



(51) International Patent Classification:

C12N 5/10 (2006.01) A61K 35/17 (2015.01)

C07K 14/275 (2006.01) A61P 35/00 (2006.01)

C07K 16/28 (2006.01)

(21) International Application Number:

PCT/SG2024/050384

(22) International Filing Date:

07 June 2024 (07.06.2024)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

10202301606S 07 June 2023 (07.06.2023) SG

(71) Applicant: AGENCY FOR SCIENCE, TECHNOLOGY AND RESEARCH [SG/SG]; 1 Fusionopolis Way, #20-10 Connexis North Tower, Singapore 138632 (SG).

(72) Inventors: WANG, Cheng-I; 8A Biomedical Grove, #03-06 Immunos, Singapore 138648 (SG). LEE, Wen-Hsin; 8A Biomedical Grove, #03-06 Immunos, Singapore 138648 (SG). WANG, Bei; 8A Biomedical Grove, #03-06 Immunos, Singapore 138648 (SG). YEAP, Yee Chin, Yvonne; 8A Biomedical Grove, #03-06 Immunos, Singapore 138648 (SG). NGOH, Zi Xian Eve; 8A Biomedical Grove, #03-06 Immunos, Singapore 138648 (SG).

(74) Agent: TECK YEOW, Lim; Donaldson & Burkinshaw LLP, P.O. Box 3667 Robinson Road Post Office, Singapore 905667 (SG).

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

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

Declarations under Rule 4.17:

— of inventorship (Rule 4.17(iv))

(54) Title: HER2 CHIMERIC ANTIGEN RECEPTOR SECRETING

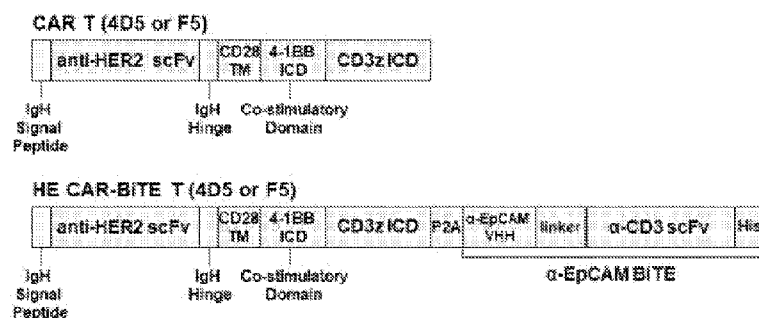


Figure 1A

(57) Abstract: There is provided a modified cell expressing (a) a chimeric antigen receptor targeting HER2, 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 comprise sequences as disclosed herein. Also disclosed are methods of producing/generating the cell as disclosed herein and methods of treatment employing the cell as disclosed herein.

Published:

- *with international search report (Art. 21(3))*
- *with sequence listing part of description (Rule 5.2(a))*

HER2 CHIMERIC ANTIGEN RECEPTOR SECRETING

TECHNICAL FIELD

5 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 cell marker.

BACKGROUND

15 The cancer incidence worldwide is expected to rise to 28.4 million cases in 2040. Among all adult human cancer cases, approximately 90% are solid tumours. Solid tumours 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, gastric adenocarcinoma accounts for 90-95% cancers of the stomach. 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%), liver (~85%), 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.

25 Compared to conventional treatment strategies, antibody and cell-based therapies are gaining increasingly 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

35

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. To overcome these challenges, there is urgent medical need to explore next-generation CAR T cell therapies to prevent cancer relapse.

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. Hence, 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

EpCAM is a well-known tumour associated antigen that is frequently over-expressed in almost all solid tumours originated from epithelium, however, neither CAR T cells nor systemically administered BiTEs targeting EpCAM 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.

The present disclosure demonstrated that HER2 targeting CAR T cells secreting anti-EpCAM BiTE (named as "HE CAR-BiTE T") exerted superior efficacy in killing human tumours originating from gastric carcinomas *in vitro*. In summary, the present inventors approach took 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 treatment. The present disclosure also provides for the possibility of expanding the anti-EpCAM BiTE secreting CAR T cell therapy to other cancer types to establish a broader scope of next-generation CAR T cell therapies.

5

In one aspect, there is provided a modified cell expressing

(a) a chimeric antigen receptor targeting HER2, 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,

10

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:

15

(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)

20

a CDR-H2 comprising:

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

25

a CDR-H3 comprising:

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

30

(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 IIPFGTA (SEQ ID NO: 2), and

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

35

(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:

- QSLHNSNGYNY (**SEQ ID NO: 9** - 1B6, 1C1, 1C11 and 1H6),
- QSLHNSRYNY (**SEQ ID NO: 17** - 1D4), or
- QSIDF (**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, 2B7, and 2D10),
- and 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)

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

10 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

15 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)

20 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:

- 25 (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)

30 a CDR-H1 comprising GGTFFSSYA (SEQ ID NO: 1),

a CDR-H2 comprising IIPFGTA (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),

35 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 QSLHNSNGYNY (SEQ ID NO: 9),
- a CDR-L2 comprising LGS (SEQ ID NO: 10), and
- 5 a CDR-L3 comprising MQALQTPYT (SEQ ID NO: 11)
- (v) a light chain variable region comprising: (1C11)
- a CDR-L1 comprising QSLHNSNGYNY (SEQ ID NO: 9),
- a CDR-L2 comprising LGS (SEQ ID NO: 10), and
- a CDR-L3 comprising MQGLQSPWT (SEQ ID NO: 15)
- 10 (vi) a light chain variable region comprising: (1D4)
- a CDR-L1 comprising QSLHNSRYNY (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)
- 15 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 QSLHNSNGYNY (SEQ ID NO: 9),
- 20 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
- 25 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)
- 30 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
 QVQLVESGGGLVQAGGSLRLSCAAS**GSIFSGND**MSWYRQAPGKGLE
 LVAVIT**SGGST**YYADSVKGRFTISRDN SKNTLYLQMNSLRAEDTAVYY
 CTNGRWSGDTYYAHHWGQGT**L** (**SEQ ID NO: 37** - hu2C4-VHH)
- 5 (ii) a heavy chain variable domain comprising
 QVQLQESGGGLVQAGGSLRLSCAD**SGSIFSGND**MAWYRRAPGVERE
 LVAVIT**SGGST**HYADSVKGRFTISRDN AQKTVYLQTN DLKPEDTAVYY
 CTNGRWSGDTYYAHHWGQGT**Q** (**SEQ ID NO: 35** - 2C4-VHH)
- 10 (iii) a heavy chain variable domain comprising
 EVQLVQSGAEVKKPGSSVKVSC**KASGGTFSS**YAISWVRQAPGQGLE
 WMGGI**PIFGT**ANYAQN FQGRVTMTADTSISTAYMELSSLRSEDTAVY
 YCARSLGGRFRYWGQGT**L** (**SEQ ID NO: 4** - 1B6, 1C1, 1C11, 1D4
 and 1H6)
- 15 (iv) a heavy chain variable domain comprising
 QVQLQQSGPGLVKPSQTL**SLTCAISGDSISSNSVA**WNWIRQSPSRGL
 EWLGR**TYR**SKWYSDYAI SVKGRLDINPDTSKNQFSLQLNSVTPEDT
 AVYYCAREVEGSSYDAFDIWGQGT**M** (**SEQ ID NO: 8** - 1E4),
- 20 (v) a light chain variable domain comprising
 DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHS**NGYNYLDWYLQKPGQ
 SPQLLIY**LGS**NRASGV PDRFSGSGSGTDFTLKISRVEAEDVGVYY**CM**
QALQTPYTFGQGTK (**SEQ ID NO: 12** - 1B6 and 1C1)
- 25 (vi) a light chain variable domain comprising
 EIVLTQSPLSLPVTPGEPASISCRSS**QSLLHS**NGYNYLDWYLQKPGQS
 PQLLIY**LGS**NRASGV PDRFSGSGSGTDFTLKISRVEAEDVGVYY**CMQ**
GLQSPWTFGQGTK (**SEQ ID NO: 16** - 1C11)
- 30 (vii) a light chain variable domain comprising
 DVVMTQSPLSLPVTPGESASISCRSS**QSLLHS**NRNYLDWYLQKPGQ
 SPQLLIY**LGS**NRASGV PDRFSGSGSGTDFTLKISRVEAEDVGVYY**CM**
QALQTPYTFGQGTK (**SEQ ID NO: 18** - 1D4)
- 35 (viii) a light chain variable domain comprising
 DIQLTQSPSSLSASVGDRVITITCRAS**QSID**FLN WYQQKPGKAPKLLIY
AASSLQTGVPSRFGSGSGTEFTLTISLQPEDLGTY**YCCQSY**IMPDT
 FGQGTK (**SEQ ID NO: 22** - 1E4)
- (ix) a light chain variable domain comprising
 DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHS**NGYNYLDWYLQKPGQ

SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYCM
QGLQTPYTFGQGTK (SEQ ID NO: 24 - 1H6)

- (x) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCADSG**SIFSG**NDMAWYRRAPGVERE
 5 LVAVIT**SGGST**HYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYY
 CTNGRW**SGD**TY**AHHW**GQGTQ (**SEQ ID NO: 28- 1A5-VHH**)

- (xi) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCAASG**SSE**RFTSVAWYRQAPGKERE
 10 LVAFIT**NGGST**RYTDPVKGRFTISRDNAAKNTVYLMNSLKAEDTAVYY
 CMAGTSW**GQGTQ (SEQ ID NO: 32 - 1B8-VHH)**

- (xii) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCADSG**SIFSG**NDMAWYRRAPGVERE
 15 LVAVIT**SGGST**HYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYY
 CTNGRW**SGD**TY**AHLW**GQGTQ (**SEQ ID NO: 34 - 2B7-VHH**)

- (xiii) a heavy chain variable domain comprising
 QVQLQESGGGLVQAGDSLRLSCADSG**SIFSG**NDMAWYRRAPGVERE
 20 LVAVIT**SGGST**HYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYY
 CTNGRW**SGD**TY**AHHW**GQGTQ (**SEQ ID NO: 36 - 2D10-VHH**); and

or fragment or variant or sequence at least 80%, 85%, 90%, 95%, 96%, 97%,
 20 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) QVQLQESGGGLVQAGGSLRLSCADSG**SIFSG**NDMAWYRRAPGVERE
 25 LVAVIT**SGGST**HYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYY
 CTNGRW**SGD**TY**AHHW**GQGTQ (**SEQ ID NO: 35 - clone 2C4-VHH**),
 or

- (ii) QVQLVESGGGLVQAGGSLRLSCAASG**SIFSG**NDM**SWY**RQAPGKGLE
 30 LVAVIT**SGGST**TYADSVKGRFTISRDN**SKNT**LYLMNSLRAEDTAVYY
 CTNGRW**SGD**TY**AHHW**GQGT**L (SEQ ID NO: 37 - hu2C4-VHH)**, or

- (iii) QVQLQESGGGLVQPGGSLRLSCADSG**SIFSG**NDMAWYRRAPGVERE
 LVAVIT**SGGST**HYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYY
 CTNGRW**SGD**TY**AHHW**GQGTQ (**SEQ ID NO: 28 - clone 1A5-VHH**), or

(iv) QVQLQESGGGLVQPGGSLRLSCAASGSSERFTSWAWYRQAPGKERE
 LVAFITNGGSTRYTDVPVKGRFTISRDNKNTVYLQMNSLKAEDTAVYY
 CMAGTSWGQGTQ (SEQ ID NO: 32 - clone 1B8-VHH), or

(v) QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 5 LVAVITSGGSTHYADSVKGRFTISRDNKNTVYLQTNDLKPEDTAVYY
 CTNGRWSGDTYYAHLWGQGTQ (SEQ ID NO: 34 - clone 2B7-VHH), or

(vi) QVQLQESGGGLVQAGDSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNKNTVYLQTNDLKPEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 36 - clone 2D10-VHH),

10 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
 15 group consisting of:

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

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

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

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

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

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

a CDR-H2 comprising TYRSKWYS (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
 25 and 1H6)

a CDR-L1 comprising:

- QSLLHSNGYNY (SEQ ID NO: 9 - 1B6, 1C1, 1C11 and 1H6),

- QSLLHSNRYNY (SEQ ID NO: 17 - 1D4), or

- QSIDF (SEQ ID NO: 19 - 1E4)

30 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),

35 • 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.

5 In some examples, 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 GGTFSYA (SEQ ID NO: 1),

10 a CDR-H2 comprising IPIFGTA (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 QSLHSNGYNY (SEQ ID NO: 9),

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

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

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

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

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

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

20 a light chain variable region comprising:

a CDR-L1 comprising QSLHSNGYNY (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)

25 a CDR-H1 comprising GGTFSYA (SEQ ID NO: 1),

a CDR-H2 comprising IPIFGTA (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 QSLHSNRYNY (SEQ ID NO: 17),

30 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 GGTFSYA (SEQ ID NO: 1),

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

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

a light chain variable region comprising:

- a CDR-L1 comprising QSLHNSGYN (SEQ ID NO: 9),
- a CDR-L2 comprising LGS (SEQ ID NO: 10), and
- a CDR-L3 comprising MQGLQTPYT (SEQ ID NO: 23); and

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

- a CDR-H1 comprising GDSISSNSVA (SEQ ID NO: 5),
- a CDR-H2 comprising TYRSKWYS (SEQ ID NO: 6), and
- a CDR-H3 comprising AREVEGSSYDAFDI (SEQ ID NO: 7); and

a light chain variable region comprising: (1E4)

- 10 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.

15 In some examples, 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)

20 EVQLVQSGAEVKKPGSSVKVCKASGGTFSSYAISWVRQAPGQGLEWM
GGIPIFGTANYAQNFGGRVTMTADTSISTAYMELSSLRSED TAVYYCARS
LGGRFRYWGGT (SEQ ID NO: 4), and

a light chain variable domain comprising

25 DVVMTQSPLSLPVTPGEPASISCRSSQSLHNSGYNLDWYLQKPGQSP
QLLIYLGSRASGVDPDRFSGSGSDFTLKISRVEAEDVGVYYCMQALQT
PYTFGQGTK (SEQ ID NO: 12)

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

30 EVQLVQSGAEVKKPGSSVKVCKASGGTFSSYAISWVRQAPGQGLEWM
GGIPIFGTANYAQNFGGRVTMTADTSISTAYMELSSLRSED TAVYYCARS
LGGRFRYWGGT (SEQ ID NO: 4), and

a light chain variable domain comprising

EIVLTQSPLSLPVTPGEPASISCRSSQSLHNSGYNLDWYLQKPGQSPQ
LLIYLGSRASGVDPDRFSGSGSDFTLKISRVEAEDVGVYYCMQGLQSP
WTFGQGTK (SEQ ID NO: 16)

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

EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWM
GGIIPFGTANYAQNFGGRVTMTADTSISTAYMELSSLRSED TAVYYCARS
LGGRFRYWGGTGL (SEQ ID NO: 4), and

a light chain variable domain comprising

5 DVVMTQSPLSLPVTPGESASISCRSSQSLHNSRNYLDWYLQKPGQSP
QLLIYLGSNRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQT
PYTFGQGTK (SEQ ID NO: 18)

(iv) a heavy chain variable domain comprising: (1E4)

10 QVQLQQSGPGLVKPSQTLSTCAISGDSISSNSVAWNWIRQSPSR
GLEWLGRTYYRSKWYSDYAI SVKGRLDINPDTSKNQFSLQLNSVT
PEDTAVYYCAREVEGSSYDAFDIWGGTGM (SEQ ID NO: 8), and

a light chain variable domain comprising

15 DIQLTQSPSSLSASVGDRVTITCRASQSISDFLNWYQQKPGKAPKLLIYAA
SSLQTGVPSRFGSGSGTEFTLTISLQPEDLGTYYCQQSYIMPDTFGQG
TK (SEQ ID NO: 22), or

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

20 EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWM
GGIIPFGTANYAQNFGGRVTMTADTSISTAYMELSSLRSED TAVYYCARS
LGGRFRYWGGTGL (SEQ ID NO: 4), and

a light chain variable domain

25 DVVMTQSPLSLPVTPGEPASISCRSSQSLHNSNGYNYLDWYLQKPGQSP
QLLIYLGSNRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYCMQGLQT
PYTFGQGTK (SEQ ID NO: 24), or

or fragment or variant or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%,
25 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:

