Genomic profiling and pre-clinical modelling of breast cancer leptomeningeal metastasis reveals acquisition of a lobular-like phenotype

Amanda Fitzpatrick[1,2], Marjan Iravani[1], Adam Mills[1], David Vicente[1], Thanussuyah Alaguthurai[3], Ioannis Roxanis[1], Nicholas C. Turner[1,4], Syed Haider[1], Andrew N.J. Tutt[1,3,5] and *Clare M. Isacke[1]

[1]The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London, UK.

[2]Comprehensive Cancer Centre, School of Cancer & Pharmaceutical Sciences, King's College London, UK

[3]Breast Cancer Now Research Unit, Guy's Hospital, King's College London, London, UK

[4]The Royal Marsden NHS Foundation Trust; London, UK

[5]Oncology and Haematology Directorate, Guy's and St Thomas' NHS Foundation Trust, London, UK

***Corresponding Author**: Clare M. Isacke, The Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, UK. Telephone +44 20 7153 5510, Email clare.isacke@icr.ac.uk

Abstract

Breast cancer leptomeningeal metastasis (BCLM), where tumour cells grow along the lining of the brain and spinal cord, is a devastating development for patients. Investigating this metastatic site is hampered by difficulty in accessing tumour material. Here, we utilise cerebrospinal fluid (CSF) cell-free DNA (cfDNA) and CSF disseminated tumour cells (DTCs) to explore the clonal evolution of BCLM and heterogeneity between leptomeningeal and extracranial metastatic sites. Somatic alterations with potential therapeutic actionability were detected in 81% (17/21) of BCLM cases, with 19% detectable in CSF cfDNA only. BCLM was enriched in genomic aberrations in adherens junction and cytoskeletal genes, revealing a lobular-like breast cancer phenotype. CSF DTCs were cultured in 3D to establish BCLM patient-derived organoids, and used for the successful generation of BCLM *in vivo* models. These data reveal that BCLM possess a unique genomic aberration profile and highlights potential cellular dependencies in this hard-to-treat form of metastatic disease.

Introduction

Despite advances in effective treatment of visceral metastatic disease, central nervous system (CNS) metastasis increasingly limits patient survival[1,2]. Leptomeningeal metastasis, where tumour cells spread over the meninges lining the brain and spinal cord and shed into the surrounding cerebrospinal fluid (CSF), most commonly occurs in breast cancer, lung cancer and melanoma[3,4]. Leptomeningeal metastasis causes debilitating neurological symptoms with often rapid clinical deterioration. Median survival in breast cancer leptomeningeal metastasis (BCLM) is 3.4 – 5.4 months despite current treatments[4-6]. Published BCLM cohorts show an enrichment for invasive lobular cancer (ILC)[7-9], the most common 'special' subtype of breast cancer[10,11] which is characterised histologically by discohesive tumour cells arranged in single files or as individual cells, driven predominantly by mutational inactivation of *CDH1* (E-cadherin)[12,13]. By contrast, the more common invasive carcinoma of no special type (IC-NST)

are E-cadherin proficient and exist in nests or sheets of cells with preserved cell-cell adhesions[14].

Despite an increased knowledge of the genomic events that evolve in metastatic breast and other solid cancers[15-20], the biology of leptomeningeal metastasis remains poorly understood since the leptomeninges are a rarely-biopsied site. It is often cited that BCLM is a late manifestation of widespread metastatic disease, however, to date no studies have interrogated BCLM tumour evolution to determine the timing of BCLM metastatic seeding. The study of the molecular landscape and evolution of leptomeningeal malignancy is key to understanding what drives this disease and identifying targetable vulnerabilities.

Previous studies have demonstrated that CSF is an abundant source of tumour-derived cell-free DNA (cfDNA) in the setting of CNS-predominant malignancies, while plasma cfDNA is predominantly derived from extracranial sites[21,22]. Therefore we have utilised CSF cfDNA to interrogate the genomics and evolution of BCLM, and generated *in vitro* and *in vivo* models of BCLM by expansion of the scarce CSF disseminated tumour cells (DTCs). We demonstrate that BCLM shows early divergent evolution from the primary tumour and a distinct genomic profile compared to the primary tumour and other metastatic sites. Particularly noteworthy is the enrichment in BCLM of adherens junction and cytoskeletal genomic alterations, predicted to result in the acquisition of features associated with lobular breast cancers.

Results

Whole exome sequencing reveals unique genomic events in BCLM

Previously we described the prospective collection of CSF and plasma cfDNA from a cohort of BCLM patients, and demonstrated the utility of using ultra-low pass whole genome sequencing (ulpWGS) to measure the tumour-derived (ctDNA) fraction for BLCM diagnosis and therapy response monitoring[21]. Here we have investigated the genomic landscape of BCLM by whole exome sequencing (WES) in 21 patients for whom both matched CSF and plasma cfDNA were available along with archival samples of 18 matched primary tumour samples and 8 available metastatic site samples (**Fig. 1a,b; Table 1; Supplementary Table 1**). A threshold of ≥10% ctDNA fraction (assessed by ulpWGS[21]) was used to select cfDNA samples for whole exome sequencing (WES). As previously reported[21,22], when metastatic disease is restricted to the CNS and/or extracranial metastases are controlled, ctDNA content in plasma is low. Consequently although all 21 CSF cfDNA samples had ≥10% ctDNA fraction, only 11 of the plasma cfDNA samples met this criteria (**Supplementary Fig. 1a)**. Plasma-derived cfDNA undergoing WES was sequenced to a higher coverage than CSF cfDNA and tumour tissues to mitigate for the lower ctDNA content (**Supplementary Fig. 1a,b; Supplementary Table 2**).

To assess the inter-tumoral heterogeneity between BCLM (CSF cfDNA), the primary tumour and extracranial metastases (plasma cfDNA or metastatic tissue), a multi-caller variant identification pipeline and post-hoc refinements approach was implemented to WES outputs to prioritise high-confidence mutations (see Methods; filtered non-silent somatic variants across all samples shown in **Supplementary Table 3**). Overall, tumour mutational burden (TMB) was as expected for the breast cancer genome[23], median CSF cfDNA 2.0, plasma cfDNA 2.0, primary tumour 1.3, metastasis 1.2 mutations/Mb with no significant difference between sample type (**Supplementary Fig. 1c,d**). Samples exhibiting high TMB are discussed later.

