a chimeric antigen receptor (CAR), wherein the immune cell engager is capable of bi-specifically binding to EpCAM (epithelial cell adhesion molecule) and an immune cell, and wherein the CAR is capable of binding to a first antigen.

In some examples, there is provided a polynucleotide comprising a sequence encoding a chimeric antigen receptor capable of recognizing, binding, and engaging a GPC3 positive cell. In some examples, the first antigen is GPC3. Therefore, in some examples, the CAR is capable of binding to GPC3.

In some examples, the polynucleotide further comprises sequences encoding one or more co-stimulatory domain, a signal peptide, a hinge, and/or a signalling domain.

In some examples, the co-stimulatory domain may include, but is not limited to 4-1BB, CD28, CD27, OX-40, and the like.

In some examples, the signal peptide may be, but is not limited to, a IgH signal peptide, a IgK signal peptide, a CD8 signal peptide, and the like.

In some examples, the hinge may be, but is not limited to, an IgH hinge, a hinge and / or transmembrane domain of an immunoglobulin-like protein (such as IgA, IgD, IgE, IgG, IgM, and the like), CD28, CD8, 4-1 BB, and the like.

In some examples, the sequence encoding CAR encodes for an antigen binding protein capable of binding GPC3 (or anti-GPC3 antigen binding protein), or fragment, or variant thereof. In some examples, the sequence encoding CAR encodes for a single-chain variable fragment. In some examples, the sequence encoding CAR encodes for an anti-GPC3 scFv.

In some examples, the sequence encoding CAR further encodes for an immune cell signalling domain. In some examples, the signalling domain may include, but is not limited to, a CD3, a TCFζ, a FcRγ, a FcRβ, a CD3γ, a CD3θ, a CD3ε, a CD3η, a CD3ζ, a CD22, a CD79a, a CD79b, a CD66d intracellular domain, and the like. In some examples, the CD3 intracellular domain may be one or more of CD3ζ, CD3ε, CD3γ, CD3δ, and the like. In some examples, the sequence encoding CAR further encodes for a CD3 intracellular domain (signalling domain).

In some examples, the sequence encoding CAR encodes for an anti-GPC3 scFv, a CD28 costimulatory domain, a 4-1BB co-stimulatory domain, an IgH signal peptide, an IgH hinge, and a CD3 signalling domain.

In some examples, the sequence encoding the immune cell engager encodes for a multi-specific antigen binding protein. In some examples, the multi-specific antigen binding protein is a bi-specific antibody. In some examples, the sequence encoding the immune cell engager encodes for an antigen binding protein capable of binding EpCAM (anti-EpCAM antigen binding protein), or fragment, or variant thereof, and an anti-immune cell antigen binding protein.

In some examples, the sequence encoding the immune cell engager encodes for a single chain variable fragment (scFv) or a single variable domain located on a heavy chain (VHH).

In some examples, the sequence encoding the immune cell engager encodes for an anti-EpCAM scFv or an anti-EpCAM VHH.

In some examples, the sequence encoding the immune cell engager encodes for an anti-immune cell antigen binding protein that binds to an immune cell activation marker. In some examples, the immune cell activation marker may include but is not limited to CD3, NKG2D, CD4, CD8, CD16, CD64, and the like.

In some examples, the immune activation marker may be in VHH form or scFv form. In some examples, the sequence encoding the immune cell engager encodes for an antigen binding protein capable of binding CD3 (anti-CD3 antigen binding protein), or fragment, or variant thereof. In some examples, the sequence encoding the immune cell engager encodes for a single chain variable fragment (scFv) or a VHH form. In some examples, the sequence encoding the immune cell engager encodes for an anti-CD3 scFv or an anti-CD3 single domain VHH.

In some examples, the immune cell engager may include a His-tag.

In some examples, the immune cell engager comprises an anti-EpCAM antigen binding protein, a linker, an anti-CD3 scFv or an anti-CD3 single domain VHH, and a His-tag. In some examples, the linker is a cleavable linker, which may include but is not limited to, P2A, T2A, F2A, and the like.

In some examples, the immune cell is a T cell, an NK cell, a macrophage, or a monocyte. In some examples, the immune cell is a T cell.

In some examples, the polynucleotide comprises the sequence encoding an immune cell engager is a sequence encoding a bispecific T cell engager (BiTE), such as an inducible BiTE, a non-inducible BiTE or a constitutive expression BiTE. In some examples, the polynucleotide comprises the sequence encoding BiTE that bi-specifically binds to EpCAM and a T cell. In some examples, the polynucleotide comprises the sequence encoding for an anti-GPC3 scFv CAR with CD3 intracellular domain, and the sequence encoding for a BiTE that binds to EpCAM and a T cell.

Also disclosed is a vector expressing the polynucleotide or polypeptide as described herein. In some examples, the vector is selected from the group consisting of a plasmid, a viral particle, a phage, a baculovirus, a yeast plasmid, a lipid based vehicle, a polymer microsphere, a liposome, and a cell based vehicle, a colloidal gold particle, lipopolysaccharide, polypeptide, polysaccharide, a viral vehicle, an adenovirus, a retrovirus, a lentivirus, an adeno-associated viruses, a herpesvirus, a vaccinia virus, a foamy virus, a cytomegalovirus, a Semliki forest virus, a poxvirus, a pseudorabies virus, an RNA virus vector, a DNA virus vector and a vector derived from a combination of a plasmid and a phage DNA, further optionally wherein said polynucleotide is operatively

linked to an expression control sequence(s) to direct peptide synthesis, even further optionally wherein the vector comprises one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells.

In some examples, the vector is a lentiviral vector.

Also disclosed is a host cell comprising the vector of as disclosed herein or polynucleotide or polypeptide as disclosed herein. In some examples, the host cell comprises cloning or expression vectors as described above and/or nucleic acid sequences encoding for the antigen binding protein, antibodies and binding fragments thereof as described herein.

The host cell can be any type of cell capable of being transformed or transfected with the nucleic acid or vector so as to produce an antigen binding protein or binding fragment/protein thereof encoded thereby. The host cell comprising the nucleic acid or vector can be used to produce the antigen binding protein or binding fragment/protein thereof, or a portion thereof (e.g., a heavy chain sequence, or a light chain sequence encoded by the nucleic acid or vector). After introducing the nucleic acid or vector into the cell, the cell is cultured under conditions suitable for expression of the encoded sequence. The antibody, antigen binding protein, or fragment, or portion of the antibody then can be isolated from the cell.

The host cells may be prokaryotic host cells (such as E. coli) or eukaryotic host cells (such as a yeast cell, an insect cell, or a vertebrate cell). The host cell, when cultured under appropriate conditions, expresses an antibody or binding fragment thereof which can subsequently be collected from the culture medium (if the host cell secretes it into the medium) or directly from the host cell producing it (if it is not secreted). Selection of an appropriate host cell will depend upon various factors, such as desired expression levels, polypeptide modifications that are desirable or necessary for activity, such as glycosylation or phosphorylation, and ease of folding into a biologically active molecule. Selection of the host cell will depend in part on whether the antibody or binding fragment thereof is to be post-transcriptionally modified (e.g., glycosylated and/or phosphorylated). The host cell may comprise a bacterial cell, a yeast cell, an animal cell e.g., a mammalian cell and/or a plant cell.

Suitable mammalian host cells include CHO, myeloma or hybridoma cells. Many are available from the American Type Culture Collection (ATCC), Manassas, Va. Examples include mammalian cells, such as Chinese hamster ovary cells (CHO) (ATCC No. CCL61), human embryonic kidney (HEK) 293 or 293T cells (ATCC No. CRL1573),

3T3 cells (ATCC No. CCL92), or PER.C6 cells. Other cell types of use in expressing antibodies include lymphocytic cell lines, e.g. NSO myeloma cells and SP2 cells, COS cells.

In one embodiment, the host cell expresses/secretes the antigen binding protein, variant or fragment thereof as disclosed herein.

In one aspect, there is provided a cell that expresses/secretes an immune cell engager that is specific to EpCAM.

In one embodiment, the cell is a stem cell, for example selected from the group comprising a mesenchymal stem cell, neural stem cell and a pluripotent stem cell, such as an induced pluripotent stem cell (iPSC). Thus, in one embodiment, the stem cell is a mesenchymal stem cell. In one embodiment, the stem cell is a neural stem cell. In one embodiment, the stem cell is a pluripotent stem cell, such as an iPSC.

In one embodiment, the cell is an immune cell.

In one embodiment, the immune cell expresses/secretes an immune cell engager which is specific to EpCAM. In one embodiment, the immune cell engager is selected from the group comprising a T cell engager, an NK cell engager, a monocyte engager and a macrophage engager.

In one embodiment, the immune cell expresses/secretes a bispecific T cell engager (BiTE) which is specific to EpCAM.

In one embodiment, the immune cell expresses/secretes a bispecific T cell engager (BiTE), such as an inducible, non-inducible or constitutive expression BiTE comprising the antigen binding protein, variant or fragment thereof as disclosed herein.

Surprisingly, the present inventors have established that immune cells, such as CAR T-cells, are able to secrete EpCAM BiTEs. It was not previously known that this was possible. Advantageously, by engineering immune cells to express EpCAM immune engagers, this allows the EpCAM immune engagers to be be secreted at the target site (for example at the site of a solid tumour), thereby minimising toxicity and/or side effects.

In some examples, the immune cell may include but is not limited to a macrophage, a dendritic cell, a T cell, a B cell, an eosinophil, a basophil, a neutrophil, a mast cell, a natural killer T cell (NKT cell), natural killer cell (NK cell), a macrophage, a monocyte, and the like. In one embodiment, the immune cell is a NK cell. In one embodiment, the immune cell is a macrophage. In one embodiment, the immune cell is a dendritic cell. In one embodiment, the immune cell is a monocyte.

In one embodiment, the immune cell is a T-cell. In one embodiment, the immune cell is a CAR T-cell, such as an anti-GPC3, anti-HER2 or anti-CD19 CAR T-cell. Thus,

in one embodiment, the CAR T-cell is an anti-GPC3 CAR T-cell. In one embodiment, the CAR T-cell is an anti-HER2 CAR T-cell. In one embodiment, the CAR T-cell is an anti-CD19 CAR T-cell. In one embodiment, the immune cell is a CAR T-, CAR NK-, CAR macrophage-, or CAR monocyte-cell,

In some examples, the immune cell may bind to more than one host cell antigen. Therefore, in some examples, the immune cell may further bind to one host cell antigen, two host cell antigens, three host cell antigens, four host cell antigens, and the like.