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

(ii) RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
35 NSQESVTEQDSKDSTYSLSTLTLSKADYEKHKVYACEVTHQGLSSPVTK
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 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**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
10 GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACACACTATG**
CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCATCACTGGGGCCAGGGGACCCAG (SEQ ID NO: 72 -
15 clone 2C4), or
- (ii) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGG
GGGCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAG**
TGGCAATGACATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACT
CGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACATACTATGC**
20 AGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCaAG
AAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGG
CCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACT**
ATGCCCATCACTGGGGCCAGGGGACCCTG (SEQ ID NO: 74 - clone
hu2C4-VHH); or
- (iii) GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAA**ACTA
CGCACAGAACTTCAGGGCAGAGTCACCATGACCGCAGACACCTC
30 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTTCG**
CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41 - for clones 1B6,
1C1, 1C11, 1D4, 1H6), or
- (iv) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
GCAGACCCTCTCACTCACCTGTGCCATCTCCGGGGACAGTATCTC

TAGTAACAGTGTTGCTTGGAACTGGATCAGGCAGTCCCCATCGAG
 AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
ACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC
 AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
 5 CCGGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG
 (SEQ ID NO: 45 - clone 1E4); or

(v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
 10 **TGGCAATGACAT**GGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
 GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ACACTATG
 CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC
 AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC
 GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA**
 15 **CTATGCCCATCACT**GGGGGCCAGGGGACCCAG (SEQ ID NO: 65 -
 clone 1A5), or

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

(vii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
 GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ACACTATG
 CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC
 30 AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
 CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
ACTATGCCCATCTCTGGGGGCCAGGGGACCCAG (SEQ ID NO: 71 -
 clone 2B7), or

(viii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGG
 35 CTGGGGACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCT**

5 **TCAGTGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTG**
GAGCGCGAGTTGGTCGCGGTTATTACTAGCGGTGGTAGTACACAC
TATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAAT
GCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAG
GACACGGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCCGA
TACTTACTATGCCCATCACTGGGGCCAGGGGACCCAG (SEQ ID
NO: 73 - clone 2D10), and/or

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

- 10 (i) **GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG**
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGTCAGAGCCTCCTGC
ATAGTAATGGATACAAC TATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
15 **TACTGCATGCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG**
ACCAAG (SEQ ID NO: 49 - clone 1B6 and 1C1), or
- (ii) **GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG**
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGTCAGAGCCTCCTGC
ATAGTAATGGATACAAC TATTTGGATTGGTACCTGCAGAAGCCAG
20 **GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC**
CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGTACAGATTT
TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCATGCAAGGTCTACAAAGTCCCTGGACGTTTCGGCCAAGGG
ACCAAG (SEQ ID NO: 53 - clone 1C11), or
- 25 (iii) **GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG**
GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGTCAGAGCCTCCTGC
ATAGTAATAGATACAAC TATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATTT
30 **TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT**
TACTGCATGCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG
ACCAAG (SEQ ID NO: 55 - clone 1D4), or
- (iv) **GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG**
GAGACAGAGTCACCATCACTTGCCGGGCAAGTCAGAGTATTAGCG
35 **ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT**

- 5 CCTGATCTAT**GCTGCATCG**AGTTTACAACTGGGGTCCCCTCAAGA
 TTCGGTGGCAGTGGATCTGGGACAGAATTCACCTCTACCATAAGCA
 GTCTACAACCTGAAGATTTGGGAACCTATTACTGT**CAACAGAGTTA**
CATTATGCCCCGACACTTTTGGCCAGGGGACGAAA (SEQ ID NO: 59 -
 clone 1E4), or
- (v) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 10 CGGGTCCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATT
 TAACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
 TTACTGCATG**CAAGGTCTACAGACTCCGTACACT**TTTGGCCAGGG
 GACCAAG (SEQ ID NO: 61 - clone 1H6),
- or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
 15 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 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
 20 comprising: (1B6 and 1C1)
- GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGGAT**CATCCCTATCTTTGGTACAGCAA**ACTA
 25 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41), and
- a light chain variable domain encoded by the nucleotide sequence
 30 comprising:
- GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 35 CGGGTCCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATT

TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCATGCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG
ACCAAG (SEQ ID NO: 49)

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

5 GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAG
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
TGAGTGGATGGGAGGGATCATCCCTATCTTTGGTACAGCAAACTA
10 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGTGCGAGATCGTTGGGTGGGAGATTTCG
CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41), and

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

15 GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGTCAGAGCCTCCTGC
ATAGTAATGGATACAACATTTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
20 CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGTACAGATTT
TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCATGCAAGGTCTACAAAGTCCCTGGACGTTCCGGCCAAGGG
ACCAAG (SEQ ID NO: 53)

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

25 GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAG
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
TGAGTGGATGGGAGGGATCATCCCTATCTTTGGTACAGCAAACTA
30 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGTGCGAGATCGTTGGGTGGGAGATTTCG
CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41), and

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

35

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATAGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 5 GGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCT**AATCGGGCCTC
 CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATT
 TAACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
 TACTGCAT**GCAAGCTCTACAACTCCGTACACT**TTTGGCCAGGGG
 ACCAAG (**SEQ ID NO: 55**)

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

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAA**CTA
 15 CGCACAGAACTTCAGGGCAGAGTCACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**), and

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

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCT**AATCGGGCCTC
 25 CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATT
 TAACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
 TTA**CTGCATGCAAGGTCTACAGACTCCGTACACT**TTTGGCCAGGG
 GACCAAG (**SEQ ID NO: 61**); and

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

CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
 GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC**
TAGTAACAGTGTTGCTTGGAACTGGATCAGGCAGTCCCCATCGAG
 AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
 35 **ACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC**

AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
 CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG
 (SEQ ID NO: 45), and

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

GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
 GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT
 10 CCTGATCTAT**GCTGCATCG**AGTTTACAACTGGGGTCCCCCTCAAGA
 TTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCACCATAAGCA
 GTCTACAACCTGAAGATTTGGGAAGTTACTGT**CAACAGAGTTA**
CATTATGCCCCGACACTTTTGGCCAGGGGACGAAA (SEQ ID NO: 59)

or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100%
 15 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
 20 GCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
 GGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ACTATGCAGACTCCG
 TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
 ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
 25 **GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCACTGG**
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
 30 GCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACTCGAGTTG
 GTCGCGGTT**ATTACTAGCGGTGGTAGTAC**A TACTATGCAGACTCCGTG
 AAGGGCCGATTCACCATCTCCAGAGACAATTCCaAGAAcACCCtATATC
 TGCAAATGAACAGCCTGAGAGCTGAGGACACGGCCGTGTATTACTGCA
 35 **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

(iii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAG
 5 GGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
 GGTGCGCGTT**ATTACTAGCGGTGGTAGTACAC**ACTATGCAGACTCCG
 TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
 ATCTGCAAACGAACGACTTGAAACCTGAGGACACGGCCGTGTATTACT
 10 **GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCACTGG**
 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
 15 GGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAGATTC**
ACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGCGCGAGTT
 GGTGCGATTT**ATTACTAATGGTGGTAGCACA**AGATATACAGACCCCGT
 GAAGGGCCGATTCACCATCTCCAGAGACAACGCCAAGAACACGGTGT
 ATCTGCAAATGAACAGCCTGAAAGCTGAGGACACGGCCGTCTATTATT
 20 **GTATGGCGGGTACGTCTCTGGGGCCAGGGGACCCAG (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**
 25 **ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT**
 GGTGCGCGTT**ATTACTAGCGGTGGTAGTACAC**ACTATGCAGACTCCG
 TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
 ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCTCTGG
 30 **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**
 35 **ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT**

GGTCGCGGTTATTACTAGCGGTGGTAGTACACACTATGCAGACTCCG
TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCACTGG
5 **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
CTCGGTGAAGGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAGCAGCT
10 **ATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCTTGAGTGG**
ATGGGAGGGATCATCCCTATCTTTGGTACAGCAAACCTACGCACAGAAC
TTCCAGGGCAGAGTCACCATGACCGCAGACACCTCCATAAGCACAGC
CTACATGGAGCTGAGCAGCCTGAGATCTGAGGACACGGCCGTGTATTA
CTGT**GCGAGATCGTTGGGTGGGAGATTTGCTACTGGGGCCAGGGAA**
15 **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
GACCTCTCACTCACCTGTGCCATCTCCGGGGACAGTATCTCTAGTAA
20 **CAGTGTTGCTTGGA**ACTGGATCAGGCAGTCCCCATCGAGAGGCCTTG
AGTGGCTGGGAAGGACATACTACAGGTCCAAGTGGTACAGTGATTAT
GCAATATCTGTGAAAGGTGCGATTAGACATCAACCCAGACACATCCAAG
AACCAGTTCTCCCTGCAGCTGAACTCTGTGACTCCCGAGGACACGGCT
GTGTATTATTGT**GCAAGAGAAGTTGAGGGCAGCAGCTATGATGCTTTT**
25 **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 examples, the multi-specific antigen binding protein, variant or fragment thereof is a bispecific antibody.

30 In some examples, 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 examples, 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
35 BiTE or a constitutive expression BiTE.

In some examples, 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 examples, wherein the multi-specific antigen binding protein is an
5 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 cell is an immune cell for example selected from the group consisting of a T cell, a macrophage, a monocyte, and an NK cell.

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

10

In some examples, the cell binds to HER2 and secretes an inducible bispecific T cell engager that targets EpCAM and CD3 (HE CAR-BiTE T).

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

15 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.

20 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.

25 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

30 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
35 (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

5 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

10 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

15 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 are identical except for possible mutations (e.g., naturally occurring

20 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

25 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

30 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 four Fvs and four inter-Fv

35 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 (CH1) 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.

10 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
15 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
20 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
25 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
30 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
35 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 xenomice 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 as disclosed herein may comprise 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 5 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 10 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 15 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 seven, or eight, or nine, or ten, or fifteen, or twenty amino acid mutations with respect 20 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.

25 "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) 30 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, sarcosine, 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 antigen binding protein" as used herein, refers to a multi-specific antigen binding protein that may bind to two or more host cell target antigens. Therefore, 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
5 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
10 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
15 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.

20 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
25 "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
30 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
35 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
5 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
10 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
15 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
20 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.

25 "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
30 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

10

The present disclosure provides engineered cells expressing
(a) a chimeric antigen receptor targeting a first target antigen, and
(b) a multi-specific antigen binding protein that binds to EpCAM (epithelial cell adhesion molecule).

15

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.

In some examples, the first target antigen is an antigen associated with a disease.

20

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 first target antigen is a diseased cell and the like.

25

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. In some examples, 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, lung cancer, stomach cancer (such as gastric adenocarcinoma), breast cancer, skin cancer, ovary cancer, kidney cancer, pancreas cancer, head and neck cancer, prostate gland cancer, esophagus cancer, bladder cancer, colon cancer, 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 tumour cell lines such as but is not limited to gastric

35

adenocarcinoma cell line (such as AGS), breast cancer cell line (such as MDA-MB468, MCF7, SKBR3, BT474, MDA-MB231, MDA-MB436), glioblastoma (A-172), and the like.

In some examples, target antigen is an epithelial marker. In some examples, the epithelial markers 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 (FR α), and the like. In some examples, the first antigen is HER2. Therefore, in some embodiments, the chimeric antigen receptor binds to HER2.

In some examples, HER2 expressing cancer may include but is not limited to, gastric / gastroesophageal cancer, breast cancer, head and neck cancer, ovarian cancer, endometrium cancer, bladder cancer, lung cancer, colon cancer, and the like.

In some examples, the CAR T cell targets HER 2 alone (HE CAR-T).

In some examples, the cell expresses a multi-specific antigen binding protein that binds to one or more markers. For example, the multi-specific antigen binding protein may bind to EpCAM and an immune marker.

Without wishing to be bound by theory, 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) showed dose-limiting toxicities and failed to achieve FDA approval.

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.

Amplification or overexpression of HER2 has been found in approximately 10–30% of gastric/gastroesophageal cancers and 15–30% of breast cancers. In addition, HER2 overexpression has also been reported in other cancers like ovary, endometrium, bladder, lung, colon, as well as head and neck cancers.

Immunotherapies such as CAR T cell therapies targeting HER2 has clinically proven to have some treatment efficacy and be generally safe, unless antibody clones with very high antigen binding affinities were used. 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 HER2, 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 tumours, 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.

Interestingly, as a broad-spectrum pan marker for epithelial cells, EpCAM is found to be co-expressed with HER2 in many cancers. Instead of using dual CAR setting, a treatment that HER2- targeting CAR T cells secreting anti-EpCAM bi-specific T cell engager (BiTE) was developed in this invention 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 “HE CAR-BiTE T” (HER2 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 proximity of tumours and at the same time being activated and subsequently help the clearance of tumour cells. Since HER2 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-HER2 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.

In the experimental data of the present disclosure, it is shown that EpCAM has the highest expression level as compared to other epithelial tumour markers in the multiple HER2 positive tumour cells that were stained. Therefore, EpCAM is chosen as the BiTE target due to its wide coverage of cancer cells of epithelial origin. The much higher expression of EpCAM on cancerous cells compared to normal epithelial cells will provide a wider therapeutic window as compared to other targets. 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.

Therefore, in one aspect, there is provided a modified cell expressing

- (a) a chimeric antigen receptor targeting HER2, 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:
- 5 • ITSGGST (**SEQ ID NO: 26** – 2C4, hu2C4, 1A5, 2B7, and 2D10), or
- ITNGGST (**SEQ ID NO: 30** – 1B8) and
- a CDR-H3 comprising:
- 10 • TNGRWSGDTYYAHH (**SEQ ID NO: 27** –2C4, hu2C4, 1A5, 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**),
- 15 a CDR-H2 comprising IPIFGTA (**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
- 20 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:
- 25 • QSLHNSNGYNY (**SEQ ID NO: 9** - 1B6, 1C1, 1C11 and 1H6),
- QSLHNSNRYNY (**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
- 30 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,

5 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)

10 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:

- 15 • TNGRWSGDTYYAHH (SEQ ID NO: 27 – 2C4, hu2C4, 1A5, and 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%,
20 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,
25 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)

30 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),

35 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
5 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
 - 10 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**)
- 15 (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)
 - 20 a CDR-L1 comprising QSLHNSNGYNY (**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)
 - 25 a CDR-L1 comprising QSLHNSNGYNY (**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)
 - 30 a CDR-L1 comprising QSLHNSNRYNY (**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**)
- 35 (viii) a light chain variable region comprising: (1H6)

- a CDR-L1 comprising QSLHNSNGYNY (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)
 5 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)
 10 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
 15 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
 QVQLQESGGGLVQAGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNQKTVYLQTNDLKPEDTAVYY
 20 CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 35 - 2C4-VHH)
- (ii) a heavy chain variable domain comprising
 QVQLVESGGGLVQAGGSLRLSCAASGSIFSGNDMSWYRQAPGKGLE
 LVAVITSGGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 37 - hu2C4-VHH)
- (iii) a heavy chain variable domain comprising
 25 EVQLVQSGAEVKKPGSSVKVSKASGGTFSSYAISWVRQAPGQGLE
 WMGGIPIFGTANYAQNFGQGRVTMTADTSISTAYMELSSLRSEDVAVY
 YCARSLGGRFRYWQGTQ (SEQ ID NO: 4 - 1B6, 1C1, 1C11, 1D4
 and 1H6)
- (iv) a heavy chain variable domain comprising
 30 QVQLQQSGPGLVKPSQTLSTCAISGDSISSNSVAWNWIRQSPSRGL
 EWLGRITYYRSKWYSDYAIKGRDLINPDTSKNQFSLQLNSVTPEDT
 AVYYCAREVEGSSYDAFDIWGQGTQ (SEQ ID NO: 8 - 1E4),
- (v) a light chain variable domain comprising
 35 DVVMTQSPLSLPVTPGEPASISCRSSQSLHNSNGYNYLDWYLQKPGQ

- SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCM
QALQTPYTFGQGTK (SEQ ID NO: 12 - 1B6 and 1C1)
- (vi) a light chain variable domain comprising
 EIVLTQSPLSLPVTGPGEPAISCRSS**QSLLH**SNGYNYLDWYLQKPGQS
 5 PQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCM**Q**
GLQSPWTFGQGTK (SEQ ID NO: 16 - 1C11)
- (vii) a light chain variable domain comprising
 DVVMTQSPLSLPVTGPGEPAISCRSS**QSLLH**SNRYNYLDWYLQKPGQ
 10 SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCM
QALQTPYTFGQGTK (SEQ ID NO: 18 - 1D4)
- (viii) a light chain variable domain comprising
 DIQLTQSPSSLSASVGDRVITICRAS**QSID**FLNWWYQQKPGKAPKLLIY
AASSLQTGVPSRFGGSGSGTEFTLTISLQPEDLGTYCYCQ**SYIMPD**
TFGQGTK (SEQ ID NO: 22 - 1E4)
- 15 (ix) a light chain variable domain comprising
 DVVMTQSPLSLPVTGPGEPAISCRSS**QSLLH**SNGYNYLDWYLQKPGQ
 SPQLLIY**LGS**NRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYCM
QGLQTPYTFGQGTK (SEQ ID NO: 24 - 1H6)
- (x) a heavy chain variable domain comprising
 20 QVQLQESGGGLVQPGGSLRLSCADSG**SIFSG**NDAWYRRAPGVERE
 LVAVITSGG**STHY**ADSVKGRFTISRDNAAQKTVYLQTNDLKPEDTAVYY
CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 28- 1A5-VHH)
- (xi) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERE
 25 LVAFITNGG**STRY**TDVPKGRFTISRDNAAKNTVYLMNSLKAEDTAVYY
CMAGTSWGQGTQ (SEQ ID NO: 32 - 1B8-VHH)
- (xii) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCADSG**SIFSG**NDAWYRRAPGVERE
 LVAVITSGG**STHY**ADSVKGRFTISRDNAAQKTVYLQTNDLKPEDTAVYY
 30 **CTNGRWSGDTYYAHLWGQGTQ (SEQ ID NO: 34 - 2B7-VHH)**
- (xiii) a heavy chain variable domain comprising
 QVQLQESGGGLVQAGDSLRLSCADSG**SIFSG**NDAWYRRAPGVERE
 LVAVITSGG**STHY**ADSVKGRFTISRDNAAQKTVYLQTNDLKPEDTAVYY
CTNGRWSGDTYYAHHWGQGTQ (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 heavy chain variable domain selected from the group consisting of:

- (i) a heavy chain variable domain comprising
 QVQLQESGGGLVQAGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNQKTVYLQTNDLPEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 35 - 2C4-VHH)
- (ii) a heavy chain variable domain comprising
 QVQLVESGGGLVQAGGSLRLSCAASGSIFSGNDMSWYRQAPGKGLE
 LVAVITSGGSTYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 37 - hu2C4-VHH)
- (iii) a heavy chain variable domain comprising
 EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLE
 WMGGIPIFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDVAVY
 YCARSLGGRFRYWGQGTQ (SEQ ID NO: 4 - 1B6, 1C1, 1C11, 1D4
 and 1H6)
- (iv) a heavy chain variable domain comprising
 QVQLQQSGPGLVKPSQTLSTCAISGDSISSNSVAWNWIRQSPSRGL
 EWLGRITYYRSKWYSDYASVKGRDLINPDTSKNQFSLQLNSVTPEDT
 AVYYCAREVEGSSYDAFDIWGQGTQ (SEQ ID NO: 8 - 1E4),
- (v) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNQKTVYLQTNDLPEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 28- 1A5-VHH)
- (vi) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCAASGSSEFRTSVAWYRQAPGKERE
 LVAFITNGGSTRYTDVPKGRFTISRDNKNTVYLQMNSLKAEDTAVYY
 CMAGTSWGQGTQ (SEQ ID NO: 32 - 1B8-VHH)
- (vii) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNQKTVYLQTNDLPEDTAVYY
 CTNGRWSGDTYYAHLWGQGTQ (SEQ ID NO: 34 - 2B7-VHH)

(viii) a heavy chain variable domain comprising

QVQLQESGGGLVQAGDSLRLSCADSGSIFSGNDMAWYRRAPGVERE
LVAVITSGGSTHYADSVKGRFTISRDNAAQKTVYLQTNDLPEDTAVYY
CTNGRWSGDITYAHHWGQGTQ (SEQ ID NO: 36 - 2D10-VHH); and

5 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

10 having a sequence:

(i) QVQLQESGGGLVQAGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
LVAVITSGGSTHYADSVKGRFTISRDNAAQKTVYLQTNDLPEDTAVYY
CTNGRWSGDITYAHHWGQGTQ (SEQ ID NO: 35 - clone 2C4-VHH),

or

15 QVQLVESGGGLVQAGGSLRLSCAASGSIFSGNDMSWYRQAPGKGLE
LVAVITSGGSTHYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYY
CTNGRWSGDITYAHHWGQGTQ (SEQ ID NO: 37 - hu2C4-VHH), or

(ii) QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
LVAVITSGGSTHYADSVKGRFTISRDNAAQKTVYLQTNDLPEDTAVYY
20 CTNGRWSGDITYAHHWGQGTQ (SEQ ID NO: 28 - clone 1A5-VHH), or

(iii) QVQLQESGGGLVQPGGSLRLSCAASGSSEFRTSVAWYRQAPGKERE
LVAFITNGGSTRYTDPAKGRFTISRDNKNTLYLQMNSLKAEDTAVYY
CMAGTSWGQGTQ (SEQ ID NO: 32 - clone 1B8-VHH), or

(iv) QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
25 LVAVITSGGSTHYADSVKGRFTISRDNAAQKTVYLQTNDLPEDTAVYY
CTNGRWSGDITYAHLWGQGTQ (SEQ ID NO: 34 - clone 2B7-VHH), or

(v) QVQLQESGGGLVQAGDSLRLSCADSGSIFSGNDMAWYRRAPGVERE
LVAVITSGGSTHYADSVKGRFTISRDNAAQKTVYLQTNDLPEDTAVYY
CTNGRWSGDITYAHHWGQGTQ (SEQ ID NO: 36 - clone 2D10-VHH),

30 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

35 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 IPIFGTA (SEQ ID NO: 2), and
 a CDR-H3 comprising ARSLGGRFRY (SEQ ID NO: 3)
- 5 (ii) a heavy chain variable region comprising: (1E4)
 a CDR-H1 comprising GDSISSNSVA (SEQ ID NO: 5),
 a CDR-H2 comprising TYYSKWYS (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
 10 and 1H6)
 a CDR-L1 comprising:
 • QSLHNSNGYNY (SEQ ID NO: 9 - 1B6, 1C1, 1C11 and 1H6),
 • QSLHNSRYNY (SEQ ID NO: 17 - 1D4), or
 • QSIDF (SEQ ID NO: 19 - 1E4)
- 15 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),
 20 • 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.
- 25 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),
 30 a CDR-H2 comprising IPIFGTA (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 QSLHNSNGYNY (SEQ ID NO: 9),
 a CDR-L2 comprising LGS (SEQ ID NO: 10), and
 35 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 IPIFGTA (SEQ ID NO: 2), and
 - a CDR-H3 comprising ARSLGGRFRY (SEQ ID NO: 3); and/or
- 5 a light chain variable region comprising:
- a CDR-L1 comprising QSLHNSNGYNY (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)
- 10 a CDR-H1 comprising GGTFSSYA (SEQ ID NO: 1),
- a CDR-H2 comprising IPIFGTA (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 QSLHNSNRYNY (SEQ ID NO: 17),
- 15 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 IPIFGTA (SEQ ID NO: 2), and
 - a CDR-H3 comprising ARSLGGRFRY (SEQ ID NO: 3); and/or
- 20 a light chain variable region comprising:
- a CDR-L1 comprising QSLHNSNGYNY (SEQ ID NO: 9),
 - a CDR-L2 comprising LGS (SEQ ID NO: 10), and
 - a CDR-L3 comprising MQGLQTPYT (SEQ ID NO: 23); and
- 25 (v) a heavy chain variable region comprising: (1E4)
- a CDR-H1 comprising GDSISSNSVA (SEQ ID NO: 5),
 - a CDR-H2 comprising TYRSKWYS (SEQ ID NO: 6), and
 - a CDR-H3 comprising AREVEGSSYDAFDI (SEQ ID NO: 7); and/or
- a light chain variable region comprising: (1E4)
- 30 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:

(vi) a heavy chain variable domain comprising: (1B6 and 1C1)

5 EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWM
GGIIPFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDVAVYYCARS
LGGRFRYWGGTL (SEQ ID NO: 4), and/or
a light chain variable domain comprising
DVVMTQSPLSLPVTPGEPASISCRSSQSLHNSNGYNYLDWYLQKPGQSP
10 QLLIYLGSNRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQT
PYTFGQGTK (SEQ ID NO: 12)

(vii) a heavy chain variable domain comprising: (1C11)

EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWM
GGIIPFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDVAVYYCARS
15 LGGRFRYWGGTL (SEQ ID NO: 4), and/or
a light chain variable domain comprising
EIVLTQSPLSLPVTPGEPASISCRSSQSLHNSNGYNYLDWYLQKPGQSPQ
LLIYLGSNRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQGLQSP
WTFGQGTK (SEQ ID NO: 16)

(viii) a heavy chain variable domain comprising: (1D4)

EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWM
GGIIPFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDVAVYYCARS
20 LGGRFRYWGGTL (SEQ ID NO: 4), and/or
a light chain variable domain comprising
DVVMTQSPLSLPVTPGESASISCRSSQSLHNSRYNYLDWYLQKPGQSP
25 QLLIYLGSNRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQT
PYTFGQGTK (SEQ ID NO: 18)

(ix) a heavy chain variable domain comprising: (1E4)

QVQLQQSGPGLVKPSQTLSTCAISGDSISNSVAWNMRQSPSR
30 GLEWLGRITYYRSKWYSDYAIKVKGRLDINPDTSKNQFSLQLNSVT
PEDTAVYYCAREVEGSSYDAFDIWGGTM (SEQ ID NO: 8),
and/or
a light chain variable domain comprising

DIQLTQSPSSLSASVGDRVITICRAS**Q****S****I****S****D****F****L****N****W****Y****Q****K****P****G****K****A****P****K****L****L****I****Y****A****A**
S**S****L****Q****T****G****V****P****S****R****F****G****G****S****G****S****G****T****E****F****T****L****T****I****S****S****L****Q****P****E****D****L****G****T****Y****Y****C****Q****Q****S****Y****I****M****P****D****T****F****G****Q****G**
 TK (**SEQ ID NO: 22**), or

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

5 EVQLVQSGAEVKKPGSSVKV**S****C****K****A****S****G****G****T****F****S****S****Y****A****I****S****W****V****R****Q****A****P****G****Q****G****L****E****W****M**
 GGII**P****I****F****G****T****A****N****Y****A****Q****N****F****Q****G****R****V****T****M****T****A****D****T****S****I****S****T****A****Y****M****E****L****S****S****L****R****S****E****D****T****A****V****Y****Y****C****A****R****S**
 LGGR**F****R****Y****W****G****Q****G****T****L** (**SEQ ID NO: 4**), and/or

a light chain variable domain

DVVM**T****Q****S****P****L****S****L****P****V****T****P****G****E****P****A****S****I****S****C****R****S****S****Q****S****L****L****H****S****N****G****Y****N****Y****L****D****W****Y****L****Q****K****P****G****Q****S****P**
 10 QLLI**Y****L****G****S****N****R****A****S****G****V****P****D****R****F****S****G****S****G****S****G****T****D****F****T****L****Q****I****S****R****V****E****A****E****D****A****G****V****Y****Y****C****M****Q****G****L****Q****T**
 PYTF**G****Q****G****T****K** (**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.

15 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:

(iii) RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSG
 NSQESVAEQDSKDSTYLSSTLTLSKADYEKHKLYACEVTHQGLSSPVTK

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

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

25 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:

30 (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
- 5 CAGAGCCTCCTGCATAGTAATGGATACAACTAT (SEQ ID NO: 46),
- a CDR-L2 comprising TTGGGTTCT (SEQ ID NO: 47), and
- a CDR-L3 comprising ATGCAAGCTCTACAACTCCGTACACT (SEQ
ID NO: 48)
- (ii) a heavy chain variable region comprising: (1C11)
- 10 a CDR-H1 comprising GGAGGCACCTTCAGCAGCTATGCT (SEQ ID
NO: 38),
- a CDR-H2 comprising ATCATCCCTATCTTTGGTACAGCA (SEQ ID
NO: 39), and
- a CDR-H3 comprising GCGAGATCGTTGGGTGGGAGATTTCGCTAC
15 (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
- 20 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),
- 25 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:
- 30 a CDR-L1 comprising
- CAGAGCCTCCTGCATAGTAATAGATACAACTAT (SEQ ID NO: 54),
- a CDR-L2 comprising TTGGGTTCT (SEQ ID NO: 47), and
- a CDR-L3 comprising ATGCAAGCTCTACAACTCCGTACACT (SEQ
ID NO: 48)
- 35 (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
5 a CDR-H3 comprising GCGAGATCGTTGGGTGGGAGATTTGCTAC (**SEQ ID NO: 40**); and/or
a light chain variable region comprising:
a CDR-L1 comprising
CAGAGCCTCCTGCATAGTAATGGATACAACTAT (**SEQ ID NO: 46**),
10 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
15 (**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
20 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**)
25 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:
30 (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
35

a CDR-H3 comprising

ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC (SEQ ID NO: 64)

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

5 a CDR-H1 comprising GGAAGCTCCGAAAGATTACATCA (SEQ ID NO: 66),

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

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

10 (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

15 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.

20 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
GGGCTCTCTGAGACTCTCCTGTGCAGACTCTGGAAGCATCTTCAG
25 TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTTATTACTAGCGGTGGTAGTACACACTATG
CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
CGGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGATACTT
30 ACTATGCCCATCACTGGGGCCAGGGGACCCAG (SEQ ID NO: 72 -
clone 2C4),

(ii) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGG
GGGCTCTCTGAGACTCTCCTGTGCAGCCTCTGGAAGCATCTTCAG
TGGCAATGACATGTCTGGTACCGCCAGGCTCCAGGGAAGGGACT
35 CGAGTTGGTCGCGGTTATTACTAGCGGTGGTAGTACATACTATGC

- AGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATTCCaAG
AAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGG
CCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGATACTTACT
ATGCCCATCACTGGGGCCAGGGGACCCTG (SEQ ID NO: 74 - clone
5 hu2C4-VHH);
- (iii) GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
TGAGTGGATGGGAGGGAT**CATCCCTATCTTTGGTACAGCAAACTA**
10 CGCACAGAACTTCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41 - for clones 1B6,
1C1, 1C11, 1D4, 1H6), or
- 15 (iv) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC**
TAGTAACAGTGTTGCTTGGAACTGGATCAGGCAGTCCCCATCGAG
AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
ACAGTGATTATGCAATATCTGTGAAAGTCGATTAGACATCAACCC
20 AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG
(SEQ ID NO: 45 - clone 1E4);
- (v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
25 AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACACACTATG**
CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC
30 GGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGATACTTA
CTATGCCCATCACTGGGGCCAGGGGACCCAG (SEQ ID NO: 65 -
clone 1A5),
- (vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
GGGGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAG**
35 **ATTCACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGC**

- GCGAGTTGGTCGCATTTATTACTAATGGTGGTAGCACAAGATATAC
 AGACCCCGTGAAGGGCCGATTCACCATCTCCAGAGACAACGCCAA
 GAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACAC
 GGCCGTCTATTATTGTATGGCGGGTACGTCCTGGGGCCAGGGGAC
 5 CCAG (**SEQ ID NO: 69** - clone 1B8),
- (vii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 AGGGTCTCTGAGACTCTCCTGTGCAGACTCTGGAAGCATCTTCAG
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
 GCGAGTTGGTCGCGGTTATTACTAGCGGTGGTAGTACACACTATG
 10 CAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAATGCCC
 AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
 CGGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGATACTT
ACTATGCCCATCTCTGGGGCCAGGGGACCCAG (**SEQ ID NO: 71** -
 clone 2B7),
- 15 (viii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGG
 CTGGGGACTCTCTGAGACTCTCCTGTGCAGACTCTGGAAGCATCT
TCAGTGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTG
 GAGCGCGAGTTGGTCGCGGTTATTACTAGCGGTGGTAGTACACAC
 TATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGACAAT
 20 GCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAG
 GACACGGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGA
TACTTACTATGCCCATCACTGGGGCCAGGGGACCCAG (**SEQ ID**
NO: 73 - clone 2D10), or
 and/or
- 25 the light chain variable domain is encoded by a nucleotide sequence comprising:
- (vi) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 30 CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATTT
 TACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
 TACTGCATGCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG
 ACCAAG (**SEQ ID NO: 49** - clone 1B6 and 1C1), or
- (vii) GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 35 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**

5 **ATAGTAATGGATACAAC**TATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCTAATCGGGCCTC**
 CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGTACAGATTT
 TAACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
 TACTGCAT**GCAAGGTCTACAAAGTCCCTGGACGTT**CGGCCAAGGG
 ACCAAG (**SEQ ID NO: 53** - clone 1C11), or

(viii) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCAC
 CCCTGGAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCT**
 10 **CCTGCATAGTAATAGATACAAC**TATTTGGATTGGTACCTGCAGAAG
 CCAGGGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCTAATCGGG**
 CCTCCGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAG
 ATTTTAACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGT
 TTATTACTGCAT**GCAAGCTCTACAACTCCGTACACTTTT**GGCCAG
 GGGACCAAG (**SEQ ID NO: 55** - clone 1D4), or