Mutational signature analysis was performed using the WES-identified single nucleotide variants (SNVs) (median of 96 filtered SNV per sample) (**Supplementary Fig. 2a-d; Supplementary Tables 4 and 5**). This displayed a higher contribution from chemotherapy-related single base substitution (SBS) signatures in plasma cfDNA compared to CSF cfDNA (**Fig. 1c**), which could reflect the 'sanctuary site' phenomenon whereby the CNS has a lower exposure to chemotherapy due to impedance by the blood-brain-barrier[24]. In keeping with prior findings in metastatic *vs.* primary breast cancer[20], CSF cfDNA samples had higher SBS13 (APOBEC) signature contribution than primary tumours, and lower SBS1 (clock-like/FFPE artefact) and SBS23 (unknown aetiology) signatures (**Supplementary Table 5**).

Indicative of a distinct genomic landscape in BCLM, 47.8% of mutations across CSF cfDNA and primary tumour were detected only in CSF cfDNA (**Fig. 1d**), with 20.4% detected only in primary tumour. Likewise, indicating a distinct mutational repertoire between BCLM and extracranial metastases, 24.6% of cfDNA mutations across the 11 paired CSF and plasma samples were detected only CSF cfDNA, with 16.8% detected only in plasma cfDNA (**Fig. 1e**). There was no significant correlation between the proportion of unique variants in CSF cfDNA *vs.* primary tumour or between CSF cfDNA *vs.* plasma cfDNA with time to development of BCLM nor to the number of previous treatment lines (**Supplementary Fig. 3)**.

<u>BCLM subclones seed early from the primary tumour</u>

To investigate metastatic evolution of BCLM, variant allele frequencies and allele-specific copy number alterations (CNA; see below for CNA analysis) were used to infer mutational subclones in CSF and plasma cfDNA, primary tumour and metastatic tissues where sequenced. Clonal ordering and estimation of cancer cell fractions using ClonEvol allowed visualisation of BCLM metastatic evolution trees (**Fig. 2; Supplementary Fig. 4**). Clonal modelling failed in one case (KCL625) due to the primary tumour and CSF cfDNA not sharing a common founder clone.

Within the limitation of WES, which may underestimate subclonal composition compared to whole genome sequencing, BCLM evolution analyses revealed: (a) BCLM metastatic seeding occurs early in primary tumour subclone evolution, predominantly at the stage of the founder clone (clone 1). In all 16 cases with primary tumour evaluated by clonal modelling, additional primary tumour subclone evolution occurred following BCLM metastatic divergence, and primary tumours were composed of 2 - 4 subclones. (**Fig 2a,b,c; Supplementary Fig. 4a-h, k-n,q; Supplementary Table 6)**. (b) For the 11 cases with matched CSF and plasma cfDNA, divergent evolution between these occurred in 9 cases (**Fig 2a,b; Supplementary Fig. 4a-e,j,p; Supplementary Table 6)**. Shared metastasis-specific branches were often seen, but CSF cfDNA, plasma cfDNA and other metastatic sites possessed distinct mutational subclones implying ongoing evolution at the leptomeningeal and extracranial sites. 2 cases showed linear co-evolution between CSF and plasma cfDNA (**Supplementary Fig. 4g,o; Supplementary Table 6)**.

Phylogenetic tree construction allowed depiction of the metastatic seeding patterns in BCLM, with the limitation that not all metastatic sites were sampled. In the assessable cases, metastatic seeding was mostly monophyletic (**Fig 2a,b; Supplementary Fig. 4a-e,g**; **Supplementary Table 6)** where BCLM and other metastatic sites were derived from the same subclone within the primary tumour. However, one case showed polyphyletic seeding with metastatic sites originating from primary tumour at different timepoints in its evolution (**Fig 2c; Supplementary Table 6**). As illustrated, KCL566 **(Fig. 2a)** displayed a monophyletic seeding pattern where metastases arose from the founding clone of the primary tumour (clone 1) and shared a common evolution branch (clone 2) before displaying divergent evolution between lymph node metastasis, plasma and CSF cfDNA. At the time of CSF and plasma sampling, progressive disease was present at extracranial sites (peritoneum, subcutaneous tissues and bone), potentially driven by *PTEN* p.L23S mutation, while CSF cfDNA depicted significant further parallel evolution, with two further subclones of high mutation count (clones 4 and 9; 429 and 371 mutations, respectively), displaying a *CDH1* truncating mutation p.L139X and

NF1 frameshift mutation. RMH010 **(Fig. 2b)** shows early monophyletic seeding of all metastases from the primary tumour (clone 1), with divergent evolution between liver metastases (6 years post primary diagnosis) and CSF and plasma cfDNA (10 years post primary diagnosis, progressive disease in leptomeninges, peritoneum, liver, bone and orbital soft tissues) implying ongoing tumour evolution at metastatic sites. KCL553 (**Fig. 2c**) showed a polyphyletic pattern of metastatic seeding with ovarian metastasis seeding from the founding clone of the primary tumour containing *ARID1A* frameshift mutation (clone 1) and BCLM (CSF cfDNA) seeding after a clonal sweep in the primary tumour comprising *TP53* frameshift mutation (clone 2). Further subclone evolution in the CSF cfDNA was apparent with development of a *KMT2D* missense variant.

7/21 BCLM cases had co-existing brain parenchymal metastasis (**Table 1**), the majority developing metachronous to BCLM. Archived brain metastasis tissue, resected 3 years prior to BCLM development, was available for RMH011 (**Supplementary Fig. 4i**). Clonal modelling revealed that BCLM and brain metastasis shared two common ancestral clones, however displayed divergent evolution with unique subclones at each site. No primary tumour was available for sequencing to determine the metastatic seeding pattern of this case.

Cancer driver gene aberrations enriched in BCLM

We identified 35 cancer driver genes (defined in Methods) which were frequently aberrated in CSF cfDNA by non-silent somatic mutation or high-level copy number change (deep deletion or amplification) (**Fig. 3; Supplementary Fig. 5)**. Driver gene aberrations occurring exclusively in CSF cfDNA and not shared with primary/metastatic tumour or plasma cfDNA (where sequenced) were found in genes involved in histone modification (*KAT6B, KMT2D, NUTM2A*) and microtubule formation (*PDE4DIP*). Frequently aberrated cancer driver genes in plasma cfDNA are shown in **Supplementary Table 7,** however, given the small number of plasma cfDNA samples sequenced, the identification of extracranial metastasis driver genes is better represented by larger published cohorts[18,25,26].

Next, we examined the alteration rate of the BCLM altered genes to a publicly available sequencing dataset of 216 non-BCLM metastatic breast cancer samples[27], herein known as MBC cohort (**Fig. 3; Supplementary Fig. 5)**. This comparison revealed 10 genes, *CDH1, MUC16, TBX3, ARHGEF10, MDM2, WRN, CTNNA1, NUTM2A, PSIP1* and *MLH1*, with a significantly different alteration rate in BCLM compared to the MBC cohort ($p \leq 0.05$, Chi-square test) (**Supplementary Table 8**).