In yet another aspect, there is provided a method of producing / generating the cell as described herein, comprising introducing the polynucleotide as described herein into the cell.

Also disclosed are methods of producing / generating the cell as described herein, wherein the one or more nucleic acid encoding the host cell target antigen are fused via a linker (such as P2A cleavable linker) in a vector; and wherein the vector is transformed or transfected into a host cell. In some examples, the culture supernatant of the host cell transformed or transfected with the vector is harvested.

In some examples, the method comprises introducing into the cell a vector as described herein. In some examples, the vector is introduced via viral transduction.

Also disclosed is a composition comprising the cell or the supernatant of the cell as described herein.

Also disclosed is a pharmaceutical composition comprising the (engineered) cell or the supernatant of the cell as described herein and suitable pharmaceutical composition thereof.

In some examples, the composition is a prophylactic and/or therapeutic composition.

Pharmaceutically acceptable agents for use in the present pharmaceutical compositions include carriers, excipients, diluents, antioxidants, preservatives, colouring, flavouring and diluting agents, emulsifying agents, suspending agents, solvents, fillers, bulking agents, buffers, delivery vehicles, tonicity agents, cosolvents, wetting agents, complexing agents, buffering agents, antimicrobials, and surfactants.

Also disclosed is a composition or pharmaceutical composition as described herein for use in therapy/medicine/vaccine, optionally the composition may further include an excipient and/or stabilizers.

In yet another aspect, there is provided a method of treating a disease in a subject in need thereof, the method comprises administering to the subject the cell of as described herein, optionally the disease is a proliferative disease.

Also disclosed is a method of treating a disease in a subject in need thereof, the method comprises administering to the subject an engineered cell or composition as described herein.

Also disclosed is a method of preventing and/or reducing the severity of symptoms caused by a disease in a subject in need thereof, the method comprises administering to the subject an engineered cell or composition as described herein.

In some examples, the cell or composition or pharmaceutical composition is to be administered to the subject through one or more routes of administration including, but not limited to, topical, intravascular, intravenous, oral, subcutaneous, intraarterial, intrathecal, intraperitoneal, intranasal, intradermal, intramuscular, and the like.

Also disclosed is the use of the cell as described herein in the manufacture of a medicament for preventing and/or treating a disease.

Also disclosed are polynucleotide, an antibody, a bi-specific T cell engager, an engineered cell, an engineered immune cell, a method, a composition, or a pharmaceutical composition as described herein.

In some examples, the disease is a proliferative disease. In some examples, the disease is tumour or cancer. In some examples, the disease is a carcinoma and may include but is not limited to liver cancer (such as hepatocellular carcinoma), lung cancer (such as lung squamous cell carcinoma), stomach cancer (such as gastric adenocarcinoma), breast cancer, skin cancer (such as melanoma), ovary cancer (such as ovarian clear cell carcinoma), kidney cancer, pancreas cancer, head and neck cancer, prostate gland cancer, oesophagus cancer, bladder cancer, colon cancer, childhood cancers (such as hepatoblastomas, nephroblastoma, yolk sac tumours, and the like), and the like.

Additionally, when describing some embodiments, the disclosure may have disclosed a method and/or process as a particular sequence of steps. However, unless otherwise required, it will be appreciated that the method or process should not be limited to the particular sequence of steps disclosed. Other sequences of steps may be possible. The particular order of the steps disclosed herein should not be construed as undue limitations. Unless otherwise required, a method and/or process disclosed herein should not be limited to the steps being carried out in the order written. The sequence of steps may be varied and still remain within the scope of the disclosure.

Furthermore, it will be appreciated that while the present disclosure provides embodiments having one or more of the features/characteristics discussed herein, one or more of these features/characteristics may also be disclaimed in other alternative embodiments and the present disclosure provides support for such disclaimers and these associated alternative embodiments.

SEQUENCES

CDRs are in bold (**CDR1**);in bold and in italics (***CDR2***); or in bold, in italics and underlined (***CDR3***). Highlighted residues/bases show differences between the 1B6 and 1C1 light chain constant domains.

1. Amino-acid Sequences

1.1.	Heavy chain variable domain of human IgG1

1B6, 1C1, 1C11, 1D4, 1H6 (SEQ ID NO: 4)

EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLEWMGG***IIPIFGT A***NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVYYC***ARSLGGRFRY***WGQGTL

1E4 (SEQ ID NO: 8)

QVQLQQSGPGLVKPSQTLSLTCAIS**GDSISSNSVA**WNWIRQSPSRGLEWLGR***TYYRS KWYS***DYAISVKGRLDINPDTSKNQFSLQLNSVTPEDTAVYYC***AREVEGSSYDAFDI***W GQGTM

1.2.	Light chain variable domain of human IgG1

1B6 (SEQ ID NO: 12)

DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQSPQLLIY***LGS*** NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC***MQALQTPYT***FGQGTK

1B6 (light chain constant domain) (SEQ ID NO: 13)

RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVA EQDSKDSTYSLSSTLTLSKADYEKHKLYACEVTHQGLSSPVTKSFNRGEC

1C1 (SEQ ID NO: 12)

DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQSPQLLIY***LGS*** NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC***MQALQTPYT***FGQGTK

1C1 (light chain constant domain) (SEQ ID NO: 14)

RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVT
EQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFSRGEC

1C11 (SEQ ID NO: 16)

EIVLTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQSPQLLIY***LGS***N
RASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC***MQGLQSPWT***FGQGTK

1D4 (SEQ ID NO: 18)

DVVMTQSPLSLPVTPGESASISCRSS**QSLLHSNRYNY**LDWYLQKPGQSPQLLIY***LGS***
NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC***MQALQTPYT***FGQGTK

1E4 (SEQ ID NO: 22)

DIQLTQSPSSLSASVGDRVTITCRAS**QSISDF**LNWYQQKPGKAPKLLIY***AAS***SLQTGVP
SRFGGSGSGTEFTLTISSLQPEDLGTYYC***QQSYIMPDT***FGQGTK

1H6 (SEQ ID NO: 24)

DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQSPQLLIY***LGS***
NRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYC***MQGLQTPYT***FGQGTK

1.3.	Single domain heavy chain variable domain of llama VHH

2C4-VHH (SEQ ID NO: 35)

QVQLQESGGGLVQAGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERELVAV***ITSGGS***
THYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYYC***TNGRWSGDTYYAHH***WGQ
GTQ

1A5-VHH (SEQ ID NO: 28)

QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERELVAV***ITSGGS***
THYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYYC***TNGRWSGDTYYAHH***WGQ
GTQ

1B8-VHH (SEQ ID NO: 32)

QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERELVAF***ITNGGS***
TRYTDPVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYYC***MAGTS***WGQGTQ

2B7-VHH (SEQ ID NO: 34)

QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERELVAV***ITSGGS***
THYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYYC***TNGRWSGDTYYAHL***WGQ
GTQ

2D10-VHH (SEQ ID NO: 36)

QVQLQESGGGLVQAGDSLRLSCADS**GSIFSGND**MAWYRRAPGVERELVAV***ITSGGS***
THYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYYC***TNGRWSGDTYYAHH***WGQ
GTQ

1.4.	Single domain heavy chain variable domain of humanized llama VHH

hu2C4-VHH (SEQ ID NO: 37)

QVQLVESGGGLVQAGGSLRLSCAAS**GSIFSGND**MSWYRQAPGKGLELVAV***ITSGGS T***YYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYC***TNGRWSGDTYYAHH***WGQ GTL

1.5.	Anti-EpCAM Okt3 BiTE

Nb01-013A (SEQ ID NO: 75)

Annotation

2C4-VHH, **2C4-VHH (CDR1)**, ***2C4-VHH (CDR2)***, ***2C4-VHH (CDR3)***

Okt3-VH

Okt3-VL

G4S or $(G4S)_3$ linker

6x His-tag (for the purpose of detecting BiTE by ELISA or Western blot)

QVQLQESGGGLVQAGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERELVAV***ITSGGS T***HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYYC***TNGRWSGDTYYAHH***WGQ GTQVTVSSGGGGSQVQLQQSGAELARPGASVKMSCKASGYTFTRYTMHWVKQRP GQGLEWIGYINPSRGYTNYNQKFKDKATLTTDKSSSTAYMQLSSLTSEDSAVYYCAR YYDDHYCLDYWGQGTTLTVSSGGGGSGGGGSGGGGSQIVLTQSPAIMSASPGEKV TMTCSASSSVSYMNWYQQKSGTSPKRWIYDTSKLASGVPAHFRGSGSGTSYSLTIS GMEAEDAATYYCQQWSSNPFTFGSGTKLEINHHHHHH

2. Nucleotide Sequences

2.1.	Heavy chain variable domain of human IgG1

1B6, 1C1, 1C11, 1D4, 1H6 (SEQ ID NO: 41)

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGTCCTCGGT GAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAGCAGCTATGCT**ATCAGCTG GGTGCGACAGGCCCCTGGACAAGGGCTTGAGTGGATGGGAGGG***ATCATCCCTAT CTTTGGTACAGCA***AACTACGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCA GACACCTCCATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGAC ACGGCCGTGTATTACTGT***GCGAGATCGTTGGGTGGGAGATTTCGCTAC***TGGGGC CAGGGAACCCTG

1E4 (SEQ ID NO: 45)

CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTCGCAGACCCTC
TCACTCACCTGTGCCATCTCC**GGGGACAGTATCTCTAGTAACAGTGTTGCT**TGGA
ACTGGATCAGGCAGTCCCCATCGAGAGGCCTTGAGTGGCTGGGAAGG***ACATACT***
ACAGGTCCAAGTGGTACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACAT
CAACCCAGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACTCCC
GAGGACACGGCTGTGTATTATTGT***GCAAGAGAAGTTGAGGGCAGCAGCTATGAT***
GCTTTTGATATCTGGGGCCAAGGGACAATG

2.2.	Light chain variable domain of human IgG1

1B6 (SEQ ID NO: 49)

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGCCGG
CCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATGGATACAACTA**
TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT***TTG***
GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGC
ACAGATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATT
ACTGC***ATGCAAGCTCTACAAACTCCGTACACT***TTTGGCCAGGGGACCAAG

1B6 (light chain constant domain) (SEQ ID NO: 50)

CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGA
AATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGC
CAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAG
TGTCGCAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGAC
GCTGAGCAAAGCAGACTACGAGAAACACAAACTCTACGCCTGCGAAGTCACCCAT
CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAACAGGGGAGAGTGT

1C1 (SEQ ID NO: 49)

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGCCGG
CCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATGGATACAACTA**
TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT***TTG***
GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGC
ACAGATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATT
ACTGC***ATGCAAGCTCTACAAACTCCGTACACT***TTTGGCCAGGGGACCAAG

1C1 (light chain constant domain) (SEQ ID NO: 51)

CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGA
AATCTGGAACTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGC
CAAAGTACAGTGGAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAG
TGTCACAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCTGAC
GCTGAGCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCAT
CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCAGCAGGGGAGAGTGT

1C11 (SEQ ID NO: 53)

GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGCCGG
CCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATGGATACAACTA**
TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT***TTG***
GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGT
ACAGATTTTACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATTA
CTGC***ATGCAAGGTCTACAAAGTCCCTGGACG***TTCGGCCAAGGGACCAAG

1D4 (SEQ ID NO: 55)

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGTCGG
CCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATAGATACAACTA**
TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT***TTG***
GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGC
ACAGATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATT
ACTGC***ATGCAAGCTCTACAAACTCCGTACACT***TTTGGCCAGGGGACCAAG

1E4 (SEQ ID NO: 59)

GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGAG
TCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCGACTTTT**TAAATTGGTACCA
GCAGAAACCAGGTAAAGCCCCGAAGCTCCTGATCTAT***GCTGCATCG***AGTTTACAA
ACTGGGGTCCCCTCAAGATTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCA
CCATAAGCAGTCTACAACCTGAAGATTTGGGAACTTATTACTGT***CAACAGAGTTAC***
ATTATGCCCGACACTTTTGGCCAGGGGACGAAA

1H6 (SEQ ID NO: 61)

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGCCGG
CCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATGGATACAACTA**
TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT***TTG***
GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGC
ACAGATTTTACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTATT
ACTGC***ATGCAAGGTCTACAGACTCCGTACACT***TTTGGCCAGGGGACCAAG

2.3.	Single domain heavy chain variable domain of llama VHH

1A5 (SEQ ID NO: 65)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAGGGTCTCT
GAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGAC**ATGGCCTG
GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT***ATTACTAGCG***
GTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGA
CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACACG
GCCGTGTATTACTGC***ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT***
CACTGGGGCCAGGGGACCCAG

1B8 (SEQ ID NO: 69)

CAGGTGCAGCTGCAGGAGTCTGGGGGGAGGCTTGGTGCAGCCTGGGGGGTCTCT
GAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAGATTCACATCA**GTGGCCTG
GTACCGCCAGGCTCCAGGAAAGGAGCGCGAGTTGGTCGCATTT***ATTACTAATGG***
TGGTAGCACAAGATATACAGACCCCGTGAAGGGCCGATTCACCATCTCCAGAGA
CAACGCCAAGAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACACG
GCCGTCTATTATTGT***ATGGCGGGTACGTCC***TGGGGCCAGGGGACCCAG

2B7 (SEQ ID NO: 71)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAGGGTCTCT
GAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGAC**ATGGCCTG
GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT***ATTACTAGCG***
GTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGA
CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACACG
GCCGTGTATTACTGC***ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT***
CTCTGGGGCCAGGGGACCCAG

2C4 (SEQ ID NO: 72)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCT
GAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGAC**ATGGCCTG
GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT***ATTACTAGCG***
GTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGA
CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACACG
GCCGTGTATTACTGC***ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT***
CACTGGGGCCAGGGGACCCAG

2D10 (SEQ ID NO: 73)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGACTCTCT
GAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGAC**ATGGCCTG
GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT***ATTACTAGCG***
GTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGA
CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACACG
GCCGTGTATTACTGC***ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT***
CACTGGGGCCAGGGGACCCAG

2.4.	Single domain heavy chain variable domain of humanized llama VHH

hu2C4-VHH (SEQ ID NO: 74)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCT
GAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAGTGGCAATGAC**ATGTCCTG
GTACCGCCAGGCTCCAGGGAAGGGACTCGAGTTGGTCGCGGTT***ATTACTAGCGG***
TGGTAGTACATACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGAC
AATTCCaAGAAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGC

CGTGTATTACTGC***ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCA C***TGGGGCCAGGGGACCCTG

DESCRIPTION OF FIGURES

Example embodiments of the disclosure will be better understood and readily apparent to one of ordinary skill in the art from the following discussions and if applicable, in conjunction with the figures. It should be appreciated that other modifications may be made without deviating from the scope of the invention. Example embodiments are not necessarily mutually exclusive as some may be combined with one or more embodiments to form new exemplary embodiments. The example embodiments should not be construed as limiting the scope of the disclosure.

Figures 1A to 1H show Hep3B xenografts in NSG mouse can be suppressed by anti-GPC3 CAR T (5C4) cells but cannot be completely eradicated as escaped tumour cells lost GPC3 expression. **Figure 1A** shows a FACS analysis of tumour markers (GPC3 and EpCAM) expressed on the Hep3B cells. **Figures 1B to 1D** show the results of two million Hep3B cells that were subcutaneously injected into the right flanks of NSG mice (Day -26). At Day 0, mice were grouped according to the tumour size and ten million anti-GPC3 CAR T (5C4) cells or Mock T cells were intravenously injected into these mice via tail vein (Day 0). **Figure 1B** shows the percentage of CAR expression on anti-GPC3 CAR T (5C4) cells detected by flow cytometry analysis. **Figure 1C** shows tumour sizes that were measured and recorded every 3 to 7 days. **Figure 1D** shows the FACS analysis of tumour markers (GPC3 and EpCAM) expressed on the re-emerged tumours from an anti-GPC3 CAR T (5C4) treated mouse at 70 days post CAR T cell infusion. **Figure 1E to 1H** shows the results of two million Hep3B cells that were subcutaneously injected into the right flanks of NSG mice (Day -21). At Day 0, mice were grouped according to the tumour size and four million sorted anti-GPC3 CAR T (5C4) cells with enriched CAR positive populations or Mock T cells were intravenously injected into these mice via tail vein (Day 0). **Figure 1E** shows that after lentiviral transduction, T cells expressing anti-GPC3 CAR T (5C4) cells were enriched by flow sorting and the percentage of CAR expression on sorted anti-GPC3 CAR T (5C4) cells were detected by flow cytometry analysis. Tumour size (**Figure 1F**) and weight (**Figure 1G**) of each mouse were measured and recorded every 3 to 7 days. **Figure 1H** shows FACS analysis of tumour markers expressed on the re-emerged tumours from three anti-GPC3 CAR T (5C4) treated mice at 71 days post CAR T infusion.

Figures 2A to 2D show the construction and *in vitro* characterization of GE CAR-BiTE T cells. **Figure 2A** shows the structure map of GE CAR-BiTE. **Figure 2B** shows the percentage of CAR expression on anti-GPC3 CAR T (5C4) and GE CAR- BiTE T (sequence of "MT110" was used as BiTE) cells detected by flow cytometry analysis. **Figure 2C** shows the percentage of cytolysis of Hep3B cells mediated by human GE CAR-BiTE T cells with time-course measurement using xCelligence impedance assay (E:T ratio = 1:1). **Figure 2D** shows the percentage of cytolysis of Hep3B cells mediated by anti-EpCAM BiTE activated naïve T cells. Anti-EpCAM BiTE were secreted by human GE CAR-BiTE T cells and naïve human T cells were isolated from PBMCs of a healthy donor. The time-course measurement was performed by xCelligence impedance assay (E:T ratio = 6:1).

Figures 3A to 3G show human GE CAR-BiTE T cells completely eradicated large Hep3B xenografts *in vivo*. **Figures 3A to 3E** show the results after lentiviral transduction, where anti-GPC3 CAR T (5C4) cells, GE CAR-BiTE T cells or mock T cells were cultured and expanded in T cell growth medium containing IL-2 (50 unit per ml). Two million Hep3B cells were subcutaneously injected into the right flank of NSG mice (Day -27). On Day 0, mice were re-grouped according to the measurable tumour size (3-5 mice per group) and eight million GE CAR-BiTE cells, CAR T (5C4) cells, or mock T cells were intravenously injected into these mice via tail vein. **Figure 3A** shows the percentage of CAR expression on anti-GPC3 CAR T (5C4) and GE CAR- BiTE T (sequence of "MT110" was used as BiTE) cells detected by flow cytometry analysis. Weight (**Figure 3B**) and tumour size (**Figure 3C**) of each mouse were measured and recorded every 3 to 7 days. **Figure 3D** shows tumour sizes of each individual mouse in each treatment group were shown. **Figure 3E** shows tumours of selected mice were excised and the expression of the tumour markers GPC3 and EpCAM were analyzed. **Figures 3F and 3G** shows results after lentiviral transduction, where anti-GPC3 CAR T (5C4) cells, GE CAR-BiTE T cells or mock T cells were cultured and expanded in T cell growth medium containing IL-7 (20 ng/ml) and IL-15 (5 ng/ml). Two million Hep3B (GPC3 positive, EpCAM positive) cells were subcutaneously injected into the right flank of NSG mice (Day -21). On Day 0, mice were re-grouped according to the measurable tumour size (3-6 mice per group) and eight million GE CAR-BiTE cells, CAR T (5C4) cells, or mock T cells were intravenously injected into these mice via tail vein. **Figure 3F** shows the tumour sizes of each mouse were measured and recorded every 3-4 days. **Figure 3G** shows the survival curves of the mice from "Tumour only" group (grey dotted line), "Mock T" group (grey solid line), "CAR T (5C4)" group (black dotted line) and "GE CAR-

BiTE T" group (black solid line). Death was defined either when autonomous death of the mice was observed or when the bearing tumour size of the mice exceeds 2700mm^3.

Figures 4A to 4G show human GE CAR-BiTE T cells completely eradicated HepG2 xenografts *in vivo*. **Figures 4A and 4B** show the results after one million HepG2 cells were subcutaneously injected into the right flank of NSG mice (Day -9). On Day 0, mice were re-grouped according to the measurable tumour size and ten million CAR T (5C4) cells or mock T cells were intravenously injected into these mice via tail vein. **Figure 4A** shows the percentage of CAR expression on anti-GPC3 CAR T (5C4) cells detected by flow cytometry analysis. **Figure 4B** shows tumour sizes of each mouse were measured and recorded every 3 to 7 days. **Figures 4C to 4G** shows the results when one million HepG2 cells were subcutaneously injected into the right flank of NSG mice (Day -4). On Day 0, mice were re-grouped according to the measurable tumour size and three million GE CAR-BiTE T cells, CAR T (5C4) cells, or mock T cells were intravenously injected into these mice via tail vein. **Figure 4C** shows the percentage of CAR expression on anti-GPC3 CAR T (5C4) and GE CAR- BiTE T (sequence of "MT110" was used as BiTE) cells detected by flow cytometry analysis. **Figure 4D** shows the tumour sizes of each mouse were measured and recorded every 3 to 7 days. **Figure 4E** shows the tumour sizes of each individual mouse in each treatment group were shown. **Figure 4F** shows the weights of each mouse were measured and recorded every 3 to 7 days. **Figure 4G** shows the survival curves of the mice. Death was defined either when autonomous death of the mice was observed or when the bearing tumour size of the mice exceeds 1700mm^3.