15 (ix) GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
 GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT
 CCTGATCTAT**GCTGCATCGAGTTT**ACAACTGGGGTCCCTCAAGA
 TTCGGTGGCAGTGGATCTGGGACAGAATCACTCTCACCATAAGCA
 20 GTCTACAACCTGAAGATTTGGGAACTTATTACTGT**CAACAGAGTTA**
CATTATGCCCCGACACTTTTGGCCAGGGGACGAAA (**SEQ ID NO: 59** -
 clone 1E4), or

(x) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
 25 **ATAGTAATGGATACAAC**TATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCTAATCGGGCCTC**
 CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATTT
 TAACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
 TTACTGCAT**GCAAGGTCTACAGACTCCGTACACTTTT**GGCCAGGG
 30 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) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGG
 5 GGGCTCTCTGAGACTCTCCTGTGCAGACTCTGGAAGCATCTTCAG
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
 GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACACACTATG**
 CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
 AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
 10 CGGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGATACTT
ACTATGCCCATCACTGGGGCCAGGGGACCCAG (SEQ ID NO: 72 -
 clone 2C4), or
- (ii) CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGG
 GGGCTCTCTGAGACTCTCCTGTGCAGCCTCTGGAAGCATCTTCAG
 15 **TGGCAATGACATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACT**
 CGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACATACTATGC**
 AGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCaAG
 AAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGG
 CCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGATACTTACT
 20 **ATGCCCATCACTGGGGCCAGGGGACCCTG (SEQ ID NO: 74 - clone**
 hu2C4-VHH); or
- (iii) GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAG
 25 **CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT**
 TGAGTGGATGGGAGGGAT**CATCCCTATCTTTGGTACAGCAA**ACTA
 CGCACAGAACTTCAGGGCAGAGTCACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTT**CG
 30 **CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41 - for clones 1B6,**
 1C1, 1C11, 1D4, 1H6), or
- (iv) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
 GCAGACCCTCTCACTCACCTGTGCCATCTCCGGGGACAGTATCTC
TAGTAACAGTGTTGCTTGGAACTGGATCAGGCAGTCCCCATCGAG
 AGGCCTTGAGTGGCTGGGAAGGAC**ATACTACAGGTCCAAGTGGT**
 35 **ACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC**

AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
 CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG

(SEQ ID NO: 45 - clone 1E4); or

5 (v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
 GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACACACTATG**
 CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
 10 AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC
 GGCCGTGTATTACTGCACAAAC**GGAAGATGGTCAGGCGATACTTA**
CTATGCCCATCACTGGGGCCAGGGGACCCAG (SEQ ID NO: 65 -
 clone 1A5), or

(vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 15 GGGGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAG**
ATTCACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGC
 GCGAGTTGGTCGCATTT**ATTACTAATGGTGGTAGCACAAGATATAC**
 AGACCCCGTGAAGGGCCGATTACCATCTCCAGAGACAACGCCAA
 GAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACAC
 20 GGCCGTCTATTATTGT**ATGGCGGGTACGTCCTGGGGCCAGGGGAC**
 CCAG (SEQ ID NO: 69 - clone 1B8), or

(vii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGG
 AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
 25 GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACACACTATG**
 CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
 AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
 CGGCCGTGTATTACTGCACAAAC**GGAAGATGGTCAGGCGATACTT**
ACTATGCCCATCTCTGGGGCCAGGGGACCCAG (SEQ ID NO: 71 -
 30 clone 2B7), or

(viii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGG
 CTGGGGACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCT**
TCAGTGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTG
 GAGCGCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTACACAC**
 35 TATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAAT

GCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAG
 GACACGGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCCGA
TACTTACTATGCCCATCACTGGGGCCAGGGGACCCAG (SEQ ID
 NO: 73 - clone 2D10), and/or

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

(i) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 10 CGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATT
 TACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
 TACTGCAT**GCAAGCTCTACAACTCCGTACACT**TTTTGGCCAGGGG
 ACCAAG (SEQ ID NO: 49 - clone 1B6 and 1C1), or

(ii) GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 15 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 CGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGTACAGATT
 TACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
 20 TACTGCAT**GCAAGGTCTACAAAGTCCCTGGACGT**TCGGCCAAGGG
 ACCAAG (SEQ ID NO: 53 - clone 1C11), or

(iii) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATAGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 25 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 CGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATT
 TACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
 TACTGCAT**GCAAGCTCTACAACTCCGTACACT**TTTTGGCCAGGGG
 ACCAAG (SEQ ID NO: 55 - clone 1D4), or

(iv) GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
 GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT
 30 CCTGATCTAT**GCTGCATCG**AGTTTACAAACTGGGGTCCCTCAAGA
 TTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCACCATAAGCA
 35 GTCTACAACCTGAAGATTTGGGAACCTTATTACTGT**CAACAGAGTTA**

CATTATGCCCGACACTTTTGGCCAGGGGACGAAA (SEQ ID NO: 59 - clone 1E4), or

(v) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
5 **ATAGTAATGGATACA**ACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATT
TACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
TTACTGCAT**GCAAGGTCTACAGACTCCGTACACTTTTGGCCAGGG**
10 **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 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**
20 **CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT**
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAA**ACTA
CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
25 **CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41), and/or**

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

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
30 **ATAGTAATGGATACA**ACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCT**AATCGGGCCTC
CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCAT**GCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG**
35 **ACCAAG (SEQ ID NO: 49)**

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

5 GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAG
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGGATCATCCCTATCTTTGGTACAGCAAATA
 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGTGCGAGATCGTTGGGTGGGAGATTTTCG
 10 **CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41), and/or**

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

GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGTCAGAGCCTCCTGC
 15 **ATAGTAATGGATACAACCTATTTGGATTGGTACCTGCAGAAGCCAG**
 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGTACAGATTT
 TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
 TACTGCATGCAAGGTCTACAAAGTCCCTGGACGTTCCGGCCAAGGG
 20 **ACCAAG (SEQ ID NO: 53)**

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

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAG
 25 **CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT**
 TGAGTGGATGGGAGGGATCATCCCTATCTTTGGTACAGCAAATA
 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGTGCGAGATCGTTGGGTGGGAGATTTTCG
 30 **CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41), and/or**

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

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCCGTACCCCTG
 GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGTCAGAGCCTCCTGC
 35 **ATAGTAATAGATACAACCTATTTGGATTGGTACCTGCAGAAGCCAG**

GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATT
 TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
 TACTGCATGCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG
 5 ACCAAG (SEQ ID NO: 55)

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

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAG
 10 CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGGATCATCCCTATCTTTGGTACAGCAAACTA
 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGTGCGAGATCGTTGGGTGGGAGATTTCG
 15 CTACTGGGGCCAGGGAACCCTG (SEQ ID NO: 41), and/or

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

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGTCAGAGCCTCCTGC
 20 ATAGTAATGGATACAACATAATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATT
 TACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
 TTACTGCATGCAAGGTCTACAGACTCCGTACACTTTTGGCCAGGG
 25 GACCAAG (SEQ ID NO: 61); and

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

CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
 GCAGACCCTCTCACTCACCTGTGCCATCTCCGGGGACAGTATCTC
 30 TAGTAACAGTGTTGCTTGGAAGTGGATCAGGCAGTCCCCATCGAG
 AGGCCTTGAGTGGCTGGGAAGGACATACTACAGGTCCAAGTGGT
 ACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC
 AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
 CCCGAGGACACGGCTGTGTATTATTGTGCAAGAGAAGTTGAGGGC

AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG

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

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

5 GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
 GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT
 CCTGATCTAT**GCTGCATCG**AGTTTACAACTGGGGTCCCCTCAAGA
 TTCGGTGGCAGTGGATCTGGGACAGAATTCACCTCTCACCATAAGCA
 10 GTCTACAACCTGAAGATTTGGGAACCTTATTACTGT**CAACAGAGTTA**
CATTATGCCCCGACACTTTTGGCCAGGGGACGAAA (**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
 15 comprises comprises a light chain constant domain encoded by the nucleotide sequence comprising

CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCA
 GTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCC
 CAGAGAGGCCAAAGTACAGTGGAAAGGTGGATAACGCCCTCCAATCGGGTA
 20 ACTCCCAGGAGAGTGTGCGCAGAGCAGGACAGCAAGGACAGCACCTACAG
 CCTCAGCAGCACCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAG
 TCTACGCCTGCGAAGTCACCCATCAGGGCCTGAGCTCGCCCGTCACAAAG
 AGCTTCAACAGGGGAGAGTGT (**SEQ ID NO: 50 - clone 1B6 – light chain**
constant domain), or
 25 CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCA
 GTTGAAATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCC
 CAGAGAGGCCAAAGTACAGTGGAAAGGTGGATAACGCCCTCCAATCGGGTA
 ACTCCCAGGAGAGTGTGACAGAGCAGGACAGCAAGGACAGCACCTACAG
 CCTCAGCAGCACCTGACGCTGAGCAAAGCAGACTACGAGAAACACAAAG
 30 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
 5 GCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
 GGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**CACTATGCAGACTCCG
 TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
 ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
 10 **GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCACTGG**
 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
 15 GCTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACTCGAGTTG
 GTCGCGGTT**ATTACTAGCGGTGGTAGTAC**A TACTATGCAGACTCCGTG
 AAGGGCCGATTCACCATCTCCAGAGACAATTCCaAGAAcACcTATATC
 TGCAAATGAACAGCCTGAGAGCTGAGGACACGGCCGTGTATTACTGCA
 20 **CAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCACTGGGGC**
 CAGGGGACCCCTG (**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
 25 GGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
 GGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**CACTATGCAGACTCCG
 TGAAGGGCCGATTCACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
 ATCTGCAAACGAACGACTTGAAACCTGAGGACACGGCCGTGTATTACT
 30 **GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCACTGG**
 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
 35 GGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAGATTC**

- ACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGCGCGAGTT**
GGTCGCATTTATTACTAATGGTGGTAGCACAAGATATACAGACCCCGT
GAAGGGCCGATTACCATCTCCAGAGACAACGCCAAGAACACGGTGT
ATCTGCAAATGAACAGCCTGAAAGCTGAGGACACGGCCGTCTATTATT
 5 **GTATGGCGGGTACGTCTGGGGCCAGGGGACCCAG (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**
GGTCTCTGAGACTCTCCTGTGCAGACTCTGGAAGCATCTTCAGTGGCA
 10 **ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT**
GGTCGCGGTTATTACTAGCGGTGGTAGTACACACTATGCAGACTCCG
TGAAGGGCCGATTACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCTCTGG
 15 **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**
ACTCTCTGAGACTCTCCTGTGCAGACTCTGGAAGCATCTTCAGTGGCA
 20 **ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT**
GGTCGCGGTTATTACTAGCGGTGGTAGTACACACTATGCAGACTCCG
TGAAGGGCCGATTACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCACTGG
 25 **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**
CTCGGTGAAGGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAGCAGCT
 30 **ATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCTTGAGTGG**
ATGGGAGGGATCATCCCTATCTTTGGTACAGCAAACCTACGCACAGAAC
TTCCAGGGCAGAGTCACCATGACCGCAGACACCTCCATAAGCACAGC
CTACATGGAGCTGAGCAGCCTGAGATCTGAGGACACGGCCGTGTATTA
CTGTGCGAGATCGTTGGGTGGGAGATTTGCTACTGGGGCCAGGGAA
 35 **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
 GACCCCTCTCACTCACCTGTGCCATCTCCGGGGACAGTATCTCTAGTAA
 5 **CAGTGTGCTTGGA**ACTGGATCAGGCAGTCCCCATCGAGAGGCCTTG
 AGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGTACAGT**GATTAT
 GCAATATCTGTGAAAGGTGCGATTAGACATCAACCCAGACACATCCAAG
 AACCAGTTCTCCCTGCAGCTGAACTCTGTGACTCCCGAGGACACGGCT
 GTGTATTATTGT**GCAAGAGAAGTTGAGGGCAGCAGCTATGATGCTTTT**
 10 **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 examples, the multi-specific antigen binding protein may comprise an
 15 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 seven, 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
 20 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
 25 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
 30 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
 35 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, 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 immune cell binds to HER2 and secretes an inducible bispecific T cell engager that targets EpCAM and CD3 (HE CAR-BiTE T).

Without wishing to be bound by theory, while the CAR 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 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 cells will concomitantly contribute to the prevention of cancer relapse and recurrence by eliminating cancer stem cells and progenitor cells.

In addition, as noted in the experimental data of the present disclosure, it is demonstrated that HER2 targeting CAR T cells secreting anti-EpCAM BiTE (HE CAR-BiTE T) exerted superior efficacy in killing human tumours originating from gastric carcinomas or breast cancers *in vitro* as compared to HE CAR-T alone (FIGURE 1D and Figure 1E). Both clones 4D5 and F5 of HE CAR-BiTE T show higher % of cytolysis of AGS (HER2^{High}, EpCAM^{High}) and MDA-MB468 (HER2^{VeryLow}, EpCAM^{High}) cells as compared to HE CAR-T alone. In MDA-MB468 (HER2^{VeryLow}, EpCAM^{High})(FIGURE 1E), clone F5 and clone 4D5 showed minimal or reduced cytolysis of the MDA-MB468 as compared to their respective HE CAR-BiTE T. This suggests that the main cytotoxic function in MDA-MB468 cells is contributed by the anti-EpCAM BiTE. These results further highlight that the HE CAR-BiTE T cells of the present invention only targets the host cells when the expression levels of EpCAM is high to avoid targeting cells with low EpCAM expression. Clone 4D5 of HE CAR-T cells could not initiate efficient killing of AGS cells (FIG. 1D) with low level of interferon γ secretion (FIGURE 1F), suggesting that the cytotoxic function of HE CAR-BiTE T cells for clone 4D5 was mainly contributed by

the secreted anti-EpCAM BiTE. The present disclosure has also shown that the modified cell as disclosed herein achieved a more efficient and prolonged tumour controls than treatment with anti-HER2 CAR T cells alone (see Figure 2).

In some examples, cell as disclosed herein may engage immune cells in the target environment that include, but is not limited to 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 immune cell is a T cell or a macrophage or a monocyte or an NK cell. In some examples, the immune cell is a T cell.

Therefore, in some examples, the cell binds to HER2 and secretes a multi-specific antigen binding protein that binds to both EpCAM and CD3 (e.g. HE CAR-BiTE T).

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 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 CD3 ζ , CD3 ϵ , CD3 γ , CD3 δ , 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 Otk3 (Nb01-013A).

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 HER2 positive cell. In some examples, the first antigen is HER2. Therefore, in some examples, the CAR is capable of binding to HER2.

In some examples, the polynucleotide further comprises sequences encoding one or more co-stimulatory domain, a signal peptide, a hinge, and/or a signaling 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 HER2 (or anti-HER2 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- HER2 scFv.

In some examples, the sequence encoding CAR further encodes for an immune
5 cell signaling domain. In some examples, the signaling domain may include, but is not limited to, a CD3, a TCF ζ , a FcR γ , a FcR β , a CD3 γ , a CD3 θ , a CD3 ϵ , a CD3 η , a CD3 ζ , 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 CD3 ζ , CD3 ϵ , CD3 γ , CD3 δ , and the like. In some examples, the sequence encoding CAR further encodes for
10 a CD3 intracellular domain (signaling domain).

In some examples, the sequence encoding CAR encodes for an anti- HER2 scFv, a CD28 costimulatory domain, a 4-1BB co-stimulatory domain, an IgH signal peptide, an IgH hinge, and a CD3 signaling domain.

In some examples, the sequence encoding the immune cell engager encodes for
15 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.

20 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.

25 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
30 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
35 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
5 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
10 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-HER2 scFv CAR with CD3 intracellular domain, and the sequence encoding for a BiTE that binds to EpCAM and a T cell.

15 In yet another aspect, there is provided 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
20 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
25 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.

30 In some embodiments, 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 above.

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
35 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
5 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
10 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
15 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
20 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
25 cells.

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

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

30 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.

35 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.

5 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.

10 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 secreted at the target site (for example at the site of a solid tumour), thereby minimising toxicity and/or side effects.

15 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,

25 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.

30

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 some examples, the cell is an immune cell.