Copy number differences between BCLM and matched samples

Copy number alteration (CNA) across the genome was determined using WES data, revealing common breast cancer events across samples such as 1q and 8q gains, and 8p and 16q losses[28] (**Fig. 4a; Supplementary Table 9**). Based on the proportion of CNA events across the exome in each sample type, we identified genomic loci that differ in CNA frequency between CSF cfDNA and matched primary tumours and between CSF and matched plasma cfDNA samples (**Fig. 4b; Supplementary Fig. 6**). Compared to primary tumours, CSF cfDNA more frequently lost 13q regions containing *PCDH9* and *PCDH17* (members of the cadherin superfamily) and *KLHL1*, an actin binding protein with a role in cytoskeleton reorganisation. Broad 16q regions containing the classical cadherin *CDH1* and other cadherins showed more frequent loss in CSF than plasma cfDNA samples. Gained regions in CSF cfDNA compared to plasma included 1q21.3 containing the intermediate filament-associated proteins *TCHH* and *FLG2,* and S100 family genes which have been implicated in poor prognosis breast cancer relapse[29].

Gene set enrichment of recurrently altered genes in BCLM

The ability of tumour cells to gain access to the leptomeningeal space via the blood-brain/blood-CSF-barrier, and successfully interact with the unique microenvironment in this space, may require genomic alterations not currently characterised as cancer driver genes. Consequently, we took an unbiased approach of performing gene set enrichment analysis of

all frequently altered (by non-silent mutation, amplification or deep deletion) genes in CSF or plasma cfDNA compared to primary tumours (**Fig. 4c,d; Supplementary Fig. 7**). CSF and plasma cfDNA samples shared alterations in the gene ontology (GO) terms 'endomembrane system organisation', and 'fibroblast growth factor response'. By contrast, GO terms showing enrichment in CSF cfDNA compared to primary tumours but not in plasma cfDNA were; 'positive chemotaxis', 'myelination' and 'supramolecular fiber organization', the latter encompassing key cytoskeletal scaffold components such as actin filaments, intermediate filaments and microtubules, and their regulators. Within the 'supramolecular fiber organization' term, *CTNNA1, MYO15A, DMTN and SPTA1* are of interest due their role in regulating adherens junctions and the actin cytoskeleton. Within 'myelination', CSF cfDNA samples were enriched for genomic alterations in *ARGHEF10* (discussed above), and genes with roles in cell-cell adhesion including neurofascin *(NFASC)* and teneurin transmembrane protein 4 (*TENM4*). Within 'positive chemotaxis', amplifications were found in scribble planar cell polarity protein (*SCRIB*), previously demonstrated to promote cell migration and metastasis[30-32]. Genomic alteration in beta-defensin genes (*DEFB4A, DEFB103B, DEFB104B*), antimicrobial peptides with multiple proposed functions including regulation of cancer cell migration[33], were also enriched in BCLM samples.

Enrichment of adherens junction components and cytoskeletal aberrations in BCLM

The high rate of *CDH1* (E-cadherin) mutations (52%) in BCLM cfDNA is consistent with the enrichment of ILCs in this cohort since deleterious *CDH1* mutations are an early driver event in the majority (63%) of ILCs, leading to defective adherens junctions[12,13]. *CDH1* aberrations are rare in non-lobular breast cancers (2.3% TCGA IC-NST cases). Mutations and copy number aberrations in *CTNNA1* (α-catenin), an indispensable adherens junction component which links the cadherin/β-catenin complex to the underlying actin cytoskeleton, has been suggested as an alternative mechanism to becoming 'lobular'[34], however *CTNNA1* genomic aberrations are rare in both ILC and non-lobular breast cancers (0.5% and 1.5% respectively, TCGA). In our BCLM cohort, we identified deleterious alterations of *CDH1* or *CTNNA1* in 55%

of non-lobular BCLM cases (**Fig. 3; Supplementary Fig. 5**). For example, KCL566 (IC-NST) harboured a stop-gain *CDH1* mutation plus *CDH1* LOH in CSF cfDNA, with neither aberration being detected in the matched plasma, primary tumour or metastatic lymph node samples. KCL617, arising from a breast cancer of apocrine morphology, displayed a frameshift *CDH1* mutation in both primary tumour and CSF cfDNA, with additional *CDH1* loss of heterozygosity (LOH) in the CSF cfDNA. KCL622 and KCL625 (both IC-NST) displayed *CTNNA1* focal deep deletions in CSF cfDNA, while *CTNNA1* had neutral copy number status in matched primary tumours. KCL658 (IC-NST) displayed a *CTNNA1* stop-gain mutation in CSF cfDNA, shared with the matched plasma cfDNA and KCL650 (IC-NST) displayed a *CTNNA1* focal deep deletion in CSF cfDNA, while the matched plasma cfDNA showed shallow copy number loss only. For KCL658 and KCL650, however, no primary tumour was available for sequencing to determine the timing of *CTNNA1* aberration. *CDH1* and *CTNNA1* aberrations were often BCLM unique events and mutually exclusive (p = 0.0352; two-sided Fisher's exact test) implying that adherens junction defects occur during leptomeningeal metastasis evolution.

In addition to the adherens junction components *CDH1* and *CTNNA1*, sequencing analysis also revealed aberrations in cytoskeletal components and their regulators **(Fig. 4d; Supplementary Fig. 7)**. For example, a member of the RhoGEF family, *ARHGEF10* was frequently aberrated in BCLM CSF cfDNA (23% of cases), comprising four deep deletions (one focal event ~1.9 Mb and 3 broader, but not arm-level, events of between 6.7 - 6.9 Mb) **(Supplementary Table 9)** and a BCLM unique missense mutation (p.A1100P) with predicted pathogenicity by SIFT, Polyphen2 and CADD prediction scores (**Supplementary Table 3**). A further, nine CSF cfDNA samples demonstrated *ARHGEF10* heterozygous losses. *ARHGEF10* facilitates activation of Rho and downstream Rho kinase (ROCK) to promote actomyosin contractility[35]. Excessive actomyosin contractility has been reported to impair adhesion to, and migration on, collagen I and is deleterious to ILC development[36], consequently it is tempting to speculate that loss of *ARGHEF10* provides a pro-survival mechanism for BCLM. Myosin 15A (*MYO15A)*, encoding an actin-based motor molecule, was

aberrated in 6/21 CSF cfDNA samples (one truncating mutation and four missense mutations with predicted pathogenicity), often accompanied by loss-of-heterozygosity (LOH). Dematin actin binding protein (*DMTN*), a regulator of cytoskeleton remodelling[37,38], was frequently aberrated in CSF cfDNA. *DMTN* downregulation has been reported to promote colorectal cancer metastasis through activation of Rac1, a key cytoskeletal regulator[39]. Although the *DMTN* copy number events were part of broader 8p deletion events (18 - 29 Mb), frameshift and missense mutations were found in 3 CSF cfDNA samples and were predominantly BCLM unique mutations. Finally, spectrin-alpha 1 (*SPTA1*), a scaffold protein linking the plasma membrane to the actin cytoskeleton, was aberrated in 4/21 BCLM CSF cfDNA samples by somatic missense mutations, of which three had high pathogenicity predictions.