Figures 5A to 5H show GE CAR-BiTE T cells using anti-EpCAM VHH constructed BiTEs has superior properties in killing potency and specificity. **Figure 5A** shows the structural format of MT110 BiTE, Nb01-013A BiTE or Nb01-013B BiTE antibodies secreted by GE CAR-BiTE cells. **Figure 5B** shows FACS analysis of tumour markers (GPC3 and EpCAM) expressed on HepG2, HT-29 and HeyA8 cells. Killing assay of HepG2 (**Figure 5C**), HT-29 (**Figure 5D**) and HeyA8 (**Figure 5E**) cells by CAR T (5C4) or GE CAR-BiTE T cells (MT110, Nb01-013A or Nb01-013B) was performed using xCelligence cell impedance assay and presented over 72 hours post co-culture of target cells and effector cells (E:T ratio = 2:1). ELISA measurement of Interferon-r was measured using culture supernatant collected at 48 hours post co-culture effector CAR T cells with HepG2 (**Figure 5F**), HT-29 (**Figure 5G**) and HeyA8 (**Figure 5H**) cells.

Figures 6A and 6B shows human GE CAR-BiTE T cells using anti-EpCAM VHH constructed BiTEs are more superior tumour killing in *in vivo* HepG2 xenografts. One

million HepG2 cells were subcutaneously injected into the right flank of 30 male NSG mice (Day -4). On Day 0, mice were re-grouped according to the measurable tumour size and three million CAR T (5C4) cells, GE CAR-BiTE T cells ("MT110", "Nb01-013A" or "Nb01-013B" was used as BiTE), or Mock T cells were intravenously injected into these mice via tail vein. **Figure 6A** shows the tumour sizes of each individual mouse in each treatment group were measured and recorded every 3 to 7 days. **Figure 6B** shows the weights of different groups of mice were measured and recorded every 3 to 7 days.

Figures 7A to 7D show GE CAR-BiTE T cells using anti-EpCAM VHH constructed BiTEs exerted superior killing potency and the secreted anti-EpCAM BiTE by GE CAR-BiTE T cells could elicit target cell killing independent of high GPC3 expression. **Figure 7A** shows the structural formats of anti-GPC3 CAR T (5C4) or GE CAR-BiTE T cells (anti-GPC3 CAR T secreting anti-EpCAM BiTE, "MT110", "Nb01-013A" or "Nb01-013B" was used as BiTE) or 19E CAR-BiTE T cells (anti-CD19 CAR T secreting anti-EpCAM BiTE, "Nb01-013A" or "Nb01-013B" was used as BiTE). *In vitro* killing assay of (**Figure 7B**) Hep3B ($GPC3^{High}$, $EpCAM^{High}$), (**Figure 7C**) HT-29 ($GPC3^{Low}$, $EpCAM^{High}$) and (**Figure 7D**) HeyA8 ($GPC3^{-ve}$, $EpCAM^{-ve}$) cells by anti-GPC3 CAR T or GE CAR-BiTE T cells or 19E CAR-BiTE T cells was performed using xCelligence cell impedance assay and presented over 72 hours post co-culture of target cells and effector cells (E:T ratio = 1:1).

Figures 8A to D show Human GE CAR-BiTE T cells using anti-EpCAM VHH constructed BiTEs showed superior tumour killing in *in vivo* Hep3B xenografts. **Figure 8A** shows a schematic representation of the timeline of the *in vivo* Hep3B xenograft model. Two million Hep3B cells were subcutaneously injected into the right flank of 50 male NSG mice (Day -21). On Day 0, mice were re-grouped according to the measurable tumour size and eight million CAR T (5C4) cells, GE CAR-BiTE T cells (Anti-EpCAM BiTE secreting anti-GPC3 CAR T cells, "MT110", "Nb01-013A" or "Nb01-013B" was used as BiTE), 19E CAR-BiTE T cells (Anti-EpCAM BiTE secreting anti-CD19 CAR T cells, "Nb01-013A" or "Nb01-013B" was used as BiTE), or Mock T cells were intravenously injected into these mice via tail vein. Nighty-two days post CAR T or CAR-BiTE T cell treatment, three surviving mice from the GE CAR-BiTE T (Nb01-013A) group were re-challenged with one million Hep3B cells (subcutaneously injected into the left flank) together with five naïve male NSG mice as controls ("Tumour only"). **Figure 8B** shows the tumour sizes of each mouse were measured and recorded every 3 to 7 days. The curve showed tumour sizes up to Day 92. **Figure 8C** shows the tumour sizes of each

individual mouse in each treatment group were shown up to Day 92. **Figure 8D** shows the tumour sizes of the re-challenged tumours at the left flank (Day 92 onwards).

Figures 9A to 9F show Human GE CAR-BiTE T cells using anti-EpCAM VHH constructed BiTEs showed superior *in vivo* T cell expansion in Hep3B mouse xenograft models. In the same experiment as in **Figure 8**, cheek bleed on weekly intervals was done for flow cytometry analysis to monitor the T cell numbers in the peripheral blood, as well as Luminex® Multiplex Assays on the cytokine levels. **Figure 9A** shows the number of CD3 positive human T cells in the mouse peripheral blood from selected groups of mice after 5 weeks of T cell infusion. **Figure 9B** shows a time-course analysis of CAR positive human T cells in the mouse peripheral blood in the group of CAR T (5C4) treatment. **Figure 9C** shows a time-course analysis of CAR positive human T cells in the mouse peripheral blood in the group of GE CAR-BiTE T (Nb01-013A) treatment. The levels of Interferon-gamma (**Figure 9D**), GM-CSF (**Figure 9E**), and Perforin (**Figure 9F**) in the peripheral blood from the selected groups of mice after 5 weeks of T cell infusion were measured by Luminex® Multiplex Assays.

Figures 10A to 10C show Human GE CAR-BiTE cells using anti-EpCAM VHH constructed BiTEs showed superior localized T cell expansion within the tumour in Hep3B mouse xenograft models. In the same experiment as in **Figure 8**, after 5 weeks of T cell infusion, selected mice were sacrificed, and the xenograft tumours were excised and dissociated for flow cytometry analysis for analysing the numbers of human T cells within the tumours. **Figure 10A** shows the numbers of CD3 positive human T cells in the mouse tumours from selected mice were presented as number of cells per gram tumour. **Figure 10B** shows the numbers of CAR positive human T cells in the mouse tumours from selected mice were presented as number of cells per gram tumour. **Figure 10C** shows the percentages of CAR positive T cells within the population of human T cells in the mouse tumours from the selected mice. * Marked the mice showing effective tumour controls.

Figure 11 shows a comparison of CAR T cells enriched in the tumours and those dispersed in the peripheral blood. In the same experiment as in **Figure 8**, after 5 weeks of T cell infusion, selected mice with observed tumour control effects (M18, M49, M38 and M40) were analyzed by flow cytometry analysis for both CAR T cells in the peripheral blood and CAR T cells from the excised tumour. The total numbers of CAR positive T cells in tumours were calculated by multiplying the total weights of the tumours, while the total numbers of CAR positive T cells in the peripheral blood were calculated by multiplying the average blood volume of the experimental mice (~2ml).

Figure 12 shows the detection of anti-EpCAM BiTE molecule Nb01-013A secreted from GE CAR-BiTE T cells by ELISA. **Figure 12A** shows the detection of anti-EpCAM BiTE molecule when anti-GPC3 CAR T (5C4) cells or GE CAR-BiTE T (Nb01-013A) cells were cultured in T cell growth medium containing IL-7 (20 ng/ml) and IL-15 (5 ng/ml). The starting cell density was 0.5 million per ml and the percentages of CAR expression in both cultures were around 60%. Cell culture supernatants were collected daily (24h, 48h, 72h, 96h) for an ELISA to detect the amount of secreted anti-EpCAM BiTE molecule Nb01-013A. **Figure 12B** shows the detection of anti-EpCAM BiTE molecule when GE CAR-BiTE T (Nb01-013A) cells were co-cultured with Hep3B ($GPC3^{High}$, $EpCAM^{High}$), HT-29 ($GPC3^{Low}$, $EpCAM^{High}$) or HeyA8 ($GPC3^{-ve}$, $EpCAM^{-ve}$) cells at an E:T ratio of 2:1. The percentage of CAR expression in GE CAR-BiTE T (Nb01-013A) cells used in all co-cultures was 8.5%. Supernatants were collected daily (24h, 48h, 72h) for an ELISA to detect the amount of secreted anti-EpCAM BiTE molecule Nb01-013A after the co-culture of GE CAR-BiTE T (Nb01-013A) cells with different target cells. Briefly, human EpCAM-Fc tag protein was used to coat the ELISA plate overnight, after blocking with Casein for 2 hours, the culture supernatants containing the anti-EpCAM BiTE molecule Nb01-013A were added to the plate. After 1-hour incubation, the plate was washed, and the bound BiTE molecules were detected by an HRP conjugated secondary antibody against the His-tag.

EXPERIMENTAL SECTION

GPC3 single targeting CAR T cells resulted in antigen loss in *in vivo* xenograft mouse models

The *in vivo* treatment efficacy of the 5C4 CAR T cells has been demonstrated using human hepatocellular carcinoma cells Hep3B engrafted NSG mouse models in two independent experiments (Figure 1). Hep3B cells in *in vitro* cultures have shown to express high level of GPC3 and EpCAM (Figure 1A). In the first experiment, 5C4 CAR T cells were successful produced with 25.9% CAR expression (Figure 1B). CAR T cells showed excellent tumour control, effectively suppressing the growth of the engrafted Hep3B tumours (Figure 1C). However, following a tumour-free period, tumours re-emerged (Figure 1C) and appeared to lose the antigen target GPC3 while EpCAM expression remained high (Figure 1D). In the second experiment, sorted anti-GPC3 CAR T (5C4) cells with highly enriched CAR positive populations (94.1%, Figure 1E) were

used to treat the same Hep3B xenograft. Despite treatment with these enriched CAR expressing cells showed enhanced tumour control as compared to that of the first experiment (Figure 1F), it still failed to prevent tumour regrowth. These escaped tumour cells were, again, $GPC3^{-ve}$ and $EpCAM^{+ve}$, which further confirmed the notion of multi-targeting approach in treating the solid tumours. Moreover, GPC3-targeted CAR T cell treatments did not cause body weight loss in these mice, suggesting low or no toxicity (Figure 1G).