Also disclosed are methods of producing / generating the cell as described
35 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
5 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
10 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
15 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
20 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
25 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
30 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,
35 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, esophagus cancer, bladder cancer, colon cancer, childhood cancers (such as hepatoblastomas, nephroblastomas, 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)

EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGGIIPIFGT
 5 ANYAQNFQGRVTMTADTSISTAYMELSSLRSEDVAVYYCARSLGGRFRYWGQGTL

1E4 (SEQ ID NO: 8)

QVQLQQSGPGLVKPSQTLSTCAISGDSISSNSVAWNWIRQSPSRGLEWLGRYYRS
KWYSDYAI SVKGRLDINPDTSKNQFSLQLNSVTPEDTAVYYCAREVEGSSYDAFDW
 GQGTM

1.2. Light chain variable domain of human IgG1
--

1B6 (SEQ ID NO: 12)

DVVMTQSPLSLPVTGPGEPAISCRSSQ**SLLHS**NGYNYLDWYLQKPGQSPQLLIY**LGS**
 NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQTPYTFGQGTK

15 **1B6 (light chain constant domain) (SEQ ID NO: 13)**

RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESV
 EQDSKDYSLSTLTLSKADYEKHKYACEVTHQGLSSPVTKSFIRGEC

1C1 (SEQ ID NO: 12)

DVVMTQSPLSLPVTGPGEPAISCRSSQ**SLLHS**NGYNYLDWYLQKPGQSPQLLIY**LGS**
 20 NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQTPYTFGQGTK

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

RTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESV
 EQDSKDYSLSTLTLSKADYEKHKYACEVTHQGLSSPVTKSFIRGEC

1C11 (SEQ ID NO: 16)

25 EIVLTQSPSLSPVTPGEPASISCRSSQ**SLLHS**NGYNYLDWYLQKPGQSPQLLIY**LGSN**
 RASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQGLQSPWTFGQGTK

1D4 (SEQ ID NO: 18)

DVVMTQSPLSLPVTGPGEPAISCRSSQ**SLLHS**NRNYLDWYLQKPGQSPQLLIY**LGS**
 NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQTPYTFGQGTK

30 **1E4 (SEQ ID NO: 22)**

DIQLTQSPSSLSASVGDRTITCRASQ**SISD**FLNWWYQQKPGKAPKLLIY**AASS**LQTGVP
 SRFGGSGSGTEFTLTISLQPEDLGTYYCQQSYIMPDTFGQGTK

1H6 (SEQ ID NO: 24)

DVVMTQSPLSLPVTGPGEPAISCRSSQSLHNSNGYNYLDWYLQKPGQSPQLLIYLGSR
NRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYCMQGLQTPYTFGQGTK

1.3. Single domain heavy chain variable domain of llama VHH

5 **2C4-VHH (SEQ ID NO: 35)**

QVQLQESGGGLVQAGGSLRLSCADSGSIFSGNDMAWYRRAPGVERELVAVITSGGS
THYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYYCTNGRWSGDTYYAHHWGQ
GTQ

1A5-VHH (SEQ ID NO: 28)

10 QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERELVAVITSGGS
THYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYYCTNGRWSGDTYYAHHWGQ
GTQ

1B8-VHH (SEQ ID NO: 32)

15 QVQLQESGGGLVQPGGSLRLSCAASGSSERFTSVAWYRQAPGKERELVAFITNGGS
TRYTDPVKGRFTISRDNAAKNTVYLMNSLKAEDTAVYYCMAGTSWGQGTQ

2B7-VHH (SEQ ID NO: 34)

QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERELVAVITSGGS
THYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYYCTNGRWSGDTYYAHLWGQ
GTQ

20 **2D10-VHH (SEQ ID NO: 36)**

QVQLQESGGGLVQAGDSLRLSCADSGSIFSGNDMAWYRRAPGVERELVAVITSGGS
THYADSVKGRFTISRDNAAKTVYLQTNDLKPEDTAVYYCTNGRWSGDTYYAHHWGQ
GTQ

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

25

hu2C4-VHH (SEQ ID NO: 37)

QVQLVESGGGLVQAGGSLRLSCAASGSIFSGNDMSWYRQAPGKGLELVAVITSGGS
TYYADSVKGRFTISRDNASKNTLYLMNSLRAEDTAVYYCTNGRWSGDTYYAHHWGQ
GTL

1.5. Anti-EpCAM Okt3 BiTE

30

Nb01-013A (SEQ ID NO: 75)

Annotation

35 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)

5 QVQLQESGGGLVQAGGSLRLSCADSGSIFSGNDMAWYRRAPGVERELVAVITSGGS
THYADSVKGRFTISRDNQKTVYLQTNLKPEDTAVYYC TNGRWSGDTYYAHHWGQ
GTQVTVSSGGGGS QVQLQQSGAELARPGASVKMSCKASGYTFTRYTMHWWKQRP
GQGLEWIGYINPSRGYTNYNQKFKDKATLTDDKSSSTAYMQLSSLTSEDSAVYYCAR
YYDDHYCLDYWGQGTTLTVSSGGGGSGGGSGGGSGGGGSQIVLTQSPAIMSASPGEKV
10 TMTCSASSSVSYMNWYQKSGTSPKRWIYDTSKLASGVPAHFRGSGSGTSYSLTIS
GMEAEDAATYYCQWSSNPFTFGSGTKLEINHHHHHH

2. Nucleotide Sequences

15 2.1. Heavy chain variable domain of human IgG1

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

GAGGTCCAGCTGGTGCACTCTGGGGCTGAGGTGAAGAAGCCTGGGTCTCGGT
GAAGTCTCCTGCAAGGCTTCTGGAGGCACCTTCAGCAGCTATGCTATCAGCTG
20 GGTGCGACAGGCCCTGGACAAGGGCTTGAGTGGATGGGAGGGATCATCCCTAT
CTTTGGTACAGCAAACTACGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCA
GACACCTCCATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGAC
ACGGCCGTGTATTACTGT GCGAGATCGTTGGGTGGGAGATTTCGCTACTGGGGC
CAGGGAACCCTG

25 1E4 (SEQ ID NO: 45)

CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTCGCAGACCCTC
TCACTCACCTGTGCCATCTCCGGGGACAGTATCTCTAGTAACAGTGTTGCTTGA
ACTGGATCAGGCAGTCCCCATCGAGAGGCCTTGAGTGGCTGGGAAGGACATACT
ACAGGTCCAAGTGGTACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACAT
30 CAACCCAGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACTCCC
GAGGACACGGCTGTGTATTATTGT GCAAGAGAAGTTGAGGGCAGCAGCTATGAT
GCTTTTGATATCTGGGGCCAAGGGACAATG

35 2.2. Light chain variable domain of human IgG1

1B6 (SEQ ID NO: 49)

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGCCGG
CCTCCATCTCCTGCAGGTCTAGT CAGAGCCTCCTGCATAGTAATGGATACAACTA

TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTG**
GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAAGTGGCAGTGGATCAGGC
 ACAGATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATT
 ACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGGACCAAG

5 **1B6 (light chain constant domain) (SEQ ID NO: 50)**

CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGA
 AATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGC
 CAAAGTACAGTGGAAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAG
 TGTC**C**CAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCCTGAC
 10 GCTGAGCAAAGCAGACTACGAGAAACACAA**C**CTACGCCTGCGAAGTCACCCAT
 CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCA**C**CAGGGGAGAGTGT

1C1 (SEQ ID NO: 49)

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCCGTCACCCCTGGAGAGCCGG
 CCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATGGATACA**ACTA
 15 TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTG**
GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAAGTGGCAGTGGATCAGGC
 ACAGATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATT
 ACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGGACCAAG

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

CGAACTGTGGCTGCACCATCTGTCTTCATCTTCCCGCCATCTGATGAGCAGTTGA
 AATCTGGAAGTGCCTCTGTTGTGTGCCTGCTGAATAACTTCTATCCCAGAGAGGC
 CAAAGTACAGTGGAAAGGTGGATAACGCCCTCCAATCGGGTAACTCCCAGGAGAG
 TGTC**C**CAGAGCAGGACAGCAAGGACAGCACCTACAGCCTCAGCAGCACCCCTGAC
 GCTGAGCAAAGCAGACTACGAGAAACACAA**C**CTACGCCTGCGAAGTCACCCAT
 25 CAGGGCCTGAGCTCGCCCGTCACAAAGAGCTTCA**C**CAGGGGAGAGTGT

1C11 (SEQ ID NO: 53)

GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCCGTCACCCCTGGAGAGCCGG
 CCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATGGATACA**ACTA
 TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTG**
 30 **GGTTCT**AATCGGGCCTCCGGGGTCCCTGACAGGTTCAAGTGGCAGTGGATCAGGT
 ACAGATTTTACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATTA
 CTGC**ATGCAAGGTCTACAAAGTCCCTGGACG**TTCCGGCCAAGGGACCAAG

1D4 (SEQ ID NO: 55)

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCCGTCACCCCTGGAGAGTCGG
 CCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATAGATACA**ACTA
 TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTG**
 35 **GGTTCT**AATCGGGCCTCCGGGGTCCCTGACAGGTTCAAGTGGCAGTGGATCAGGC
 ACAGATTTTACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTATT
 ACTGC**ATGCAAGCTCTACAAACTCCGTACACT**TTTGGCCAGGGGACCAAG

1E4 (SEQ ID NO: 59)

GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAGGAGACAGAG
 TCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCGACTTTTTAAATTGGTACCA**
 GCAGAAACCAGGTAAAGCCCCGAAGCTCCTGATCTAT**GCTGCATCGAGTTTACAA**
 5 ACTGGGGTCCCCTCAAGATTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCA
 CCATAAGCAGTCTACAACCTGAAGATTTGGGAACCTATTACTGT**CAACAGAGTTAC**
ATTATGCCCGACACTTTTGGCCAGGGGACGAAA

1H6 (SEQ ID NO: 61)

GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTGGAGAGCCGG
 10 CCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGCATAGTAATGGATACAAC**
 TTTGGATTGGTACCTGCAGAAGCCAGGGCAGTCTCCACAGCTCCTGATCTAT**TTG**
GGTTCTAATCGGGCCTCCGGGGTCCCTGACAGGTTCAAGTGGCAGTGGATCAGGC
 ACAGATTTTACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTATT
 ACTGC**ATGCAAGGTCTACAGACTCCGTACACT**TTTGGCCAGGGGACCAAG

15

2.3. Single domain heavy chain variable domain of llama VHH

1A5 (SEQ ID NO: 65)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAGGGTCTCT
 GAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGACATGGCCTG**
 20 GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT**ATTACTAGCG**
GTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGA
 CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACACG
 GCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT**
CACTGGGGCCAGGGGACCCAG

1B8 (SEQ ID NO: 69)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGGGGGTCTCT
 GAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAGATTCACATCAGTGGCCTG**
 GTACCGCCAGGCTCCAGGAAAGGAGCGCGAGTTGGTCGCATTT**ATTACTAATGG**
TGGTAGCACAAGATATACAGACCCCGTGAAGGGCCGATTACCATCTCCAGAGA
 30 CAACGCCAAGAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACACG
 GCCGTCTATTATTGT**ATGGCGGGTACGTCCT**GGGGCCAGGGGACCCAG

2B7 (SEQ ID NO: 71)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAGGGTCTCT
 GAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGACATGGCCTG**
 35 GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT**ATTACTAGCG**
GTGGTAGTACACACTATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGA
 CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACACG
 GCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT**
CTCTGGGGCCAGGGGACCCAG

2C4 (SEQ ID NO: 72)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCT
 GAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGACATGGCCTG**
 GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT**ATTACTAGCG**
GTGGTAGTACACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGA
 5 CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACACG
 GCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT**
CACTGGGGGCCAGGGGACCCAG

2D10 (SEQ ID NO: 73)

CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGACTCTCT
 10 GAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCAATGACATGGCCTG**
 GTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTTGGTCGCGGTT**ATTACTAGCG**
GTGGTAGTACACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGA
 CAATGCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACACG
 GCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCAT**
 15 **CAC**TGGGGGCCAGGGGACCCAG

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

hu2C4-VHH (SEQ ID NO: 74)

CAGGTGCAGCTGGTGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGGGCTCTCT
 20 GAGACTCTCCTGTGCAGCCTCT**GGAAGCATCTTCAGTGGCAATGACATGTCCTG**
 GTACCGCCAGGCTCCAGGGAAGGGACTCGAGTTGGTCGCGGTT**ATTACTAGCGG**
TGGTAGTACA TACTATGCAGACTCCGTGAAGGGCCGATTCACCATCTCCAGAGAC
 AATTCCaAGAAcACcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGGC
 25 CGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCA**
CTTGGGGGCCAGGGGACCCTG

DESCRIPTION OF FIGURES

Example embodiments of the disclosure will be better understood and readily
 30 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
 35 not be construed as limiting the scope of the disclosure.

Figure 1A shows the construction and structure map of HE CAR-BiTE T (clone
 4D5 or F5) using anti-EpCAM VHH BiTE (Nb01-013A). **Figure 1B** shows the FACS
 analysis of tumour markers (HER2 and EpCAM) expressed on AGS and MDA-MB468

cells. **Figure 1C** shows the percentage of CAR expression on anti-HER2 CAR T (4D5 or F5) and HE CAR- BiTE T (4D5 or F5) cells detected by flow cytometry analysis. **Figure 1D** shows the percentage of cytolysis of AGS cells mediated by anti-HER2 CAR T (4D5 or 5F) and HE CAR-BiTE T (4D5 or F5) cells with time-course measurement using xCelligence impedance assay (E:T ratio = 2:1). **Figure 1E** shows the percentage of cytolysis of MDA-MB468 cells mediated by anti-HER2 CAR T (4D5 or 5F) and HE CAR-BiTE T (4D5 or F5) cells with time-course measurement using xCelligence impedance assay (E:T ratio = 1:1). **Figure 1F** shows bar graphs of ELISA measurement of Interferon- γ and IL-2 as measured using culture supernatant collected at 36 hours post co-culture effector CAR T cells with AGS cells. **Figure 1G** shows bar graphs of ELISA measurement of Interferon- γ and IL-2 as measured using culture supernatant collected at 36 hours post co-culture effector CAR T cells with MDA-MB468 cells.

Figure 2 shows human HE CAR-BiTE T cells using anti-EpCAM VHH constructed BiTEs showed superior tumour killing in *in vivo* AGS xenografts. 1.3 million AGS (Her2^{High}, EpCAM^{High}) cells were subcutaneously injected into the right flank of female NSG mice (Day -9). On Day 0, mice were re-grouped according to the measurable tumour size and five million CAR T (4D5) cells, CAR T (F5) cells, HE CAR-BiTE T (4D5) cells ("Nb01-013A" was used as BiTE), HE CAR-BiTE T (F5) cells ("Nb01-013A" was used as BiTE), or mock T cells were intravenously injected into these mice via tail vein. **Figure 2A** shows tumour sizes of each mouse were measured and recorded every 3 to 7 days. **Figure 2B** shows tumour sizes of each individual mouse in each treatment group were measured and recorded every 3 to 7 days. **Figure 2C** shows weights of each mouse were measured and recorded every 3 to 7 days.

Figure 3 shows the detection of anti-EpCAM BiTE molecule Nb01-013A secreted from HE CAR-BiTE T cells by ELISA. Anti-HER2 CAR T (4D5 or F5) cells or anti-HER2 CAR T cells secreting anti-EpCAM BiTE (Nb01-013A) (named as "HE CAR-BiTE T (4D5 or F5)") were cultured in T cell growth medium containing IL-7 (20 ng/ml) and IL-15 (5 ng/ml). The starting cell density was 1.5 million per ml and the percentages of CAR expression ranged from 36.8% to 57.2%. Cell culture supernatants were collected at 24h and 48h for an ELISA to detect the amount of secreted anti-EpCAM BiTE molecule Nb01-013A. 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

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

In this invention, the inventors have constructed "HE CAR-BiTE T" by fusing the anti-EpCAM BiTE gene (sequence of "Nb01-013A" was used) to the anti-HER2 CAR (clone 4D5 from "Trastuzumab" or our in-house anti-HER2 clone F5) lentiviral construct via a P2A cleavable linker (Figure 1A). Following the lentiviral transduction and expansion (Figure 1C), both AGS (HER2^{High}, EpCAM^{High}) and MDA-MB468 (HER2^{VeryLow}, EpCAM^{High}) cells (Figure 1B) were used to test the *in vitro* killing efficacy of CAR T cells. The HE CAR-BiTE T cells were shown to be able to kill AGS cells at a much faster rate than the anti-HER2 CAR alone (Figure 1D), suggesting the participation of the BiTE in target cell killing. Surprisingly, clone 4D5 constructed anti-HER2 CAR T cells could not initiate efficient killing of AGS cells (Figure 1D) with low level of Interferon- γ secretion (Figure 1F), suggesting that the cytotoxic function of the HE CAR-BiTE T cells (clone 4D5) was mainly contributed by the secreted anti-EpCAM BiTE. More intriguingly, anti-HER2 CAR T cells (clone 4D5) showed comparable level of killing of MDA-MB468 (HER2^{VeryLow}, EpCAM^{High}) cells, while the anti-HER2 CAR T cells constructed using our patented clone F5 showed high specificity by sparing the cells expressing very low level of HER2 (Figure 1E and 1G).