Actionable mutations discovered by CSF cfDNA sequencing

Actionable variants (OncoKB therapeutic alterations level 1 - 4) were identified in 17/21 (81%) CSF cfDNA samples, and were private to CSF in 4/21 (19%) of cases (**Supplementary Table 10** - with potentially actionable variants highlighted in green in **Fig. 2** and **Supplementary Fig. 4**). Recurrent actionable variants across the cohort were *PIK3CA* activating mutations (*n* = 5), *ARID1A* deleterious mutations (*n* = 3), *MDM2* amplifications (*n* = 3) *and PTEN* deleterious mutations (*n* = 2). Variants emergent in CSF which were not found in primary tumours were: a *BRCA2* frameshift mutation p.N1599Ifs*18 accompanied by LOH in KCL610 predictive of response to PARP inhibitors; a mono-allelic *BRCA1* frameshift mutation p.Y655Vfs*18 in KCL148 with more speculative prediction of response to PARP inhibition; an *NF1* frameshift mutation p.Y1586Ifs*17 in a late BCLM-subclone of KCL566 potentially targetable by MEK inhibition; and an *ESR1* p.Y537 mutation private to KCL680 BCLM subclones and predictive of a loss of response to aromatase inhibitor but retaining sensitivity to the oestrogen receptor (ER) degrader, fulvestrant. *ERBB2* amplification/gain in CSF cfDNA were found only in the 3 known HER2+ cases. In 2 of these *ERBB2* mutations were also found; p.L755S in RMH008 (actionable) and p.L800F in RMH011 (a kinase domain mutation of unknown actionability). A further *ERBB2* variant (p.V597M, unknown actionability) was detected in KCL148, without

ERBB2 amplification. All *ERBB2* mutations were early clonal events (**Fig. 2, Supplementary Fig. 4**), as were all actionable *PIK3CA* mutations.

Of note, two CSF samples, KCL566 and KCL148, exhibited high TMB (>10 mutations/Mb) (**Fig. 3; Supplementary Fig. 1d**). Both these cases had truncal aberration in the mismatch repair gene *MLH1* and high primary tumour TMB. However, together with KCL658, both showed substantial further elevation of TMB in CSF cfDNA, indicating continuing accumulation of mutational events during evolution in the leptomeningeal space. High TMB presents an actionability target through prediction of response to immune checkpoint inhibition[40].

BCLM patient-derived organoids (PDOs)

A major challenge in developing preclinical models of BCLM is the lack of access to clinical biopsies. To overcome this challenge, we optimised methodologies for *in vitro* expansion of the small number of viable DTCs present in CSF samples. This included miniaturising the initial culture conditions and inclusion of meningeal cell conditioned medium during the early PDO culture (see Methods for full details), resulting in the successful development of 5 BCLM PDOs grown and passaged as Matrigel-embedded cultures (**Fig. 5**). Of note, although it has been questioned whether DTCs within the CSF represent the entire leptomeningeal metastatic population, Remsik and colleagues[41] have recently reported a high degree of plasticity between 'floating' and adherent tumour cells in the CSF with the floating cells displaying a more aggressive phenotype. The BCLM PDOs were subjected to WES and a comparison of CNAs with their matched CSF cfDNA samples revealed a high level of concordance (**Fig. 6a** upper panels). Similarly, comparison of the mutational landscape revealed a median of 81.5% of CSF cfDNA detected variants were present in their matched BCLM PDOs (**Fig. 6a** lower panels). Comparison of BCLM PDOs and primary tumours revealed a lower concordance in both CNA and mutations (**Fig. 6b; Supplementary Fig. 8**). Overall these data indicate that CSF cfDNA and DTCs represent a valid surrogate for the rarely available BCLM material.

Immunohistochemical staining of PDOs and matched primary tumours (**Fig. 5**) revealed a reduction of E-cadherin levels in PDO KCL566 (IC-NST with acquired *CDH1* truncating mutation in CSF cfDNA and BCLM PDO), commensurate with the reduced *CDH1* expression compared to its primary tumour (RNAseq; **Fig. 6c**), the lack of PDO *CDH1* expression detected by RTqPCR (**Fig. 6d**) and the lack of E-cadherin staining when grown as patient-derived xenografts (PDXs; see later). Similarly, there was a loss of membrane associated α-catenin staining and reduced *CTNNA1* expression in PDO KCL622 and PDO KCL625, both IC-NST showing acquired *CTNNA1* deep deletions in CSF cfDNA and BCLM PDO. p120-catenin in a cytoplasmic, as opposed to membranous, location is a hallmark of classical ILC[42] and was observed in the primary ILC tumours and PDOs KCL320 and KCL450 (**Fig. 5**). By contrast, PDO KCL566 showed mixed cytoplasmic and membranous p120-catenin and PDOs KCL622 and KCL625 showed membranous p120-catenin. These findings agree with published reports, that *CTNNA1* loss in the presence of *CDH1* expression, does not lead to p120-catenin cytoplasmic localisation despite *CTNNA1* deficient cells acquiring 'lobular-like' features of a rounded cell morphology and anoikis resistance[43].

We noted that the 3 PDOs derived from patients with ER+ primary tumours (KCL450, KCL566, KCL320) displayed loss of ER expression (**Fig. 5; Supplementary Table 11**). Although loss of ER expression is found in ~25% of PDO derived from ER+ tumour samples, likely secondary to establishing PDO culture, reduced ER expression in metastases from ER+ primary breast cancers is also reported[44-46], and linked to more frequent *TP53* mutation in endocrine therapy-resistant metastatic diease[47]. Our BCLM samples showed higher rates of *TP53* mutation and copy number loss in CSF cfDNA (12/21) than the primary tumour (6/18) (**Supplementary Fig. 5**).