Construction of GE CAR-BiTE T constructs and *in vitro* functional characterization

The construct of GE CAR-BiTE T was designed by fusing the anti-EpCAM BiTE gene (sequence of "MT110" was used) to the anti-GPC3 CAR (clone 5C4) lentiviral construct via a P2A cleavable linker (Figure 2A). Following lentiviral transduction and expansion (Figure 2B), CAR T cells were tested for their *in vitro* killing efficacy. Interestingly, the GE CAR-BiTE T cells killed the Hep3B ($GPC3^{High}$, $EpCAM^{High}$) target cells at a slightly faster rate than the anti-GPC3 CAR alone (Figure 2C), suggesting the crucial role of the BiTE in facilitating target-cell killing. The GE CAR-BiTE cytotoxic function was further tested in an assay in which only the culture supernatant harvested from the GE CAR-BiTE T cells were shown to direct efficient T cells (naïve T cells isolated from a donor PBMC) mediated target cell killing (Figure 2D).

Human GE CAR-BiTE T cells completely eradicated large Hep3B xenografts *in vivo*.

The efficacy of tumour control by the GE CAR- BiTE T was assessed with Hep3B xenograft in the NSG mice, with the transduction rate assessed by flow cytometry analysis (Figure 3A). While both the CAR T (5C4) and GE-CAR-BiTE T treatments reduced tumour growth, only the GE-CAR-BiTE treated mice achieved complete tumour regression (Figure 3C, 3D) and similarly exhibited good safety profile without causing weight loss of mice (Figure 3B). Further analysis revealed that the tumour cells that survived anti-GPC3 CAR T (5C4) treatment lost GPC3 while still expressing EpCAM (Figure 3E). Nevertheless, when the inventors analysed cells isolated from the residual tissues from the tumour engraftment site from a representative GE CAR-BiTE T cell treated mouse, the inventors could not detect any cells expressing GPC3 or EpCAM (Figure 3E). It is highly possible that the double negative cells are mouse stromal cells residing in the engraftment site and all human Hep3B tumour cells have been eradicated upon GE CAR-BiTE T cell treatment. Next, the inventors have slightly modified the T cell culture protocol by changing IL-2 to a combination of IL-7 and IL-15 in the culture

medium, which is now a widely accepted standard protocol in CAR T manufacturing in clinical settings. The repeated experiment using GE CAR-BiTE T cells expanded in culture medium containing IL-7 and IL-15 recapitulated the results of superior efficacy in tumour control (Figure 3F) and long-term mouse survival (Figure 3G).

Human GE CAR-BiTE T cells completely eradicated HepG2 xenografts *in vivo*.

Similar to Hep3B, another HCC cell line HepG2 similarly expressed high level of GPC3 and EpCAM on the cell surface. However, compared to the Hep3B model, anti-GPC3 targeting CAR T alone was relatively ineffective in treating HepG2 derived xenograft even if CAR T cells were infused at a much earlier date post tumour inoculation (9 days in the HepG2 model, compared to 21-27 days in the Hep3B model) (Figure 4A and 4B). The inventors suspected that this is because HepG2 xenograft had a much faster growth rate than Hep3B xenograft in mice. The inventors wondered if the GE CAR-BiTE T cell treatment could be successful in this more aggressive mouse model. After lentiviral transduction (Figure 4C) and four days post HepG2 tumour cell inoculation, CAR T (5C4) alone or GE CAR-BiTE T cells were infused to the mice by intravenous injection. Different from the inefficiency of CAR T (5C4) treated group, tumours from all three GE CAR-BiTE T cell treated mice were successfully cured (Figure 4D, Figure 4E) and the mice all survived (Figure 4G) without any significant weight loss (Figure 4F) until the experiment was voluntarily ended at Day 80.

Human GE CAR-BiTE cells using new BiTEs showed superior properties in killing potency and specificity.

The inventors have noticed that the GE CAR-BiTE T cells using "MT110" generally have lower T cell viability and poor T cell expansion rate compared to CAR T (5C4) cells. Hence "MT110" was replaced with a different version of BiTE where the anti-EpCAM arm was replaced with an in-house identified anti-EpCAM VHH (clone 2C4) and was paired with two different anti-CD3 scFv (clone Okt3 or the anti-CD3 clone used in "MT110"). They were named as "Nb01-013A" and "Nb01-013B" respectively (Figure 5A). To compare their cytotoxicity effect with GE CAR-BiTE T cells using "MT110", three representative cell lines including HepG2 (express high levels of both GPC3 and EpCAM), HT-29 (express very low level of GPC3 but high level of EpCAM), and HeyA8 cells (do not express GPC3 or EpCAM) (Figure 5B) were chosen. Using the three cells expressing different levels of surface antigens, the results shown that the GE CAR-BiTE cells equipped with the new BiTEs (Nb01-013A or Nb01-013B) displayed superior properties in killing HepG2 and HT-29 cells at a comparable potency as MT110 BiTE.

Interestingly, while T cells secreting MT110 also induced strong cytotoxicity against HeyA8 cells, GE CAR-BiTE T cells using new BiTEs (Nb01-013A or Nb01-013B) spared this target negative cells (Figure 5C, 5D and 5E). This might be due to a non-specific binding of the anti-EpCAM arm of MT110 to a different unknown cell surface protein, or a specific binding to EpCAM protein, which is expressed at an extremely low level falling out of the detection limit by flow cytometry. This may be attributed to the high affinity of MT110 against EpCAM. The T cell activation mediated target cell killing can also be reflected by a parallel interferon-γ ELISA measurement (Figure 5F, 5G and 5H).

In summary, the present inventors have developed a novel CAR T cell therapy in targeting GPC3 positive solid tumours by arming the single antigen targeting CAR T cells with an additional targeting of EpCAM positive cancer cells by localized secretion of anti-EpCAM BiTE to overcome the challenges of tumour heterogeneity. By using CAR T cells secreting anti-EpCAM BiTE, the efficacy of CAR T cell therapies could be maximised by recruiting the nearby bystander T cells (T cell populations without expressing a CAR) to contribute to the tumour clearance and concurrently avoid the systematic toxicity that might be brought by other means of delivery of anti-EpCAM BiTEs. The present disclosure has shown the development of a prototype (GE CAR-BiTE), which also serves as a proof-of-concept. While GE CAR-BiTE has been demonstrated to work successfully, the present technology is not limited to GPC3- targeting CAR but can be extended to many other tumour antigens which will be applied as a CAR target for various types of carcinomas happened in multiple organs.

Human GE CAR-BiTE cells using VHH BiTE molecules showed superior properties in tumour killing potency in HepG2 mouse xenograft models.

After constructing GE CAR-BiTE T cells using different anti-EpCAM BiTE molecules ("MT110", Nb01-013A", and "Nb01-013B"), the inventors proceeded to test their tumour control potency using the *in vivo* xenograft models. HepG2 xenograft model was first tested (Figure 6). While CAR T (5C4) failed to control the tumour growth, all three groups of GE CAR-BiTE T cell treatments showed some *in vivo* tumour control efficacies (Figure 6A) without significant weight loss (Figure 6B). Tumour shrinkage was observed in 40% (2/5 mice), 60% (3/5 mice) and 80% (4/5 mice) of mice, respectively from the "MT110", "Nb01-013A" and "Nb01-013B" group (Figure 6A).

Human GE CAR-BiTE cells using VHH BiTE molecules showed superior CAR antigen independent *in vitro* cytotoxicity in killing EpCAM expressing cells.

Next, the GE CAR-BiTE T cells using different anti-EpCAM BiTEs ("MT110", "Nb01-013A", and "Nb01-013B"), together with anti-CD19 CAR T cells secreting anti-EpCAM BiTEs ("Nb01-013A" or "Nb01-013B"), named as "19E CAR-BiTE T" cells (Figure 7A), were tested for their *in vitro* killing efficacies using Hep3B ($GPC3^{High}$, $EpCAM^{High}$), HT-29 ($GPC3^{Low}$, $EpCAM^{High}$) and HeyA8 ($GPC3^{-ve}$, $EpCAM^{-ve}$) cells.

An E:T ratio of 1:1 (adjusted to 10% CAR expression level) was applied in the *in vitro* killing assay with the xCelligence RTCA system. All GE CAR-BiTE T cells including the GE CAR-BiTE T (MT110) cells did not show killing of HeyA8 (GPC3-ve, EpCAM-ve) cells at the low E:T ratio (Figure 7D), however, a strong cytotoxicity was observed in killing EpCAM positive cells Hep3B (Figure 7B) and HT-29 (Figure 7C).The 19E CAR-BiTE T cells ("Nb01-013A" used as BiTE) showed partial killing of both EpCAM positive cells Hep3B ($GPC3^{High}$, $EpCAM^{High}$)(Figure 7B) and HT-29 ($GPC3^{Low}$, $EpCAM^{High}$)(Figure 7C), indicating that the amount of anti-EpCAM BiTEs secreted by the 19E CAR-BiTE T (Nb01-013A) cells in the current *in vitro* settings exceeded the threshold for BiTE mediated cell killing and was independent of the CAR-specific antigen expression level. On the contrary, the 19E CAR-BiTE T cells ("Nb01-013B" used as BiTE) only showed a minimal killing of HT-29 ($GPC3^{Low}$, $EpCAM^{High}$) cells, suggesting that "Nb01-013B" BiTE has a much weaker potency in mediating T cell killing than "Nb01-013A" BiTE (Figure 7C).

Human GE CAR-BiTE cells using VHH BiTE molecules showed superior CAR antigen dependent tumour killing potency in Hep3B mouse xenograft models.