In summary, we have developed a novel CAR T cell therapy in targeting carcinomas (solid tumours originated from epithelial cells) 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, we could also maximize the efficacy of CAR T cell therapies 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. We have shown the development of a prototype (HE CAR-BiTE), which also serve as a proof-of-concept. While HE CAR-BiTE have been demonstrated to work successfully, our technology is not limited to HER2- 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 HE CAR-BiTE cells showed superior killing potency in mouse xenograft models.

The anti-EpCAM BiTE "Nb01-013A" was constructed by pairing the anti-EpCAM VHH clone 2C4 and the anti-human CD3 agonist clone Okt3. After testing the *in vitro* killing potency of HE CAR-BiTE T ("Nb01-013A" was used as BiTE) cells using different CAR clones (4D5 or F5), we proceeded to test their tumour control potency using the *in vivo* AGS (Her2^{High}, EpCAM^{High}) xenograft models (Figure 2). In line with the *in vitro* data, AGS appeared to be resistant to 4D5 CAR-T. While CAR T (4D5) failed to control the tumour growth, the other groups of CAR T (F5) or CAR-BiTE T cell treatments showed some *in vivo* tumour control efficacies (Figure 2A) without significant weight loss (Figure 2C). CAR T (F5) treatment showed effective control of tumour growth in all 5 mice. However, after 5-6 weeks of tumour shrinkage, the remaining tumours started to grow back and the tumour sizes reached more than 1000mm³, suggesting the tumour escape from the Her2^{Negative} cell populations. Effective tumour control was observed in 3 and 2 mice, respectively from the HE CAR-BiTE T (4D5) and HE CAR-BiTE T (F5) group, and these mice remained tumour-free up to 181 days post T cell infusion (Figure 2B), showing prolonged tumour controls by clearing Her2^{Negative}EpCAM^{Positive} tumour cells.

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

The Nb01-013A BiTE molecules can also be detected from HE CAR-BiTE T (4D5) and HE CAR-BiTE T (F5) cells from the *in vitro* cultures. Anti-HER2 CAR T (4D5 or F5) cells or HE CAR-BiTE T (4D5 or F5) cells were cultured in T cell growth medium containing IL-7 and IL-15 at the starting cell density of 1.5 million per ml. Cell culture supernatants were collected at 24h and 48h followed by an ELISA to detect the amount of secreted anti-EpCAM BiTE molecule Nb01-013A. For both HE CAR-BiTE T (4D5) and HE CAR-BiTE T (F5), the amount of secreted Nb01-013A molecules increased overtime, suggesting a continuous secretion and accumulation of anti-EpCAM VHH BiTE from both cultures (Figure 3).

In summary, we have shown that the usage of BiTE molecules constructed from an anti-EpCAM VHH (Nb01-013A) in HE CAR-BiTE T cells achieved a more efficient and prolonged tumour controls than treatment with anti-Her2 CAR T cells alone.

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-
5 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.

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

1. Anti-HER2 CAR T cells secreting anti-EpCAM BiTE can be used to treat HER2 expressing. 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, esophagus, bladder, and colon cancers.
- 15 2. Anti-EpCAM BiTE secreting CAR T cells may be configured to target other tumour specific antigens including but not limited to GPC3, Claudin18.2, ROR1, DLL3, CEA, MUC1, MUC16, CEACAM7, CD133, CD147, PSCA, PSMA, MSLN, c-Met, FR α , and the like.

This present disclosure provides for an improvement to treatment efficacy of
20 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
25 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
30 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 HER2, and

5 (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,

10 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:

- 15
- 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
- 20

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
- 25 • 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:

- 30 • 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:

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

10 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:

15 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

20 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:

25 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,

30 a CDR-L2 comprising:

- 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

35 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**) 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.

20

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

30

(ii) a heavy chain variable region comprising:

- a CDR-H1 comprising GGTFSYA (**SEQ ID NO: 1**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 5 a CDR-H2 comprising IIPFGTA (**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 ARSLGGRFY (**SEQ ID NO: 3**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto
- 10 (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,
- 15 a CDR-H2 comprising TYRSKWYS (**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
- 20 (iv) a light chain variable region comprising:
- a CDR-L1 comprising QSLHSNGYNY (**SEQ ID NO: 9**) or a sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identity thereto,
- 25 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
- 30 (v) a light chain variable region comprising:
- a CDR-L1 comprising QSLHSNGYNY (**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

5 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:

10 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

15 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:

20 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

25 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:

30 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:

5 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

10 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:

15 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

20 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.

25

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/or a light chain variable region selected from the group consisting of:

30 (i) a heavy chain variable region comprising:

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

- a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto, and
- 5 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto; and/or
- a light chain variable region comprising:
- a CDR-L1 comprising QSLHNSNGYNY (**SEQ ID NO: 9**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto,
- 10 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto, and
- a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto
- 15 (ii) a heavy chain variable region comprising:
- a CDR-H1 comprising GGTFSSYA (**SEQ ID NO: 1**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto,
- 20 a CDR-H2 comprising IIPIFGTA (**SEQ ID NO: 2**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto, and
- a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto; and/or
- 25 a light chain variable region comprising:
- a CDR-L1 comprising QSLHNSNGYNY (**SEQ ID NO: 9**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto,
- 30 a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto, and

a CDR-L3 comprising MQGLQSPWT (**SEQ ID NO: 15**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto

(iii) a heavy chain variable region comprising:

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

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

10 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto; and/or

a light chain variable region comprising:

15 a CDR-L1 comprising QSLLSNRYNY (**SEQ ID NO: 17**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto,

a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto, and

20 a CDR-L3 comprising MQALQTPYT (**SEQ ID NO: 11**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto

(iv) a heavy chain variable region comprising:

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

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

30 a CDR-H3 comprising ARSLGGRFRY (**SEQ ID NO: 3**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto; and/or

a light chain variable region comprising:

- a CDR-L1 comprising QSLHNSGYNY (**SEQ ID NO: 9**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto,
- a CDR-L2 comprising LGS (**SEQ ID NO: 10**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto, and
- a CDR-L3 comprising MQGLQTPYT (**SEQ ID NO: 23**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto; and
- (v) a heavy chain variable region comprising:
- a CDR-H1 comprising GDSISSNSVA (**SEQ ID NO: 5**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto,
- a CDR-H2 comprising TYYRSKWYS (**SEQ ID NO: 6**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto, and
- a CDR-H3 comprising AREVEGSSYDAFDI (**SEQ ID NO: 7**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto; and/or
- a light chain variable region comprising:
- a CDR-L1 comprising QSISDF (**SEQ ID NO: 19**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto,
- a CDR-L2 comprising AAS (**SEQ ID NO: 20**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical thereto, and
- a CDR-L3 comprising QQSYIMPDT (**SEQ ID NO: 21**) or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical 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 – 1B8**) 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:

5 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

10 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:

15 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

20 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.

25 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

30 QVQLVESGGGLVQAGGSLRLSCAASGSIFSGNDMSWYRQAPGKGLE
LVAVITSGGSTYYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYY
CTNGRWSGDTYYAHHWGQGT (SEQ ID NO: 37) 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
35 acids substitutions

- (ii) a heavy chain variable domain comprising
 QVQLQESGGGLVQAGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNAAQKTVYLQTNDLKPEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (**SEQ ID NO: 35**) or a fragment,
 5 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
 EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLE
 10 WMGGIPIFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDNAVY
 YCARSLGGRFRYWGGTL (**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
- (iv) a heavy chain variable domain comprising
 QVQLQQSGPGLVKPSQTLTLTCAISGDSISSNSVAWNWIRQSPSRGL
 EWLGRITYYRSKWYSDYAIKGRDLINPDTSKNQFSLQLNSVTPEDT
 15 AVYYCAREVEGSSYDAFDIWGQGT (**SEQ ID NO: 8**) or a fragment,
 variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,
 20 98%, 99%, or 100% identical thereto and/or having two or three amino
 acids substitutions,
- (v) a light chain variable domain comprising
 DVVMTQSPLSLPVTPGEPASISCRSSQSLLHSNGYNYLDWYLQKPGQ
 SPQLLIYLGSNRASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVYYCM
 25 **QALQTPYTFGQGTK** (**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
- (vi) a light chain variable domain comprising
 30 EIVLTQSPLSLPVTPGEPASISCRSSQSLLHSNGYNYLDWYLQKPGQS
 PQLLIYLGSNRASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQ
GLQSPWTFGQGTK (**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
 35 substitutions

- (vii) a light chain variable domain comprising
 DVVMTQSPLSLPVTPGESASISCRSS**QSLLHSNRYNYLDWYLQKPGQ**
 SPQLLIY**LGS**NRASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVYY**CM**
QALQTPYTFGQGTK (SEQ ID NO: 18) or a fragment, variant or
 5 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
- (viii) a light chain variable domain comprising
 DIQLTQSPSSLSASVGDRVITITCRAS**QSIDFLN**WYQQKPGKAPKLLIY
 10 **AASSLQ**TGVPSRFGSGSGTEFTLTISLQPEDLGTY**YCQQSYIMPDT**
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
- (ix) a light chain variable domain comprising
 15 DVVMTQSPLSLPVTPGEPASISCRSS**QSLLHSNGYNYLDWYLQKPGQ**
 SPQLLIY**LGS**NRASGVDPDRFSGSGSGTDFTLQISRVEAEDAGVYY**CM**
QGLQTPYTFGQGTK (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
 20 substitutions
- (x) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCADSG**SIFSGND**MAWYRRAPGVERE
 LVAVITSGG**STHY**ADSVKGRFTISRDN**AQKTVYLQ**TNDLKPEDTAVYY
CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 28) or a fragment,
 25 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
- (xi) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERE
 30 LVAFIT**NGGSTR**YTDVPKGRFTISRDN**AKNTVYLQ**MNSLKAEDTAVYY
CMAGTSWGQGTQ (SEQ ID NO: 32) 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

(xii) a heavy chain variable domain comprising
 QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNQKTVYLQTNLKPEDTAVYY
 CTNGRWSGDTYYAHLWGQGTQ (SEQ ID NO: 34) 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

(xiii) a heavy chain variable domain comprising
 QVQLQESGGGLVQAGDSLRLSCADSGSIFSGNDMAWYRRAPGVERE
 LVAVITSGGSTHYADSVKGRFTISRDNQKTVYLQTNLKPEDTAVYY
 CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 36) 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.

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:

EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWM
 GGIPIFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDVAVYYCARS
 LGGRFRYWQGTQ (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

DVVMVTQSPLSLPVTGPGEPAISCRSSQSLHNSNGYNYLDWYLQKPGQSP
 QLLIYLGSNRASGVDPDRFSGSGSGTDFTLKISRVEAEDVGVYYCMQALQT
 PYTFGQGTK (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:

EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWM
 GGIPIFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDVAVYYCARS

- 5 **LGGRFRYWQGQTL (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
 EIVLTQSPLSLPVTGPGEASISCRSS**QSLH**SNGYNYLDWYLQKPGQSPQ
 LLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQGLQSP**
WTFGQGTK (SEQ ID NO: 16) or a fragment, variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or 100% identical
 10 thereto and/or having two or three amino acids substitutions
 (iii) a heavy chain variable domain comprising:
 EVQLVQSGAEVKKPGSSVKVSCKAS**GGTFSS**YAIWVRQAPGQGLEWM
 GGII**PIFGT**ANYAQNFGGRVTMTADTSISTAYMELSSLRSEDVAVYYC**CARS**
LGGRFRYWQGQTL (SEQ ID NO: 4) or a fragment, variant or sequence
 15 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
 DVVMTQSPLSLPVTGPGEASISCRSS**QSLH**SNRYNYLDWYLQKPGQSP
 20 QLLIY**LGS**NRASGVPDRFSGSGSGTDFTLKISRVEAEDVGVYYC**MQALQT**
PYTFGQGTK (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:
 25 QVQLQQSGPGLVKPSQTLSTCAIS**GDSSNS**VAWNWIRQSPSR
 GLEWLGR**TYRSKWY**SDYAIKGRDLINPDTSKNQFSLQLNSVT
 PEDTAVYYC**AREVEGSSYDAFDI**WGQGT**M (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
 30 or three amino acids substitutions, and/or
 a light chain variable domain comprising
 DIQLTQSPSSLSASVGDRTITCRAS**QSISD**FLNWWYQQKPGKAPKLLIY**AA**
SSLQTGVPSRFGSGSGTEFTLTISSLQPEDLGTYYC**QQSYIMPDTFGQG**
TK (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:

EVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWM
GGIIPFGTANYAQNFQGRVTMTADTSISTAYMELSSLRSEDVAVYYCARS

5

LGGRFRYWQQT (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

10

a light chain variable domain

DVVMTQSPSLPVTGPGEPAISCRSSQSLHNSNGYNYLDWYLQKPGQSP
QLLIYLGSNRASGVPDRFSGSGSGTDFTLQISRVEAEDAGVYYCMQGLQT
PYTFGQGTK (SEQ ID NO: 24) or a fragment, variant or sequence thereof

15

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:

20

(i) QVQLQESGGGLVQAGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
LVAVITSGGSTHYADSVKGRFTISRDNAAQKTVYLQTNDLPEDTAVYY
CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 35) 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

25

(ii) QVQLVESGGGLVQAGGSLRLSCAASGSIFSGNDMSWYRQAPGKGLE
LVAVITSGGSTYADSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYY
CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 37) 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

30

(iii) QVQLQESGGGLVQPGGSLRLSCADSGSIFSGNDMAWYRRAPGVERE
LVAVITSGGSTHYADSVKGRFTISRDNAAQKTVYLQTNDLPEDTAVYY
CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 28) or a fragment,
variant or sequence thereof at least 80%, 85%, 90%, 95%, 96%, 97%,

35

98%, 99%, or 100% identical thereto and/or having two or three amino acids substitutions, or

(iv) QVQLQESGGGLVQPGGSLRLSCAAS**GSSERFTS**VAWYRQAPGKERELVAFIT**NGG**STRYTDPAVKGRFTISRDNKNTVYLMNSLKAEDTAVYY

5 **CMAGTSWGQGTQ (SEQ ID NO: 32)** 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) QVQLQESGGGLVQPGGSLRLSCADSG**SIFSG**NDMAWYRRAPGVERELVAVIT**SGG**STHYADSVKGRFTISRDNKNTVYLMNSLKAEDTAVYY

10 **CTNGRWSGDTYYAHLWGQGTQ (SEQ ID NO: 34)** 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

(vi) QVQLQESGGGLVQAGDSLRLSCADSG**SIFSG**NDMAWYRRAPGVERELVAVIT**SGG**STHYADSVKGRFTISRDNKNTVYLMNSLKAEDTAVYY

15 **CTNGRWSGDTYYAHHWGQGTQ (SEQ ID NO: 36)** 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.

20

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:

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

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

a CDR-H3 comprising

30 ACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCAC (**SEQ ID NO: 64**)

(ii) a heavy chain variable region comprising:

a CDR-H1 comprising GGAAGCTCCGAAAGATTACATCA (**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:

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

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

10 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.