Finally, immunohistochemical staining revealed enhanced HER2/ERBB2 levels in 4/5 PDO (ranging from 1+ to 3+ on IHC scoring), where in all cases the matched primary tumours were HER2 negative (0 on IHC scoring) (**Fig. 5; Supplementary Table 11).**

The standard-of-care treatment for patients with BCLM is delivery of methotrexate into CSF, intrathecally via repeated lumbar puncture or an intraventricular device such as an Ommaya reservoir, however randomised trial evidence for its effectiveness is lacking[48]. Compared to a breast cancer cell line (DU4475, grown as 3D spheroids in PDO culture conditions) which shows complete response to methotrexate, PDO KCL622 shows no response and the remaining 4 BCLM PDOs show only a partial response despite the high drug concentrations used (**Fig. 6e**).

PDX models reveal distinct BCLM organotropism

Although informative, to date mouse models of BCLM have been limited to selecting non-BCLM-derived breast cancer cell lines for their ability to grow in the leptomeninges[49]. To study the organotropism of CSF DTC-derived PDOs, PDOs were transduced to express mCherry and luciferase 2 (mChLuc2) (**Supplementary Fig. 9**) prior to inoculation into immunocompromised mice. Orthotopic inoculation into the mammary fat pad revealed a variable ability of BCLM PDOs to form primary breast tumour xenografts (**Fig. 7a**). For PDOs KCL320 and KCL450 only small tumour masses were detectable in a subset of inoculated animals, with no evidence of spontaneous metastasis to secondary sites. Of note, these PDOs were derived from the two patients with luminal, *de novo* ILCs (KCL320 and KCL450). By contrast, PDOs KCL566, KCL622 and KCL625 readily formed primary tumours with KCL566 and KCL622 showing widespread spontaneous metastasis including, in a subset of mice, to the leptomeninges. PDO KCL320 is of interest as it is derived from CSF DTCs of a patient with metastatic disease restricted to the leptomeninges, and *in vivo* showed preference for growth in the CNS environment. In addition to its failure to grow robustly after orthotopic inoculation (**Fig. 7a**), PDO KCL320 did not grow following intraductal inoculation into the mouse mammary gland (MIND model) (**Fig. 7b**) or by intraperitoneal inoculation (**Fig. 7c**). However, following intracardiac inoculation, 4 out of 5 BCLM PDOs, including PDO KCL320, resulted in colonisation of the leptomeninges and/or the brain parenchyma (**Fig. 8a**), and

intracerebroventricular (ICV) inoculation of KCL320 PDO directly into the CSF resulted in two out of the four mice developing metastatic disease restricted to the leptomeninges (**Fig. 8b**). In addition to the leptomeninges, BCLM PDOs display a strong predilection for metastasising to endocrine organs such as the ovaries, adrenal glands and pituitary gland, a feature associated with the metastatic spread of ILC[50]. This pattern of metastatic spread in the mice was distinct from the metastatic involvement in the patients from whom the PDO cohort was derived, with only 1/5 patients (KCL625) recorded to have metastatic disease in endocrine tissues (**Table 1**). The *in vivo* endocrine tissue organotropism likely reflects the enrichment for alterations in E-cadherin, the associated junctional component α-catenin and/or in cytoskeletal regulators during BCLM evolution, indicating that colonisation of the leptomeninges strongly selects for lobular-like features.

Discussion

Recent studies in metastatic breast and other solid tumours, have demonstrated specific genomic alterations in metastases beyond those observed in primary tumours[15-19,51]. However, to date, metastasis to the leptomeninges has been relatively unstudied due to difficulty accessing material from this rarely biopsied site. Here we present genomic profiling of a cohort of BCLM samples (both CSF cfDNA and CSF-DTCs expanded in culture as PDOs) with concurrently sequenced plasma cfDNA and archival tumour tissues, allowing comparison of BCLM with the primary tumour and extracranial metastatic sites. Through phylogenetic modelling we determined that BCLM exhibits early clonal divergence from the primary tumour, and that metastases spatially separated by the blood-brain/blood-CSF-barrier bear distinct genomic changes, indicating that BCLM evolve biological features favourable for tropism to and/or survival within the leptomeningeal niche. It was also revealed that leptomeningeal and extracranial metastases have different evolutionary pressures, with plasma cfDNA bearing increased marks from previous chemotherapy treatment compared to CSF cfDNA, implying that the blood-brain/blood-CSF-barrier may effectively create a sanctuary to CNS disease.

Sequencing CSF cfDNA uncovered additional genomic alterations with predicted or potential actionability over primary tumour sequencing alone in 19% of cases. This finding supports the wider prospective assessment of the clinical utility of CSF liquid biopsy for precision medicine approaches in BCLM, however important caveats to interpreting the potential actionability of these identified variants are (a) the uncertainty of drug penetration through the blood-brain and blood-CSF barriers to reach therapeutic concentrations within the CSF, and (b) limited licensed indications for agents that target the biology associated with these variants in a breast cancer setting. High TMB was discovered in 2 cases with substantial elevation of TMB in the CSF cfDNA setting in 3 cases; a finding that has therapeutic relevance given that high TMB is predictive of response to immune checkpoint inhibitors[40], and recent studies showing promising clinical efficacy of pembrolizumab[52] and that, despite being administered intravenously, pembrolizumab increases the number of cytotoxic T-cells and IFN-γ signalling in CSF[53]. CSF DTCs expanded into PDO culture allowed for immunohistochemical profiling and revealed changes in receptor levels, such as increased HER2 compared to the primary tumour. We did not identify copy number level changes in *ERBB2*, in agreement with published cohorts in breast cancer brain metastasis which show no amplification of *ERBB2* but increased *ERBB2* expression in brain metastases compared to their matched primary tumours[54,55]. The increase in HER2 levels in non-amplified (HER2-low) cancers has therapeutic implications for CNS disease given the efficacy of systemically-delivered HER2-ADCs in both HER2-amplified and HER2-low brain metastases, with trials ongoing in BCLM[56,57].