With this observation, it is hypothesised if the *in vivo* tumour control is dependent or independent of the CAR-specific antigen expression (Figure 8). Interestingly, Hep3B xenografts from the two 19E CAR-BiTE T cell treated groups of mice showed a similar growth rate as the control groups ("Tumour only" group and "Mock T" group), suggesting that unlike the *in vitro* setting, the *in vivo* tumour control efficacy by CAR-BiTE T cells is dependent on the CAR-specific antigen expression. In this experiment, the inventors did not observe the effect of GE CAR-BiTE T (MT110) cells in controlling the Hep3B xenografts as they had always observed in their previous experiments. This was because the mice were voluntarily culled when the humane point was hit (tumour sizes exceeded 3000 mm^3), before the *in vivo* expansion of GE CAR-BiTE T (MT110) cells reached the threshold points in executing a significant tumour control effect. This may be due to the donor variations in the T cells used to generate the CAR T cells in different experiments. Nevertheless, GE CAR-BiTE T ("Nb01-013A" or "Nb01-013B" used as BiTE) cells

generated using the same T cell donor showed much potent tumour control efficacies, suggesting that replacing the anti-EpCAM arm with our anti-EpCAM 2C4-VHH significantly improved the *in vivo* expansion as well as the *in vivo* killing potency of the GE CAR-BiTE T cells. GE CAR-BiTE T (Nb01-013A) group of mice achieved efficient tumour control for all 9 tested mice. In addition, three mice lived in a healthy status for more than 90 days after the treatment (Figure 8B, Figure 8C). At Day 92 post the GE CAR-BiTE T (Nb01-013A) cell treatment, a re-challenge test was performed with 1 million Hep3B cells injected sub-cutaneously into the left flank of the three remaining live mice. Five naïve NSG mice were also included as "Tumour only" controls. After about 20 days, tumours started to grow from all naïve mice injected with 1 million Hep3B cells and reached an average of 445.3 mm^3 at Day 133 (41 days post s.c. injection of 1 million Hep3B cells). However, tumour growth was not observed at the left flanks of any of the three re-challenged mice from the GE CAR-BiTE T (Nb01-013A) group (Figure 8D).

Human GE CAR-BiTE cells using VHH BiTE molecules showed superior *in vivo* T cell expansion in Hep3B mouse xenograft models.

To monitor the *in vivo* expansion and cytotoxic functions of the CAR T cells, the inventors have performed cheek bleed on weekly intervals for flow cytometry analysis on T cell numbers, as well as Luminex® Multiplex Assays on cytokine levels (Figure 9). A significant expansion of human T cells could be detected from the mouse blood from week 4 onwards (Day 28 post T cell treatment) and the numbers continued to increase before reaching the peak levels and subsided afterwards (Figure 9B, Figure 9C). As expected, GE CAR-BiTE T (Nb01-013A) and GE CAR-BiTE T (Nb01-013B) group of mice had much more profound T cell expansion than GE CAR-BiTE T (MT110) group of mice when the level of CD3 positive T cells across each different groups were compared after 5 weeks of T cell infusion (Figure 9A). In addition, the levels of serum cytokines such as human Interferon-gamma (Figure 9D), GM-CSF (Figure 9E) and perforin (Figure 9F) secreted by human T cells from the former two groups of mice were comparable to the levels from the CAR T (5C4) group of mice, which were much higher than GE CAR-BiTE T (MT110) and three control groups including Mock T, 19E CAR-BiTE T (Nb01-013A) and 19E CAR-BITE T (Nb01-013B). As a result of efficient T cell expansion in the mouse body, all tumours from the CAR T (5C4) group and GE CAR-BiTE T (Nb01-013A) group, as well as five tumours from the GE CAR-BiTE T (Nb01-013B) group were observed to shrink before they reached 3000 mm^3 (Figure 8C).

Human GE CAR-BiTE cells using VHH BiTE molecules showed superior localized T cell expansion within the tumour in Hep3B mouse xenograft models.

On Day 33-35, all mice from "Tumour only", "Mock T", "GE CAR-BiTE T (MT110)", "19E CAR-BiTE T (Nb01-013A)" and "19E CAR-BiTE T (Nb01-013B)" groups were voluntarily culled as most of their tumour sizes are close to the humane point. The inventors have selected 1 to 2 mice from each group to analyse the tumour infiltrating human T cells from the excised tumours. As in CAR T (5C4) and GE CAR-BiTE T (Nb01-013A) group, all treated mice showed good control of tumour growth and hence the inventors randomly chose one mouse from each of these two groups for FACS analysis but leaving the rest 8 mice for continuous monitoring. However, for the GE CAR-BiTE T (Nb01-013B) group, not all mice showed tumour shrinkage at the date of analysis. Hence, the inventors have included a total of five mice for FACS analysis, where effective tumour controls were observed in M38 and M40 but not in M34, M37 and M39. Interestingly, the inventors observed an abundant accumulation of human T cells (CD3 positive cells, Figure 10A) and CAR positive T cells (Figure 10B) in M18, M49, M38 and M40, suggesting that the effective CAR T cell expansion within the tumour is necessary for effective tumour control activities. In addition, a moderate to high percentage of anti-GPC3 CAR T cells can be detected in M18 (CAR T group) and all mice from GE CAR-BiTE T groups (M29, M49, M34, M37, M38, M39, and M40), but not in 19E CAR-BiTE T groups (M27, M22, and M23) (less than 10% CAR positive T cell populations) (Figure 10C), suggesting that CAR positive T cells will only be enriched *in vivo* when the CAR T cells encountered their CAR targeting antigens.

More intriguingly, when the inventors compared the total numbers of CAR positive T cell population in the excised tumours and those in the peripheral blood from the four mice showing effective tumour controls (Figure 11), the amount of CAR positive T cells located in the tumours were observed to be significantly higher than those circulated in the peripheral blood.

Anti-EpCAM llama VHH BiTE secreted by GE CAR-BiTE T cells can be detected from the cell cultures

Both anti-GPC3 CAR T (5C4) and GE CAR-BiTE T (Nb01-013A) cells were cultured in T cell growth medium containing IL-7 and IL-15 at the starting cell density of 0.5 million per ml. Cell culture supernatants were collected daily for a consecutive four days, followed by an ELISA to detect the amount of secreted anti-EpCAM BiTE molecule Nb01-013A. The results clearly showed that the amount of secreted Nb01-013A

molecules increased over time, suggesting a continuous secretion and accumulation of anti-EpCAM VHH BiTE from the GE CAR-BiTE T (Nb01-013A) cell cultures (Figure 12A).

Interestingly, in another experiment, GE CAR-BiTE T (Nb01-013A) cells were co-cultured with different target cells, including Hep3B ($GPC3^{High}$, $EpCAM^{High}$), HT-29 ($GPC3^{Low}$, $EpCAM^{High}$) and HeyA8 ($GPC3^{-ve}$, $EpCAM^{-ve}$). The percentage of CAR expression in GE CAR-BiTE T (Nb01-013A) cells used in all co-cultures was 8.5%. Culture supernatants from the co-cultures were also collected daily over three days and the ELISA demonstrated that only after co-culture with Hep3B cells, the secretion of Nb01-013A increased overtime, indicating that an expansion of CAR positive T cell populations only happened when the GE CAR-BiTE T (Nb01-013A) cells encountered the GPC^{High} target cells (Figure 12B).

In summary, the present disclosure has shown that the usage of VHH BiTE molecules constructed from an anti-EpCAM VHH clone 2C4 ("Nb01-013A" and "Nb01-013B") in GE CAR-BiTE T cells achieved a more efficient and potent tumour killing than GE CAR-BiTE T cells using "MT110", with the best efficacy observed when using "Nb01-013A" as the BiTE molecule.

APPLICATIONS

Embodiments of modified cells as disclosed herein takes advantage of the following factors: 1) tumour- directed T cell expansion, and 2) tumour-restricted anti-EpCAM BiTE secretion, the latter being unachievable by the traditional approach of systemic delivery. The anti-EpCAM BiTE secreting CAR T cell as disclosed herein can be expanded to other cancer types to establish a broader scope of next-generation CAR T cell therapies.

This technology described in the current TDF may have the following applications:

1. Anti-GPC3 CAR T cells secreting anti-EpCAM BiTE can be used to treat GPC3 expressing cancers such as hepatocellular carcinoma, ovarian clear cell carcinoma (OCCC), lung squamous cell carcinomas, melanomas, and some childhood cancers (hepatoblastomas, nephroblastoma, and yolk sac tumours), as well as certain stomach cancers.

2. Anti-EpCAM BiTE secreting CAR T cells may be configured to target other tumour specific antigens including but not limited to HER2, Claudin18.2, ROR1, DLL3, CEA, MUC1, MUC16, CEACAM7, CD133, CD147, PSCA, PSMA, MSLN, c-Met, FRα, and the like. These anti-EpCAM BiTE secreting CAR T cells can be used to target various solid cancers including liver, breast, head and neck, gastric, pancreas, lung, prostate, kidney, ovarian, oesophagus, bladder, and colon cancers.

This present disclosure provides for an improvement to treatment efficacy of the current single targeting CAR T cells against solid tumours by advantageously overcoming the tumour heterogeneity and treatment associated tumour escape using anti-EpCAM BiTE. Also, by applying local secretion of anti-EpCAM BiTE by CAR T cells, it converts a non-druggable target into a druggable target as it diminishes the on-target, off-tumour toxicity that might be brought by systematic delivery. In addition, by targeting cancer progenitor and cancer stem cells that express EpCAM, it further contributes to complete tumour eradication and tumour relapse prevention.

It will be appreciated by a person skilled in the art that other variations and/or modifications may be made to the embodiments disclosed herein without departing from the spirit or scope of the disclosure as broadly described. For example, in the description herein, features of different exemplary embodiments may be mixed, combined, interchanged, incorporated, adopted, modified, included etc. or the like across different exemplary embodiments. The present embodiments are, therefore, to be considered in all respects to be illustrative and not restrictive.