15 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**
20 **TGGCAATGAC**ATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ACACTATG
CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
CGGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTT**
25 **ACTATGCCCATCACT**GGGGCCAGGGGACCCAG (**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**
30 **TGGCAATGAC**ATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACT
CGAGTTGGTCGCGGTT**ATTACTAGCGGTGGTAGTAC**ATACTATGC
AGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATTCCaAG
AAcACCcTATATCTGCAAATGAACAGCCTGAGAGCTGAGGACACGG
CCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTACT**
35 **ATGCCCATCACT**GGGGGCCAGGGGACCCTG (**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
 5 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAA**ACTA
 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTT**CG
 10 **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, or

- (iv) CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
 GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC**
 15 **TAGTAACAGTGTTGCTT**GGAAGTGGATCAGGCAGTCCCCATCGAG
 AGGCCTTGAGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGT**
ACAGTGATTATGCAATATCTGTGAAAGTTCGATTAGACATCAACCC
 AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
 CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
 20 **AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG**
(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
 25 AGGGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
 GCGAGTTGGTCGCGTT**ATTACTAGCGGTGGTAGTACAC**ACTATG
 CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
 AGAAGACCGTATATCTGCAAACGAACGACTTGAAACCTGAGGACAC
 30 GGCCGTGTATTACTGC**ACAAACGGAAGATGGTCAGGCGATACTTA**
CTATGCCCATCACTGGGGCCAGGGGACCCAG **(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
 35 GGGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAG**

ATTCACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGC
GCGAGTTGGTCGCATTTATTACTAATGGTGGTAGCACAAGATATAC
AGACCCCGTGAAGGGCCGATTACCATCTCCAGAGACAACGCCAA
GAACACGGTGTATCTGCAAATGAACAGCCTGAAAGCTGAGGACAC
 5 **GGCCGTCTATTATTGTATGGCGGGTACGTCCTGGGGCCAGGGGAC**
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**
 10 **AGGGTCTCTGAGACTCTCCTGTGCAGACTCTGGAAGCATCTTCAG**
TGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGC
GCGAGTTGGTCGCGGTTATTACTAGCGGTGGTAGTACACACTATG
CAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAATGCCC
AGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAGGACA
 15 **CGGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGATACTT**
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**
 20 **CTGGGGACTCTCTGAGACTCTCCTGTGCAGACTCTGGAAGCATCT**
TCAGTGGCAATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTG
GAGCGCGAGTTGGTCGCGGTTATTACTAGCGGTGGTAGTACACAC
TATGCAGACTCCGTGAAGGGCCGATTACCATCTCCAGAGACAAT
GCCCAGAAGACCGTATATCTGCAAACGAACGACCTGAAACCTGAG
 25 **GACACGGCCGTGTATTACTGCACAAACGGAAGATGGTCAGGCGA**
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

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

(i) **GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG**
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGTCAGAGCCTCCTGC
ATAGTAATGGATACAAC TATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
 35 **CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATTT**

- 5 TACTGCAAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCATGCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG
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, or
- (ii) GAAATTGTGCTGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
10 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGTACAGATTT
TACTGCAAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCATGCAAGGTCTACAAAGTCCCTGGACGTTTCGGCCAAGGG
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, or
- (iii) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGTCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATAGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC
20 CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATTT
TACTGCAAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCATGCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG
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, or
- (iv) GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT
CCTGATCTAT**GCTGCATCG**AGTTTACAACTGGGGTCCCTCAAGA
30 TTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCACCATAAGCA
GTCTACAACCTGAAGATTTGGGAACCTTATTACTGT**CAACAGAGTTA**
CATTATGCCCCGACACTTTTGGCCAGGGGACGAAA (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

(v) GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
 GGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCT**AATCGGGCCTC
 5 CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATT
 TACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
 TTA**CTGCATGCAAGGTCTACAGACTCCGTACACTTTT**GGCCAGGG
 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-
 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
 15 group consisting of:

(i) a heavy chain variable domain encoded by the nucleotide sequence
 comprising:

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
 GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
 20 **CAGCTATGCTAT**CAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
 TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAA**ACTA
 CGCACAGAACTTCCAGGGCAGAGTCACCATGACCGCAGACACCTC
 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
 CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
 25 **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
 GGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCT**AATCGGGCCTC
 CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATT
 30 TACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT

TACTGCATGCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG
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

- 5 (ii) a heavy chain variable domain encoded by the nucleotide sequence
comprising:

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
10 TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAA**CTA
CGCACAGAACTTCAGGGCAGAGTCACCATGACCGCAGACACCTC
CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGCCAGGGAACCCTG (**SEQ ID NO: 41**) or a sequence at
15 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
20 GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAACTATTTGGATTGGTACCTGCAGAAGCCAG
GGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCTAATCGGGCCTC**
CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGTACAGATTT
TACACTGAAAATAAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
25 TACTGCATGCAAGGTCTACAAAGTCCCT**GGACGTT**CGGCCAAGGG
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

- 30 (iii) a heavy chain variable domain encoded by the nucleotide sequence
comprising:

GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAA**CTA
35 CGCACAGAACTTCAGGGCAGAGTCACCATGACCGCAGACACCTC

- 5 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGGCCAGGGAACCCTG (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**
10 **ATAGTAATAGATACAAC TATTTGGATTGGTACCTGCAGAAGCCAG**
GGCAGTCTCCACAGCTCCTGATCTATTT**GGGTTCTAATCGGGCCTC**
CGGGGTCCCTGACAGGTTCA GTGGCAGTGGATCAGGCACAGATTT
TACACTGAAAATCAGCAGAGTGGAGGCTGAGGATGTTGGGGTTTAT
TACTGCAT**GCAAGCTCTACAACTCCGTACACTTTTGGCCAGGGG**
15 **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
- (iv) a heavy chain variable domain encoded by the nucleotide sequence
comprising:
20 GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGG
GTCCTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAG**
CAGCTATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCT
TGAGTGGATGGGAGGG**ATCATCCCTATCTTTGGTACAGCAAACTA**
CGCACAGAACTTCAGGGCAGAGTCACCATGACCGCAGACACCTC
25 CATAAGCACAGCCTACATGGAGCTGAGCAGCCTGAGATCTGAGGA
CACGGCCGTGTATTACTGT**GCGAGATCGTTGGGTGGGAGATTTCG**
CTACTGGGGGCCAGGGAACCCTG (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
30 a light chain variable domain encoded by the nucleotide sequence
comprising:
GATGTTGTGATGACTCAGTCTCCACTCTCCCTGCCCGTCACCCCTG
GAGAGCCGGCCTCCATCTCCTGCAGGTCTAGT**CAGAGCCTCCTGC**
ATAGTAATGGATACAAC TATTTGGATTGGTACCTGCAGAAGCCAG
35 **GGCAGTCTCCACAGCTCCTGATCTATTTGGGTTCTAATCGGGCCTC**

CGGGGTCCCTGACAGGTTCACTGGCAGTGGATCAGGCACAGATT
TACACTGCAAATCAGCAGAGTGGAGGCTGAGGATGCTGGGGTTTA
TACTGCA**TGCAAGGTCTACAGACTCCGTACACTTTT**GGCCAGGG
GACCAAG (**SEQ ID NO: 61**) or a sequence at least 80%, 85%, 90%,
5 95%, 96%, 97%, 98%, 99%, or 100% identical thereto and/or having 10-
20 nucleic acid substitutions; and

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

CAGGTACAGCTGCAGCAGTCAGGTCCAGGGCTGGTGAAGCCCTC
10 GCAGACCCTCTCACTCACCTGTGCCATCTCC**GGGGACAGTATCTC**
TAGTAACAGTGTGCTTGGAACTGGATCAGGCAGTCCCCATCGAG
AGGCCTTGAGTGGCTGGGAAGGAC**ATACTACAGGTCCAAGTGGT**
ACAGTGATTATGCAATATCTGTGAAAGGTCGATTAGACATCAACCC
AGACACATCCAAGAACCAGTTCTCCCTGCAGCTGAACTCTGTGACT
15 CCCGAGGACACGGCTGTGTATTATTGT**GCAAGAGAAGTTGAGGGC**
AGCAGCTATGATGCTTTTGATATCTGGGGCCAAGGGACAATG

(**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

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

GACATCCAGTTGACCCAGTCTCCATCCTCCCTGTCTGCATCTGTAG
GAGACAGAGTCACCATCACTTGCCGGGCAAGT**CAGAGTATTAGCG**
ACTTTTTAAATTGGTACCAGCAGAAACCAGGTAAAGCCCCGAAGCT
25 CCTGATCTAT**GCTGCATCG**AGTTTACAAACTGGGGTCCCCTCAAGA
TTCGGTGGCAGTGGATCTGGGACAGAATTCACTCTCACCATAAGCA
GTCTACAACCTGAAGATTTGGGAACCTTATTACTGT**CAACAGAGTTA**
CATTATGCCCCGACACTTTTGGCCAGGGGACGAAA (**SEQ ID NO: 59**)

or a sequence at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, or
30 100% 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**
ATGACATGTCCTGGTACCGCCAGGCTCCAGGGAAGGGACTCGAGTTG
GTGCGCGTT**ATTACTAGCGGTGGTAGTACA**TACTATGCAGACTCCGTG
5 AAGGGCCGATTACCATCTCCAGAGACAATTCCaAGAAcACCCtTATATC
TGCAAATGAACAGCCTGAGAGCTGAGGACACGGCCGTGTATTACTGCA
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
10 nucleic acid substitutions; or
- (ii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGG
GCTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
GGTCGCGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG
15 TGAAGGGCCGATTACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCACTGG
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
20 10-20 nucleic acid substitutions, or
- (iii) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAG
GGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
GGTCGCGTT**ATTACTAGCGGTGGTAGTACA**CACTATGCAGACTCCG
25 TGAAGGGCCGATTACCATCTCCAGAGACAATGCCCAGAAGACCGTAT
ATCTGCAAACGAACGACTTGAAACCTGAGGACACGGCCGTGTATTACT
GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCACTGG
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
30 10-20 nucleic acid substitutions, or
- (iv) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGGG
GGTCTCTGAGACTCTCCTGTGCAGCCTCT**GGAAGCTCCGAAAGATTC**
ACATCAGTGGCCTGGTACCGCCAGGCTCCAGGAAAGGAGCGCGAGTT
GGTCGCATTT**ATTACTAATGGTGGTAGCACA**AGATATACAGACCCCGT
35 GAAGGGCCGATTACCATCTCCAGAGACAACGCCAAGAACACGGTGT

ATCTGCAAATGAACAGCCTGAAAGCTGAGGACACGGCCGTCTATTATT
 GTATGGCGGGTACGTCTCTGGGGCCAGGGGACCCAG (**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

5 (v) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTGGTGCAGCCTGGAG
 GGTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
 GGTCGCGGTT**ATTACTAGCGGTGGTAGTACAC**ACTATGCAGACTCCG
 TGAAGGGCCGATTCAACATCTCCAGAGACAATGCCCAGAAGACCGTAT
 10 ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCTCTGG
 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

15 (vi) CAGGTGCAGCTGCAGGAGTCTGGGGGAGGATTGGTGCAGGCTGGGG
 ACTCTCTGAGACTCTCCTGTGCAGACTCT**GGAAGCATCTTCAGTGGCA**
ATGACATGGCCTGGTACCGCCGGGCTCCAGGGGTGGAGCGCGAGTT
 GGTCGCGGTT**ATTACTAGCGGTGGTAGTACAC**ACTATGCAGACTCCG
 TGAAGGGCCGATTCAACATCTCCAGAGACAATGCCCAGAAGACCGTAT
 20 ATCTGCAAACGAACGACCTGAAACCTGAGGACACGGCCGTGTATTACT
GCACAAACGGAAGATGGTCAGGCGATACTTACTATGCCCATCACTGG
 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,
 25 GAGGTCCAGCTGGTGCAGTCTGGGGCTGAGGTGAAGAAGCCTGGGTC
 CTCGGTGAAGGTCTCCTGCAAGGCTTCT**GGAGGCACCTTCAGCAGCT**
ATGCTATCAGCTGGGTGCGACAGGCCCTGGACAAGGGCTTGAGTGG
 ATGGGAGGGAT**CATCCCTATCTTTGGTACAGCA**AACTACGCACAGAAC
 TTCCAGGGCAGAGTCACCATGACCGCAGACACCTCCATAAGCACAGC
 30 CTACATGGAGCTGAGCAGCCTGAGATCTGAGGACACGGCCGTGTATTA
CTGTGCGAGATCGTTGGGTGGGAGATTTGCTACTGGGGCCAGGGAA
 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
 GACCCTCTCACTCACCTGTGCCATCTCCGGGGACAGTATCTCTAGTAA
CAGTGTTGCTTGGAAGTGGATCAGGCAGTCCCCATCGAGAGGCCTTG
 AGTGGCTGGGAAGG**ACATACTACAGGTCCAAGTGGTACAGT**GATTAT
 5 GCAATATCTGTGAAAGGTGCGATTAGACATCAACCCAGACACATCCAAG
 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
 10 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.
- 15 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.
- 20 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.
- 25 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.
- 30 17. The cell of any one of the preceding claims, wherein the multi-specific 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 an immune cell for example 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 binds to HER2 and
5 secretes an inducible immune cell engager that targets EpCAM and CD3 (HE CAR-BiTE T).
21. A polynucleotide encoding the cell of any one of the preceding claims.
- 10 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,
15 comprising introducing the polynucleotide of claim 21 into a 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
20 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.