A striking finding relevant to BCLM biology was the enrichment for alteration of genes involved in cell-cell adhesion, cytoskeletal regulation and cell migration. It is well established that ILCs have a distinct biology, defined by a lack of cell-cell adhesion as a result of inactivation of *CDH1*, displaying a discohesive and migratory phenotype[58,59] and a metastatic site predilection to leptomeninges and other fluid-filled serosal-lined cavities including peritoneum, pleural and pericardial spaces[60]. We found that BCLM developing in non-ILC cases, was associated with ILC-like genomic alterations such as *CDH1*, *CTNNA1* and *ARHGEF10* loss-

of-function alterations. *CTNNA1* is a key component of the E-cadherin adherens junction while *ARHGEF10* is a GTP exchange factor for RhoGTPases with *in vitro* studies demonstrating increased cell motility and invasiveness in *ARHGEF10* depleted pancreatic cancer cells[61]. We propose that *CTNNA1* and *ARHGEF10* defects may be an alternative mechanism to become 'lobular-like' and hence more suited to colonisation of the leptomeningeal serosal membrane niche. In support of BCLM cells displaying a 'lobular-like' biology, the *in vivo* models of tumour dissemination via haematogenous route, showed BCLM cells had grew readily in the ovaries, and pituitary and adrenal glands, at higher rates than in the patients these derived from. This suggests that the biological adaptation which BCLM tumour cells undergo in the leptomeningeal space, confer an advantage to growth in endocrine organs[50], reflecting their shift to a more 'lobular-like' phenotype through the adherens junction and cytoskeletal defects.

Advances in the treatment of BCLM and leptomeningeal metastasis from other solid tumours are desperately needed. The standard-of-care BCLM intrathecal agent, methotrexate, was not particularly effective in BCLM PDOs treated *in vitro*, despite being used at concentrations similar to that achieved by intra-CSF injection. Although originally investigated in the adjuvant setting, methotrexate is not a commonly used chemotherapeutic agent in breast cancer, and the only randomised trial of intrathecal methotrexate in BCLM, failed to show a significant improvement in survival using this treatment[48]. Given the pressing requirement to discover novel approaches to treat BCLM, we describe the potential of CSF sampling and subsequent PDO derivation to enable discovery and functional validation of actionable targets. The genomic findings also highlight the potential clinical value of CSF cfDNA profiling for patients affected by this condition, and support the notion that BCLM may be better treated by discovery of blood-brain/blood-CSF-barrier penetrant therapies rather than intrathecal cytotoxic chemotherapy. Importantly, the identification of cell-cell adhesion and cytoskeleton alterations provides insight to the biology of BCLM and its ability to flourish in this metastatic site, but also presents a potential 'Achilles-heel' for therapeutic targeting. The methodologies and models described here serve as a timely resource for the ongoing research into

leptomeningeal metastasis with the overarching aim to improve therapeutic options and patient survival.

Methods

Patients and clinical sample collection

CSF and blood were concurrently collected from breast cancer patients at initial evaluation of BCLM (*n* = 21) as previously described[21]. Archived primary tumour and any metastatic recurrence tissues were retrieved. Written informed consent was obtained from all individuals in accordance with the Declaration of Helsinki under the following research ethics committee approved studies (REC ID 13/LO/1248, South East London Cancer Research Network, UK, and REC ID 14/LO/0292, Royal Marsden NHS Foundation Trust, UK). Histological subtype of primary tumours and BCLM PDOs were classified by independent breast pathologist review (I.R.).

Sample handing and processing

CSF samples were collected in standard universal containers and stored transported on wet ice to the laboratory, for centrifugation within 1 hour of collection at 300*g* for 10 minutes. CSF supernatant was stored at -80°C. The resultant CSF cell pellet was used for immediate entry to organoid culture (see below) or cells were re-suspended in 100 µL of 20% DMSO/FBS plus 100 µL OcellO primary organoid media (see below), and frozen to -80°C at a cooling rate of -1°C per minute using a Nalgene Cryogenic Freezing container. Extraction of CSF and plasma cfDNA, extraction of DNA from germline (buffy layer) and archival tissues, and ulpWGS for tumour purity assessment of cfDNA, has been previously described[21]. DNA from BCLM PDOs was extracted using Qiagen Blood and Tissue Kit.

Whole exome sequencing (WES)

Germline, tumour tissue and organoid DNA were fragmented to 200 base pair (bp) length using a Covaris E Series instrument (Covaris, MA, USA), while cfDNA was not further fragmented. Sequencing libraries were made using SureSelect XT Low Input Library Preparation kit and Agilent Bravo Automated Liquid Handling Platform. SureSelect XT Low

Input Dual Index P5 Indexed Adaptors were used for cfDNA to minimise index hopping[62]. Number of pre-capture PCR cycles was 8 - 14 based on input DNA quantity and quality (assessed by Agilent TapeStation) as advised by the manufacturers' protocol and run on an AB Veriti 96-well Thermocycler (Applied Biosystems, CA, USA). The target amount for hybridisation was 1000 ng. Exome capture was performed using the SureSelect XT Low Input Target Enrichment kit, Dynabeads MyOne Streptavidin T1 magnetic beads and the SureSelect XT Low Input Human All Exome V6 panel (60 Mb). The resultant libraries were run on the NovaSeq 6000 with an S2 flowcell (Illumina, San Diego, CA, USA) acquiring 2 × 100 bp paired-end reads. WES analysis was run through a bespoke pipeline using Nextflow v20.10.0[63]. Further analysis and plotting were performed in R 3.6.0.

(1) Alignment, duplicate removal and quality control: Reads from exome sequencing FASTQ files were first trimmed with Trim Galore v0.6.6 (https://github.com/FelixKrueger/TrimGalore), a wrapper script for Cutadapt (https://doi.org/10.14806/ej.17.1.200) and FastQC (https://qubeshub.org/resources/fastqc). Reads were then aligned against the human assembly build GRCh38 using Burrows-Wheeler Aligner (BWA) mem v0.7.17, run with default settings. Duplicate reads were removed with Picard Tools v2.23.8 according to Genome Analysis Toolkit (GATK) Best Practices recommendations[64]. GATK v4.1.9.0[65] was used for base score recalibration. Sample coverage statistics were calculated with the CollectHsMetrics command in Picard. Coverage is displayed in **Supplementary Fig. 1c** and **Supplementary Table 2**. NGSCheckMate (downloaded March 15th 2018, commit number 1079908) was run to check for concordance between tumour and normal pairs for all patients, thus verifying sample identity[66].

(2) Variant calling: A consensus variant calling method was used that included four variant callers: MuTect2[67], MuSE[68], LoFreq[69], and Strelka[70].

MuTect2: Before variant calling, a panel of normals was created by running MuTect2 (embedded in GATK v4.1.9.0) in unpaired mode on all germline samples in the study and aggregating the results with the CreateSomaticPanelOfNormals command. Paired tumour/germline variant calling was performed with MuTect2, which was provided with both the panel of normals and the Genome Aggregation Database (gnomAD) v3.1 for filtering. Resulting variants were further filtered on contamination fractions and orientation biases with the FilterMutectCalls command. Only biallelic calls were retained for further processing. Variants were annotated with ANNOVAR (2016-02-01 version)[71]. PolyPhen-2[72] and SIFT[73] algorithms, embedded in ANNOVAR, were used for predicting the effects of non-synonymous variants on protein function.