CLAIMS

1. A modified cell expressing

(a) a chimeric antigen receptor targeting GPC3 or CD19, and

(b) a multi-specific antigen binding protein, variant or binding fragment thereof that binds one or more target, comprising a first antigen binding protein, variant or binding fragment thereof that binds to EpCAM (epithelial cell adhesion molecule) and a second antigen binding protein, variant or binding fragment thereof that binds to an immune cell marker,

wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region and/or a light chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising:

a CDR-H1 comprising:

- GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
- GSSERFTS (**SEQ ID NO: 29**)

a CDR-H2 comprising:

- ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
- ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and

a CDR-H3 comprising:

- TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
- TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

(ii) a heavy chain variable region comprising:

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2)** or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

(iii) a heavy chain variable region comprising:

a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5)** or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6)** or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7)** or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

(iv) a light chain variable region comprising:

a CDR-L1 comprising:

- QSLLHSNGYNY (**SEQ ID NO: 9)** or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- QSLLHSNRYNY (**SEQ ID NO: 17**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
- QSISDF (**SEQ ID NO: 19**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

a CDR-L2 comprising:

- LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or

- AAS (**SEQ ID NO: 20** - 1E4) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising:

- MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- MQGLQSPWT (**SEQ ID NO: 15**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- QQSYIMPDT (**SEQ ID NO: 21**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
- MQGLQTPYT (**SEQ ID NO: 23**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

2. The cell of claim 1, wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region and/or a light chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising:

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

(ii) a heavy chain variable region comprising:

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

(iii) a heavy chain variable region comprising:

a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

(iv) a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(v) a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

(vi) a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNRYNY (**SEQ ID NO: 17**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

(vii) a light chain variable region comprising:

a CDR-L1 comprising QSISDF (**SEQ ID NO: 19**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising AAS (**SEQ ID NO: 20**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising QQSYIMPDT (**SEQ ID NO: 21**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

(viii) a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(ix) a heavy chain variable region comprising:

a CDR-H1 comprising GSSERFTS (**SEQ ID NO: 29**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

(x) a heavy chain variable region comprising:

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

3. The cell of any one of the preceding claims, wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region and a light chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising:

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and/or

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(ii) a heavy chain variable region comprising:

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and/or

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(iii) a heavy chain variable region comprising:

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and/or

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNRYNY (**SEQ ID NO: 17**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto

(iv) a heavy chain variable region comprising:

a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and/or

a light chain variable region comprising:

a CDR-L1 comprising QSLLHSNGYNY (**SEQ ID NO: 9**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

(v) a heavy chain variable region comprising:

a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and/or

a light chain variable region comprising:

a CDR-L1 comprising QSISDF (**SEQ ID NO: 19**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-L2 comprising AAS (**SEQ ID NO: 20**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-L3 comprising QQSYIMPDT (**SEQ ID NO: 21**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

4. The cell of any one of the preceding claims, wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region comprising:

a CDR-H1 comprising:

- GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
- GSSERFTS (**SEQ ID NO: 29**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising:

- ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
- ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

a CDR-H3 comprising:

- TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, or
- TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

5. The cell of any one of the preceding claims, wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable region selected from the group consisting of:

(i) a heavy chain variable region comprising:

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising TNGRWSGDTYYAHH (**SEQ ID NO: 27**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

(ii) a heavy chain variable region comprising:

a CDR-H1 comprising GSSERFTS (**SEQ ID NO: 29**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising ITNGGST (**SEQ ID NO: 30**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising MAGTS (**SEQ ID NO: 31**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto; and

(iii) a heavy chain variable region comprising:

a CDR-H1 comprising GSIFSGND (**SEQ ID NO: 25**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,

a CDR-H2 comprising ITSGGST (**SEQ ID NO: 26**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto, and

a CDR-H3 comprising TNGRWSGDTYYAHL (**SEQ ID NO: 33**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

6. The cell of any one of the preceding claims, wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain and/or a light chain variable domain selected from the group consisting of:

(i) a heavy chain variable domain comprising QVQLVESGGGLVQAGGSLRLSCAAS**GSIFSGND**MSWYRQAPGKGLE LVAV**ITSGGST**YYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYY C**TNGRWSGDTYYAHH**WGQGTL (**SEQ ID NO: 37**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions

(ii) a heavy chain variable domain comprising QVQLQESGGGLVQAGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE LVAV**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 35**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions

(iii) a heavy chain variable domain comprising EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLE WMGG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTAVY YC**ARSLGGRFRY**WGQGTL (**SEQ ID NO: 4**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions

(iv) a heavy chain variable domain comprising QVQLQQSGPGLVKPSQTLSLTCAIS**GDSISSNSVA**WNWIRQSPSRGL EWLGR**TYYRSKWYS**DYAISVKGRLDINPDTSKNQFSLQLNSVTPEDT AVYYC**AREVEGSSYDAFDI**WGQGTM (**SEQ ID NO: 8**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions,

(v) a light chain variable domain comprising DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQ SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**M QALQTPYT**FGQGTK (**SEQ ID NO: 12**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions

(vi) a light chain variable domain comprising EIVLTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQS PQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQ GLQSPWT**FGQGTK (**SEQ ID NO: 16**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions

(vii) a light chain variable domain comprising DVVMTQSPLSLPVTPGESASISCRSS**QSLLHSNRYNY**LDWYLQKPGQ SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**M QALQTPYT**FGQGTK (**SEQ ID NO: 18**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions

(viii) a light chain variable domain comprising DIQLTQSPSSLSASVGDRVTITCRAS**QSISDF**LNWYQQKPGKAPKLLIY **AAS**SLQTGVPSRFGGSGSGTEFTLTISSLQPEDLGTYYC**QQSYIMPDT** FGQGTK (**SEQ ID NO: 22**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions

(ix) a light chain variable domain comprising DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQ SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYC**M QGLQTPYT**FGQGTK (**SEQ ID NO: 24**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions

(x) a heavy chain variable domain comprising QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE LVA**VITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 28**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions

(xi) a heavy chain variable domain comprising QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERE LVAF**ITNGGST**RYTDPVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYY C**MAGTS**WGQGTQ (**SEQ ID NO: 32**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions

(xii) a heavy chain variable domain comprising QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE LVA**VITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY C**TNGRWSGDTYYAHL**WGQGTQ (**SEQ ID NO: 34**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions

(xiii) a heavy chain variable domain comprising QVQLQESGGGLVQAGDSLRLSCADS**GSIFSGND**MAWYRRAPGVERE LVAV**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 36**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions.

7. The cell of any one of the preceding claims, wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain and/or a light chain variable domain selected from the group consisting of:

(i) a heavy chain variable domain comprising: EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLE WMGG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTA VYYC**ARSLGGRFRY**WGQGTL (**SEQ ID NO: 4**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and/or

a light chain variable domain comprising DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKP GQSPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGV YYC**MQALQTPYT**FGQGTK (**SEQ ID NO: 12**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions

(ii) a heavy chain variable domain comprising: EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLE WMGG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTA VYYC**ARSLGGRFRY**WGQGTL (**SEQ ID NO: 4**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and/or

a light chain variable domain comprising EIVLTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPG

QSPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVY YC**MQGLQSPWT**FGQGTK (**SEQ ID NO: 16**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions

(iii) a heavy chain variable domain comprising:

EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLE WMGG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTA VYYC**ARSLGGRFRY**WGQGTL (**SEQ ID NO: 4**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and/or

a light chain variable domain comprising

DVVMTQSPLSLPVTPGESASISCRSS**QSLLHSNRYNY**LDWYLQKP GQSPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGV YYC**MQALQTPYT**FGQGTK (**SEQ ID NO: 18**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions

(iv) a heavy chain variable domain comprising:

QVQLQQSGPGLVKPSQTLSLTCAIS**GDSISSNSVA**WNWIRQSPSR GLEWLGR**TYYRSKWYS**DYAISVKGRLDINPDTSKNQFSLQLNSVT PEDTAVYYC**AREVEGSSYDAFDI**WGQGTM (**SEQ ID NO: 8**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and/or

a light chain variable domain comprising

DIQLTQSPSSLSASVGDRVTITCRAS**QSISDF**LNWYQQKPGKAPKLLIY **AAS**SLQTGVPSRFGGSGSGTEFTLTISSLQPEDLGTYYC**QQSYIM PDT**FGQGTK (**SEQ ID NO: 22**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, or

(v) a heavy chain variable domain comprising:

EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSSYA**ISWVRQAPGQGLE WMGG**IIPIFGTA**NYAQNFQGRVTMTADTSISTAYMELSSLRSEDTA VYYC**ARSLGGRFRY**WGQGTL (**SEQ ID NO: 4**) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, and/or

a light chain variable domain

DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNY**LDWYLQKPGQ SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYY C**MQGLQTPYT**FGQGTK (**SEQ ID NO: 24** or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions).

8. The cell of any one of the preceding claims, wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a single domain heavy chain variable domain having a sequence:

(i) QVQLQESGGGLVQAGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE LVAV**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 35**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions, or

(ii) QVQLVESGGGLVQAGGSLRLSCAAS**GSIFSGND**MSWYRQAPGKGLE LVAV**ITSGGST**YYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYY C**TNGRWSGDTYYAHH**WGQGTL (**SEQ ID NO: 37**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions, or

(iii) QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE LVAV**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 28**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions, or

(iv) QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERE LVAF**ITNGGST**RYTDPVKGRFTISRDNAKNTVYLQMNSLKAEDTAVYY C**MAGTS**WGQGTQ (**SEQ ID NO: 32**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions, or

(v) QVQLQESGGGLVQPGGSLRLSCADS**GSIFSGND**MAWYRRAPGVERE LVAV**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY C**TNGRWSGDTYYAHL**WGQGTQ (**SEQ ID NO: 34**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions, or

(vi) QVQLQESGGGLVQAGDSLRLSCADS**GSIFSGND**MAWYRRAPGVERE LVAV**ITSGGST**HYADSVKGRFTISRDNAQKTVYLQTNDLKPEDTAVYY C**TNGRWSGDTYYAHH**WGQGTQ (**SEQ ID NO: 36**) or a fragment, variant or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto and/or having two or three amino acids substitutions.

9. The cell of any one of the preceding claims, wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain encoded by a nucleotide sequence comprising:

(i) a heavy chain variable region comprising:

a CDR-H1 comprising GGAAGCATCTTCAGTGGCAATGAC (**SEQ ID NO: 62**),

a CDR-H2 comprising ATTACTAGCGGTGGTAGTACA (**SEQ ID NO: 63**), and

a CDR-H3 comprising ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC (**SEQ ID NO: 64**)

(ii) a heavy chain variable region comprising:

a CDR-H1 comprising GGAAGCTCCGAAAGATTCACATCA (**SEQ ID NO: 66**),

a CDR-H2 comprising ATTACTAATGGTGGTAGCACA (**SEQ ID NO: 67**), and

a CDR-H3 comprising ATGGCGGGTACGTCC (**SEQ ID NO: 68**); and

(iii) a heavy chain variable region comprising:

a CDR-H1 comprising GGAAGCATCTTCAGTGGCAATGAC (**SEQ ID NO: 62**),

a CDR-H2 comprising ATTACTAGCGGTGGTAGTACA (**SEQ ID NO: 63**), and

a CDR-H3 comprising ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCTC (**SEQ ID NO: 70**)

or fragment or variation or sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto.