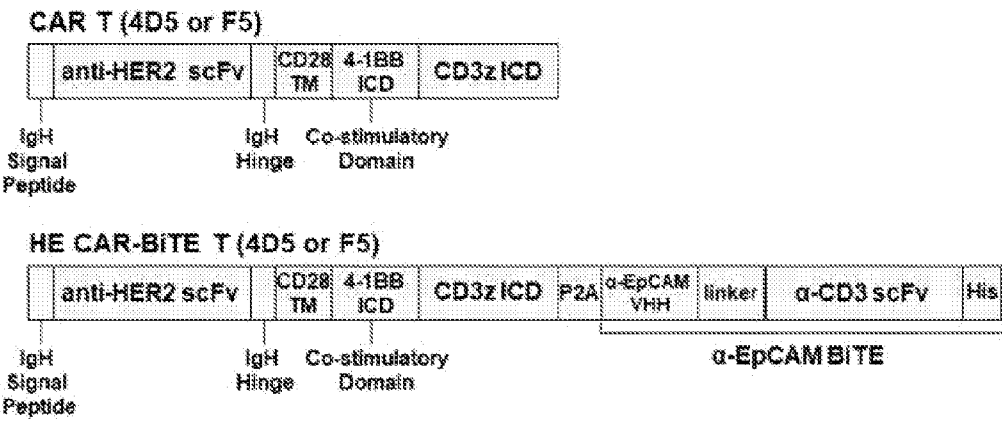


Figure 1A

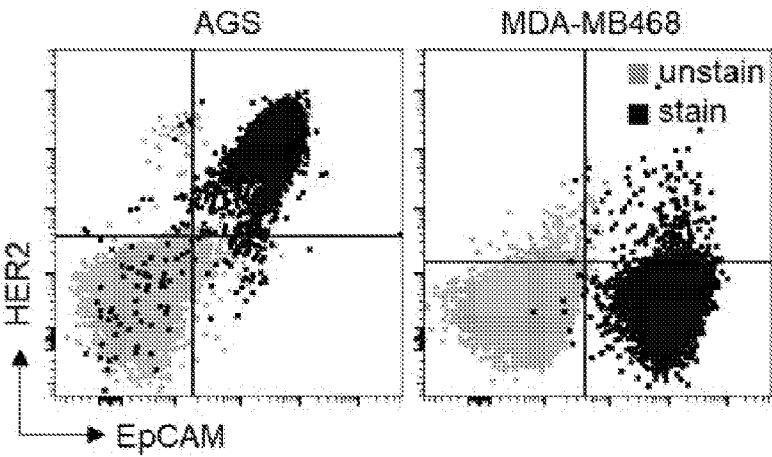


Figure 1B

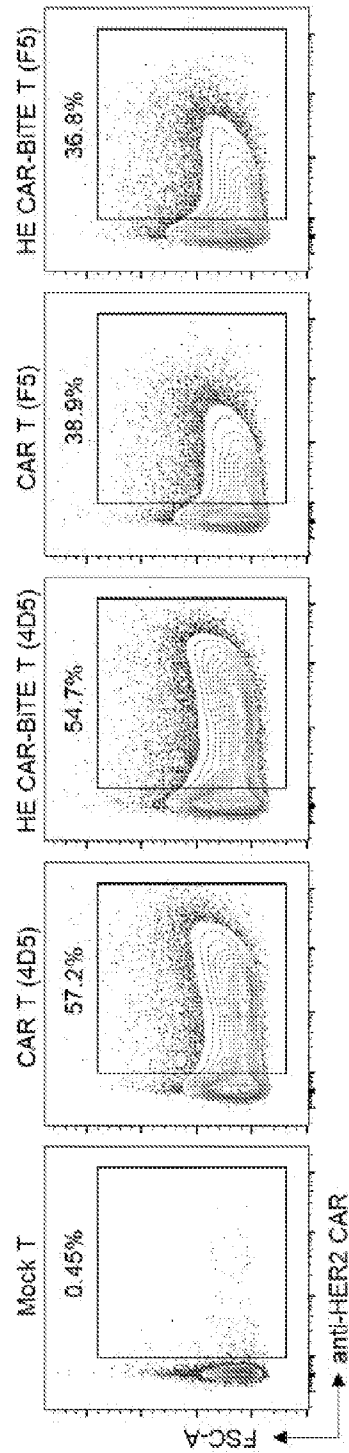


Figure 1C

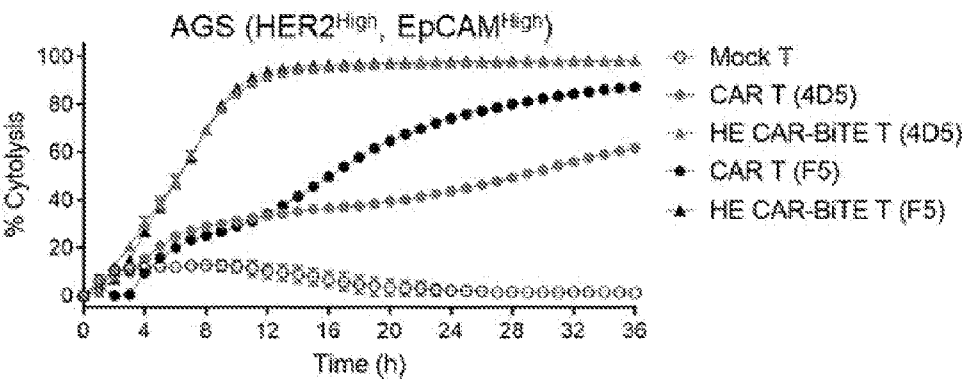


Figure 1D

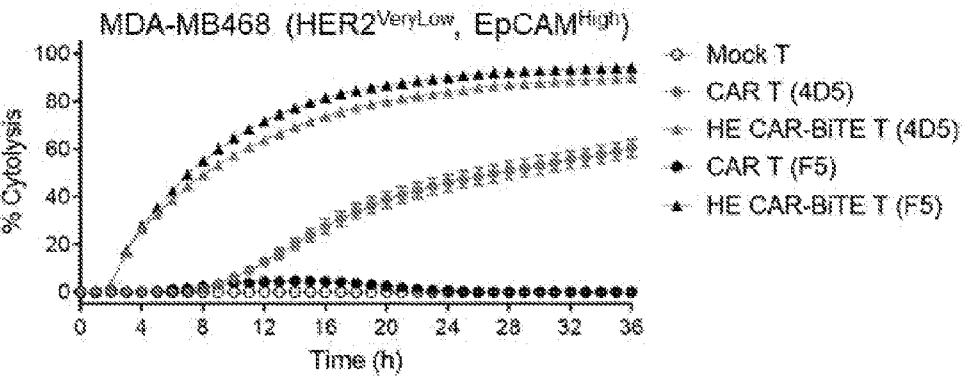


Figure 1E

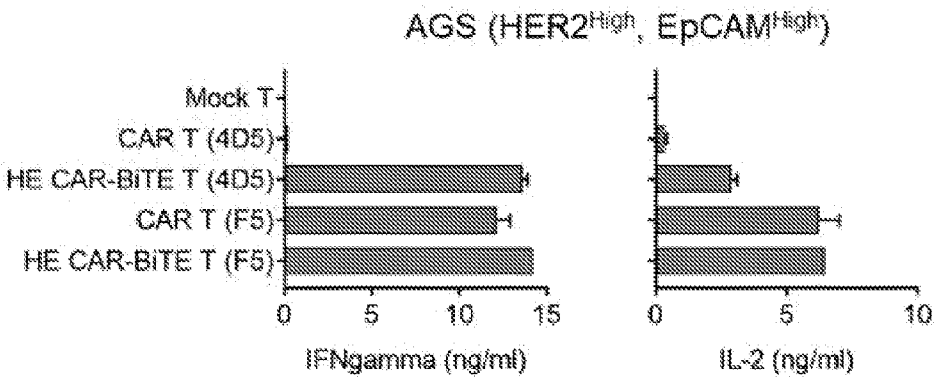


Figure 1F

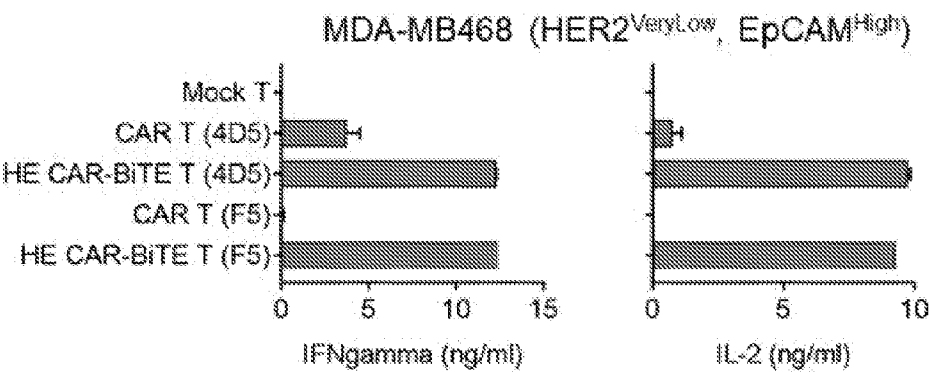


Figure 1G

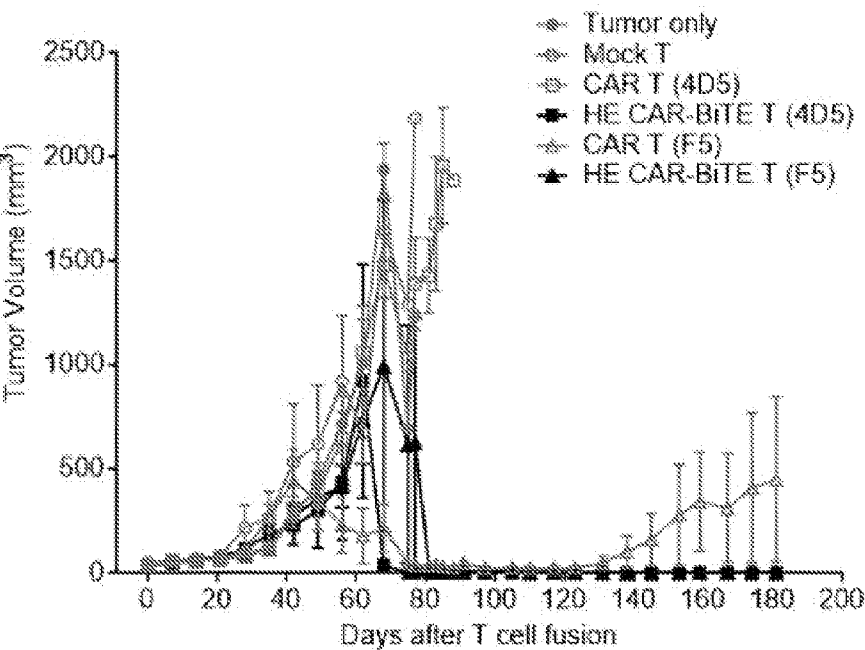


Figure 2A

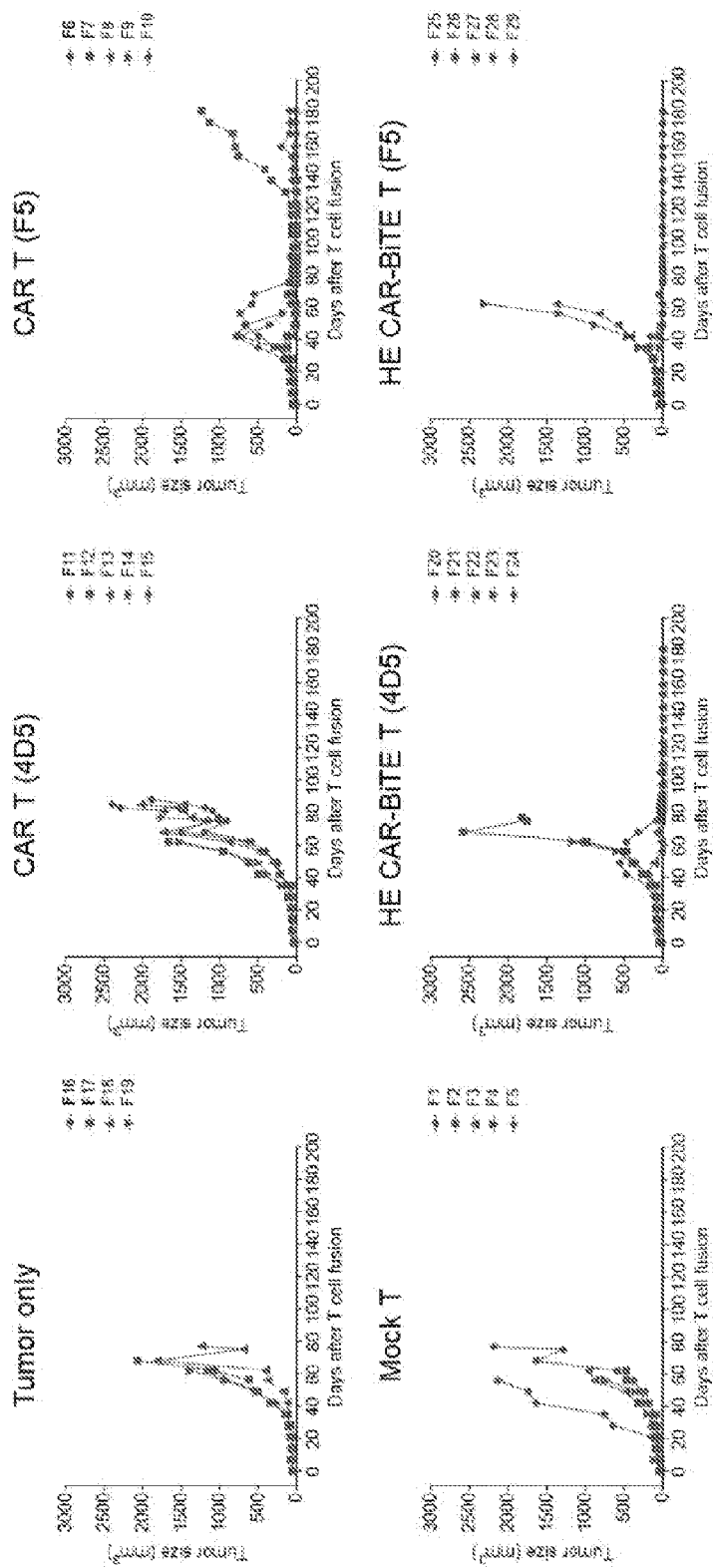


Figure 2B

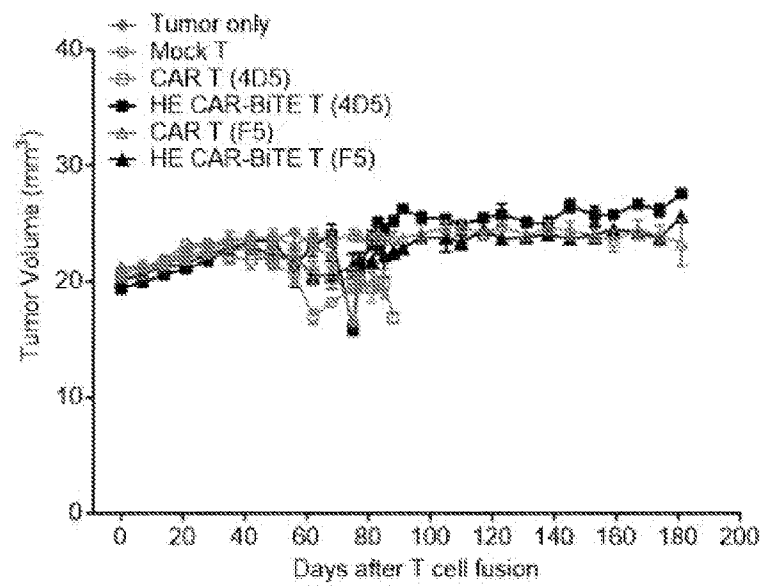


Figure 2C

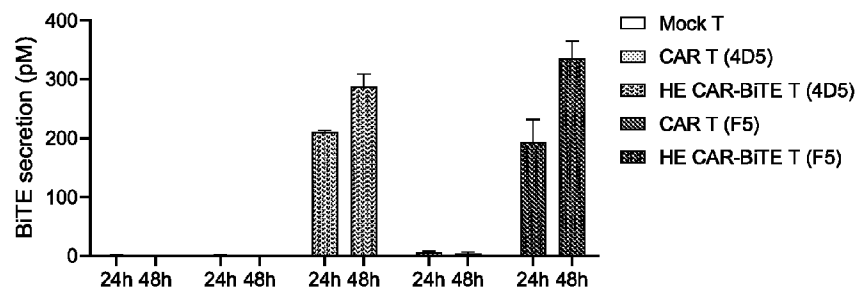


Figure 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SG2024/050384

A. CLASSIFICATION OF SUBJECT MATTER

See Supplemental Box

According to International Patent Classification (IPC)

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

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

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

DATABASES: FAMPAT, PATSNAP, MEDLINE, EMBASE, BIOSIS

SEARCH TERMS: EpCAM, CD326, HER2, ERBB2, Neu, Chimeric antigen receptor, CART, NKCAR, bispecific, multispecific, bite, T cell engager, secreting, bicistronic and related terms.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HOSKING, M. ET AL., Abstract 116: Multi-antigen targeting of heterogenous solid tumors using CAR T cells secreting bi-specific T-cell engagers. <i>Journal for ImmunoTherapy of Cancer</i> , Vol. 8 (Suppl. 3); Page A72. 10 December 2020 [Retrieved on 2024-07-12] <DOI: 10.1136/JITC-2020-SITC2020.0116> (Whole document especially Figure 1D)	1-26
X	WO 2022/018262 A1 (CELLECTIS SA) 27 January 2022 (Page 27, Line 7; Page 32, Lines 1-4; Examples; Figures 2-3, 9; Table 13)	1-26
X	WO 2022/109611 A1 (SIMCERE INNOVATION INC) 27 May 2022 (Paragraphs 0007-0016, 0019, 0150, 0194-0195, 0260-0261, 0310, 0324, 0327-0328)	1-26

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

*Special categories of cited documents:

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

"D" document cited by the applicant in the international application

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search
12/07/2024 (day/month/year)

Date of mailing of the international search report
26/07/2024 (day/month/year)

Name and mailing address of the ISA/SG

Intellectual Property Office of Singapore
1 Paya Lebar Link, #11-03
PLQ 1, Paya Lebar Quarter
Singapore 408533

Email: pct@ipos.gov.sg

Authorized officer

Tan Terence (Dr)

IPOS Customer Service Tel. No.: (+65) 6339 8616

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SG2024/050384

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 113481165 A (SHANDONG BOAN BIOTECHNOLOGY CO LTD) 8 October 2021 (Page 5) of the machine translation	1-26
X	WO 2021/050591 A1 (CYTOIMMUNE THERAPEUTICS INC) 18 March 2021 (Paragraphs 0008, 0018, 0336; Examples).	1-26
X	US 2022/0331416 A1 (LI, Z. ET AL.) 20 October 2022 (Paragraphs 0033-0046, 0216-0217, 0282, 0291-0292, 0375; Examples 10-11; Figure 3)	1-26

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SG2024/050384

Supplemental Box
(Classification of Subject Matter)

Int. Cl.
C12N 5/10 (2006.01)
C07K 14/725 (2006.01)
C07K 16/28 (2006.01)
A61K 35/17 (2015.01)
A61P 35/00 (2006.01)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/SG2024/050384

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☐ forming part of the international application as filed.
 - b. ☐ furnished subsequent to the international filing date for the purposes of international search (Rule 13~~ter~~.1(a)),
 - ☐ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

Although a sequence listing has been furnished, it was not used for the purposes of this search.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/SG2024/050384

Note: This Annex lists known patent family members relating to the patent documents cited in this International Search Report. This Authority is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2022/018262 A1	27/01/2022	EP 4185616 A1 US 2023/0248825 A1	31/05/2023 10/08/2023
WO 2022/109611 A1	27/05/2022	US 2023/0277592 A1 JP 2023550148 A US 2022/0160766 A1 EP 4247850 A1 TW 202237649 A	07/09/2023 30/11/2023 26/05/2022 27/09/2023 01/10/2022
CN 113481165 A	08/10/2021	NONE	
WO 2021/050591 A1	18/03/2021	KR 20220070228 A JP 2022547220 A EP 4028032 A1 CO 2022004603 A2 US 2022/0281982 A1 CA 3150550 A1 TW 202124441 A IL 291186 A BR 112022004407 A2 AU 2020344551 A1 CN 114901294 A MX 2022002914 A	30/05/2022 10/11/2022 20/07/2022 08/07/2022 08/09/2022 18/03/2021 01/07/2021 01/05/2022 21/06/2022 14/04/2022 12/08/2022 14/06/2022
US 2022/0331416 A1	20/10/2022	WO 2021/048724 A1 CA 3150839 A1 CN 112480263 A AU 2020345133 A1 JP 2022548623 A EP 4028525 A1 CN 114616337 A	18/03/2021 18/03/2021 12/03/2021 28/04/2022 21/11/2022 20/07/2022 10/06/2022