MuSE: MuSE was run in exome mode on the tumour-germline pairs. The gnomAD database was provided for removal of common variants.

LoFreq: Tumour and germline BAM files were first run through LoFreq's indelqual command to add the required information for calling indels. Then the somatic command was used to perform the variant calling. The gnomAD database was once again used as a resource of common variants.

Manta and Strelka: Manta and Strelka were run, in that order, to perform indel calling and SNV calling, respectively. Both were restricted to the exome. Indel and SNV data were combined at the end of the run.

The resulting variants were merged and restricted to only those reported by two or more callers. Those variants were further filtered; (a) by limiting them to exonic or splicing regions and (b) by applying a set of quality filters as follows: a minimum of 5 supporting reads in the tumour sample, a maximum of 2 reads in the germline sample, a minimum coverage depth of 10 reads, and a minimum VAF threshold (as defined below). The two exceptions to this were

low coverage (median coverage ≤50x) samples (CSF samples RMH010, RMH011 and KCL680) that had reduced read thresholds of 3, 1, and 7, respectively, and LoFreq variants for which the number of supporting reads in the normal sample is not reported. (c) A higher threshold was applied in each purity stratum to C>T and G>A variants to remove FFPE artefacts and (d), in cases where a variant passed these filters in one sample of a patient but not in another, the non-passing variant call was exempted from the filtering due to the additional evidence. Tumour mutational burden (TMB) was calculated for each sample by dividing the total number of these variants by the total megabases covered by the exome target panel.

Tumour purity and ploidy estimates were obtained from PureCN v1.16.0[74] except in the following situations (a) cfDNA samples, which had also undergone ulpWGS, where difference between PureCN and ulpWGS purity estimate >± 0.15, (b) samples with PureCN ploidy >3.5 or <1.5, (c) tumour tissue DNA with purity estimate ≤0.35. For these samples, manual curation of purity and ploidy was performed to reach a consensus purity and ploidy estimate, by examining the ichorCNA and CNVkit copy number profiles, and the median VAF of somatic variants present in the sample.

High purity samples (purity >0.5) were subjected to the highest filtering thresholds of 0.05 minimum VAF for non-FFPE samples and 0.07 minimum VAF for C>T and G>A SNVs in FFPE samples. For low purity samples (purity ≤0.2), the minimum VAF filters were adjusted to 0.01 (non-FFPE) and 0.03 (FFPE). For samples with purities between 0.2 and 0.5, intermediate VAF thresholds were set to 0.03 (non-FFPE) and 0.05 (FFPE).

(3) Mutational signatures analysis: Consensus SNV calls from each sample were converted back into VCF format using the MutationalPatterns R package[75]. Cosine similarity and signature contribution were calculated based on the COSMIC SBS signatures v3.2[76].

(4) Cancer driver gene lists (used in **Fig. 3 and Supplementary Fig. 5**): Cancer Gene Census (CGC) list[77] was downloaded from COSMIC database (https://cancer.sanger.ac.uk/census) on 4th July 2018 (719 genes). Breast cancer driver gene list (89 genes) was created by combining breast cancer drivers identified in two published studies[78,79]. OncoKB potentially actionable variants (Therapeutic Level 1 - 4) were identified by accessing https://www.oncokb.org[80] on 17th February 2022.

(5) Copy number alteration calling: Copy number $\log_2$ ratios were obtained from CNVkit v0.9.9[81], with each tumour sample run against a normal reference built from all matched germline samples in the study. The segmented ratios from each sample were then adjusted for tumour purity and ploidy, which were determined as described above. A copy number ratio was assigned to each gene using CNVkit's genemetrics function. To call copy number states, $\log_2$ ratios were first standardised (normalised to standard deviation) within each sample and copy number states were called for each sample as follows:

Copy number alterations (CNAs) were defined using study specific $\log_2$ ratio (standardised $\log_2$ ratio (z)) thresholds as follows: amplification >2.0; gain >1.0 to ≤2; deletion <-2.5; loss <-1.5 to ≤2.5; neutral copy number state ≥-1.5 to ≤1.0.

For CNA proportion test, differences in CNA frequency between sample types were analysed with two-sided tests of equal proportions in R. Separately, GATK v4.1.9.0 haplotype caller was run on both tumour and normal samples to obtain B-allele frequencies. ASCAT v2.5.1[82] was provided with the same purity and ploidy estimates used with CNVkit and used to obtain allele-specific copy number calls for use in PyClone (see below).

(6) Genomic alterations in BCLM *vs.* metastatic breast cancer (MBC) cohort: Mutation and copy number data from a cohort of 216 MBC samples which had undergone WES biopsy in the context of the SAFIR01 and SAFIR02 prospective trials[27]

(https://www.cbioportal.org/study/summary?id=brca_igr_2015; downloaded 9th August 2022). Rate of genomic alteration (including non-silent mutation, amplification and deletion) in each cohort (BCLM *vs.* MBC) were compared by proportion Chi-square test, using R.

(7) Gene set enrichment analysis: Genes enriched for alterations (including mutation, amplification and deletion) *vs.* matched primary tumour, 257 genes in CSF (altered in ≥4/21 of CSF samples and ≤2/21 matched archival tissue samples) and 160 genes in plasma (altered in ≥3/11 plasma samples and ≤2/21 matched archival tissues) were used as input lists for statistical overrepresentation test using the GO Biological Process (complete) annotation set of PANTHER database (http://www.pantherdb.org, analysis performed 12th July 2022). Enriched gene sets were retained if $-\log_{10}$ p-value (Fisher's exact test) was ≥3.0 in either CSF or plasma list, and if the reference list (size of GO family) was ≥10.

(8) Subclonal reconstruction and clonal evolution modelling: Subclonal reconstruction was performed with PyClone v0.13.1[83], clustering all filtered non-silent variants based on their allele frequencies, with adjustment for allele-specific copy number states (from ASCAT) and sample tumour purity estimates as described above. Variant clusters with corresponding cellular prevalence identified by PyClone were used as input to ClonEvol (Version 0.99.11)[84] run in R for clonal ordering and phylogenetic tree construction. Unless otherwise stated below, analysis parameters were: number of mutations per cluster ≥10; cluster VAF of >0.05; monoclonal cancer initiation model; bootstrap, non-parametric subclonal test with a minimum probability that cluster is non-negative with a p-value set at 0.05. Adjustments to parameters for individual models were: number of mutations per cluster ≥2 (KCL650), ≥5 (KCL523), ≥19 (KCL499), ≥30 (KCL566); and cluster VAF > 0.01 (KCL499) and >0.02 (KCL622).