10. The cell of any one of the preceding claims, wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain encoded by a nucleotide sequence comprising:

(i) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG GGGCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG TGGCAATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT ACTATGCCCATCAC**TGGGGCCAGGGGACCCAG (**SEQ ID NO: 72**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(ii) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGG GGGCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAG TGGCAATGAC**ATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACT CGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**TACTATGC AGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCaAG AAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGG CCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACT ATGCCCATCAC**TGGGGCCAGGGGACCCTG (**SEQ ID NO: 74**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions; or

(iii) GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG CAGCTATGCT**ATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG CTAC**TGGGGCCAGGGAACCCTG (**SEQ ID NO:** 41) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(iv) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC TAGTAACAGTGTTGCT**TGGAACTGGATCAGGCAGTCCCCATCGAG AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT ACAGT**GATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC AGCAGCTATGATGCTTTTGATATC**TGGGGCCAAGGGACAATG (**SEQ ID NO: 45**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions; or

(v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG TGGCAATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA CTATGCCCATCAC**TGGGGCCAGGGGACCCAG (**SEQ ID NO: 65**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG GGGGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAG ATTCACATCA**GTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGC GCGAGTTGGTCGCATTT**ATTACTAATGGTGGTAGCACA**AGATATAC

AGACCCCGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAA
GAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACAC
GGCCGTCTATTATTGT**ATGGCGGGTACGTCC**TGGGGCCAGGGGAC
CCAG (**SEQ ID NO: 69**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(vii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATG
CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCATCTCTGGGGCCAGGGGACCCAG (**SEQ ID NO: 71**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(viii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGG
CTGGGGACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCT**
TCAGTGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTG
GAGCGCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CAC
TATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAAT
GCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAG
GACACGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGA**
TACTTACTATGCCCATCACTGGGGCCAGGGGACCCAG (**SEQ ID NO: 73**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, and/or

the light chain variable domain is encoded by a nucleotide sequence comprising:

(xvi) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCAC
CCCTGGAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCT**
CCTGCATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAG
CCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGG
CCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAG
ATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGT
TTATTACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAG

GGGACCAAG (**SEQ ID NO: 49**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(xvii) GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCAC CCCTGGAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCT CCTGCATAGTAATGGATACAACTAT**TTGGATTGGTACCTGCAGAAG CCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGG CCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGTACAG ATTTTACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGT TTATTACTGC**ATGCAAGGTCTACAAAGTCCCTGGACG**TTCGGCCA AGGGACCAAG (**SEQ ID NO: 53**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(xviii) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCAC CCCTGGAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCT CCTGCATAGTAATAGATACAACTAT**TTGGATTGGTACCTGCAGAAG CCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGG CCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAG ATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGT TTATTACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAG GGGACCAAG (**SEQ ID NO: 55**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(xix) GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATC TGTAGGAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTAT TAGCGACTTT**TTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCG AAGCTCCTGATCTAT**GCTGCATCG**AGTTTACAAACTGGGGTCCCCT CAAGATTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCACCAT AAGCAGTCTACAACCTGAAGATTTGGGAACTTATTACTGT**CAACAG AGTTACATTATGCCCGACACT**TTTGGCCAGGGGACGAAA (**SEQ ID NO: 59**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(xx) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCAC CCCTGGAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCT**

CCTGCATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAG
CCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGG
CCTCCGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAG
ATTTTACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGG
TTTATTACTGC**ATGCAAGGTCTACAGACTCCGTACACT**TTTGGCCA
GGGGACCAAG (**SEQ ID NO: 61**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions.

11. The cell of any one of the preceding claims, wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises heavy chain and/or light chain variable domains encoded by nucleotide sequences selected from the group consisting of:

(vi) a heavy chain variable domain encoded by the nucleotide sequence comprising:

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 49**) or a sequence at least 80%, 85%, 90%, 95%,

96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions

(vii)a heavy chain variable domain encoded by the nucleotide sequence comprising:

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGTACAGATTT
TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGGTCTACAAAGTCCCTGGACG**TTCGGCCAAGGG
ACCAAG (**SEQ ID NO: 53**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions

(viii) a heavy chain variable domain encoded by the nucleotide sequence comprising:

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**

CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATAGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGG
ACCAAG (**SEQ ID NO: 55**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions

(ix) a heavy chain variable domain encoded by the nucleotide sequence comprising:

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTAT**TTGGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCAGTGGCAGTGGATCAGGCACAGATTT
TACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA

TTACTGC**ATGCAAGGTCTACAGACTCCGTACACT**TTTGGCCAGGG GACCAAG (**SEQ ID NO: 61**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions; and

(x) a heavy chain variable domain encoded by the nucleotide sequence comprising:

CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC TAGTAACAGTGTTGCT**TGGAACTGGATCAGGCAGTCCCCATCGAG AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT ACAGT**GATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC AGCAGCTATGATGCTTTTGATATC**TGGGGCCAAGGGACAATG (**SEQ ID NO: 45**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, and/or

a light chain variable domain encoded by the nucleotide sequence comprising:

GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG ACTTT**TTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT CCTGATCTAT**GCTGCATCG**AGTTTACAAACTGGGGTCCCCTCAAGA TTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCACCATAAGCA GTCTACAACCTGAAGATTTGGGAACTTATTACTGT**CAACAGAGTTA CATTATGCCCGACACT**TTTGGCCAGGGGACGAAA (**SEQ ID NO: 59**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions.

12. The cell of any one of the preceding claims, wherein the first antigen binding protein, variant or binding fragment thereof that binds to EpCAM comprises a heavy chain variable domain encoded by the nucleotide sequence comprising:

(i) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGG GCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAGTGGCA ATGAC**ATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACTCGAGTTG

GTCGCGGTT**ATTACTAGCGGTGGTAGTACA**TACTATGCAGACTCCGTG
AAGGGCCGATTCACCATCTCCAGAGACAATTCCaAGAAcACCcTATATC
TGCAAATGAACAGCCTGAGAGCTGAGGACACGGCCGTGTATTACTGC**A**
CAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCACTGGGGC
CAGGGGACCCTG (**SEQ ID NO: 74**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions; or

(ii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGG
GCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG
TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC**TGG
GGCCAGGGGACCCAG (**SEQ ID NO: 72**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(iii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAG
GGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG
TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACTTGAAACCTGAGGACACGGCCGTGTATTACT
GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC**TGG
GGCCAGGGGACCCAG (**SEQ ID NO: 65**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(iv) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGGG
GGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAGATTC**
ACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGCGCGAGTT
GGTCGCATTT**ATTACTAATGGTGGTAGCACA**AGATATACAGACCCCGT
GAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGAACACGGTGT
ATCTGCAAATGAACAGCCTGAAAGCTGAGGACACGGCCGTCTATTATT
GT**ATGGCGGGTACGTCC**TGGGGCCAGGGGACCCAG (**SEQ ID NO: 69**)

or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAG GGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA ATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCTC**TGG GGCCAGGGGACCCAG (**SEQ ID NO: 71**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGG ACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA ATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT GGTCGCGGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT GC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC**TGG GGCCAGGGGACCCAG (**SEQ ID NO: 73**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGTC CTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAGCAGCT ATGCT**ATCAGCTGGGTGCGACAGGCCCCTGGACAAGGGCTTGAGTGG ATGGGAGGG**ATCATCCCTATCTTTGGTACAGCA**AACTACGCACAGAAC TTCCAGGGCAGAGTCACCATGACCGCAGACACCTCCATAAGCACAGC CTACATGGAGCTGAGCAGCCTGAGATCTGAGGACACGGCCGTGTATTA CTGT**GCGAGATCGTTGGGTGGGAGATTTCGCTAC**TGGGGCCAGGGAA CCCTG (**SEQ ID NO:** 41) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions, or

(vii) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTCGCA GACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTCTAGTAA CAGTGTTGCT**TGGAACTGGATCAGGCAGTCCCCATCGAGAGGCCTTG

AGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGTACAGT**GATTAT
GCAATATCTGTGAAAGGTCGATTAGACATCAACCCAGACACATCCAAG
AACCAGTTCTCCCTGCAGCTGAACTCTGTGACTCCCGAGGACACGGCT
GTGTATTATTGT**GCAAGAGAAGTTGAGGGCAGCAGCTATGATGCTTTT**
GATATCTGGGGCCAAGGGACAATG (**SEQ ID NO: 45**) or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-20 nucleic acid substitutions.

13. The cell of any one of the preceding claims, wherein the multi-specific antigen binding protein, variant or fragment thereof is a bispecific antibody.

14. The cell of any one of the preceding claims, wherein the multi-specific antigen binding protein, variant or fragment thereof is an immune cell engager selected from the group comprising a T cell engager, an NK cell engager, a monocyte engager and a macrophage engager.

15. The cell of any one of the preceding claims, wherein the multi-specific antigen binding protein, variant or fragment thereof is a bispecific T cell engager (BiTE), such as an inducible BiTE, a non-inducible BiTE or a constitutive expression BiTE.

16. The cell of any one of the preceding claims, wherein the second antigen binding protein, variant or binding fragment thereof binds to the immune marker selected from the group consisting of CD3, NKG2D, CD4, CD8, CD16, and CD64.

17. The cell of any one of the preceding claims, wherein the multispecific antigen binding protein is an inducible bispecific T cell engager comprising a Heavy chain antibody variable region (i.e. VHH) or a single chain variable fragment (scFv).

18. The cell of any one of the preceding claims, wherein the cell is selected from the group consisting of a T cell, a macrophage, a monocyte, and an NK cell.

19. The cell of any one of the preceding claims, wherein the cell is a T cell, optionally a CAR T-cell.

20. The cell of any one of the preceding claims, wherein the cell

i. binds to GPC3 and secretes an inducible bispecific T cell engager that targets EpCAM and CD3 (GE CAR-BiTE T); or

ii. binds to CD19 and secretes an inducible bispecific T cell engager that targets EpCAM and CD3 (CD19 CAR-BiTE T).

21. A polynucleotide encoding the cell of any one of the preceding claims.

22. A vector expressing the polynucleotide of claim 21.

23. A host cell comprising the vector of claim 22.

24. A method of producing / generating the cell of any one of claims 1 to 20, comprising introducing the polynucleotide of claim 21 into the cell.

25. A composition comprising the cell of any one of claims 1 to 20.

26. A method of treating a disease in a subject in need thereof, the method comprises administering to the subject the cell of any one of claims 1 to 20 or the composition of claim 25, optionally the disease is a proliferative disease, optionally a cancer.