RNAseq analysis of primary tumour and matched BCLM PDOs

Total RNA from archival primary tumours (FFPE) was extracted by AllPrep DNA/RNA FFPE Kit following macrodissection of involved tumour areas. Cell lines and early passage (passage

number 3 - 5) PDOs (8 - 16 days post seeding), were harvested in RLT buffer and total RNA extracted using the RNeasy kit (QIAGEN) according to the manufacturer's instructions. RNA samples were quantified using Qubit 2.0 Fluorometer and RNA integrity was checked with Agilent TapeStation. RNA from PDOs underwent library preparation using NEBNext Ultra RNA Library Prep Kit for Illumina and sequenced on Illumina HiSeq acquiring paired end reads. RNA from primary tumours underwent library preparation using NEBNext Ultra II Directional RNA Library Prep Kit for Illumina and sequenced on an Illumina NovaSeq 6000 S2 platform acquiring paired end reads. To evaluate the quality of paired-end RNA-sequencing data, FastQC and FastQ Screen[85] were run on fastq files and a summarised report was generated using MultiQC (v1.9)[86]. Reads were trimmed using Trim Galore (v0.6.6). Trimmed reads were aligned to the human reference genome GRCh38, using STAR 2.7.6a[87] with *--quantMode GeneCounts* and *--twopassMode Basic* alignment settings. Annotation file used for feature quantification was downloaded from GENCODE (v22) in GTF file format.

Cell lines

DU4475, MDA-MB-134-VI, T47D and SUM44PE were Isacke laboratory stocks, MCF7 $^{CDH1+/+}$ and MCF7 $^{CDH1-/-}$ cell lines were obtained from Professor Chris Lord (Institute of Cancer Research, UK), derived as previously published[88]. Identity of all cell lines were confirmed by short tandem repeat testing (GenePrint 10 ID System; Promega) (**Supplementary Table 12**). All cell lines were negative for mycoplasma contamination on regular testing using MycoAlert detection kit (Lonza). Cells were cultured in RPMI plus 10% FBS in 2D on tissue culture plates or in 3D in non-adherent suspension plates in a 1:1 mix of Matrigel (growth factor-reduced, BD Biosciences - 354230) and RPMI plus 10% FBS. Cell lines used for RNAseq were cultured in 3D in non-adherent suspension plates in a 1:1 mix of growth factor-reduced Matrigel and OcellO primary organoid media (see below). Total RNA was extracted as above.

Conditioned medium from primary meningeal cells

Primary human meningeal cells (ScienCell) were cultured in 2D monolayers on poly-L-lysine (ScienCell) coated flasks, using meningeal cell culture media (1401 and 1452, ScienCell). Once the cells reached 70-80% confluency, the cells were cultured in OcellO primary organoid medium for 24 hours after which the conditioned medium was collected, centrifuged and passed through a 0.2 µm pore syringe filter and stored at -20°C.

Generation of BCLM patient-derived organoids (PDOs)

The freshly collected CSF sample was transferred to sterilised 2 mL V bottom tubes and centrifuged at 300*g* for 10 minutes. The cell pellet was resuspended in OcellO primary organoid medium to 10 µL. The composition of the OcellO primary organoid medium is as follows: Advanced DMEM/F12, 10 mM HEPES, 1× Glutamax, 1× B-27 without retinoic acid, 1× N-2 supplement, 20 ng/ml fibroblast growth factor 2 (FGF2), 50 ng/ml EGF (epidermal growth factor), 10 µM Y-27632, 1.25 mM N-acetylcysteine, 5 µM A83-01, 1x penicillin and streptomycin. An equivalent volume of Matrigel was added and mixed by pipetting. 10 µL of sample mixture was pipetted per well of 96-well round bottom suspension culture plate (pre-warmed to 37°C) and placed in a tissue culture incubator for 30 minutes, prior to adding 150 µL PDO culture medium. PDO culture medium was composed of 50% primary organoid medium (OcellO), 50% meningeal conditioned medium (see above) supplemented with 10 ng/mL neuregulin-1, 10 ng/mL EGF, and for ER+ breast cancers (from primary tumour histology), 17β-oestradiol (10^{-11} M). PDO culture media was changed every 7 days. Typically, organoids developed within 7 - 21 days after initial seeding and were first passaged (1:2) by dissociation using TrypLE Express Enzyme (ThermoFisher) and resuspension in OcellO primary organoid medium. From then on organoids were cultured in OcellO primary organoid medium supplemented with oestradiol, neuregulin-1 and EGF (without meningeal conditioned medium). Once cultures were expanded, there were re-passaged at 1:3 - 1:6 every 7 days. All established organoids were regularly tested for mycoplasma, and identity confirmed by short tandem repeat (STR) testing using GenePrint 10 (Promega) (**Supplementary Table 12**), with comparison to patient germline or primary tumour DNA. For lentiviral transductions,

organoids were dissociated into single cells using TrypLE, resuspended in OcellO primary organoid medium with the addition of lentiviral particles containing the PGK-H2BmCherry-IRES-Luc2 construct and polybrene (8 µg/mL). Cell/lentivirus mixtures were centrifuged at 800g at 32°C for 30 minutes, prior to culturing as suspension cells in low adherence 6-well plates for 48 hours. Cells were then collected and centrifuged to remove media containing virus and cultured in 50% Matrigel droplets. After 14 days cells were dissociated to single cells and FACS sorted using a FACSDiva (BD Biosciences) to select for mCherry-positive cells and returned to PDO culture.

In vitro PDO drug assays

Using Hamilton cold stage dissociated, BCLM PDO (3750 cells/well) or DU4475 (1500 cells/well) were plated in 384-well clear flat bottom plates in 50% Matrigel:50% media droplets. 35 µL of OcellO (PDOs) or cell line (DU4475) media were added to each well. 24 hours later, media containing methotrexate (Enzo, ALX-440-045) in 0.2% DMSO or 0.2% DMSO alone (vehicle alone) was added to the wells and replenished on day 7. Cell viability was measured by CellTiter-Glo after 14 days.

RTqPCR

Quantitative PCR reactions were performed using Taqman Gene Expression Assay probe (*B2M*: Hs00187842_m1; *CDH1*: Hs01023895_m1). The QuantStudio6-Flex sequence detection system was used to perform relative quantification, with all reactions performed in triplicate. Data analysis was performed using Applied Biosystems Design & Analysis software 2.6.0. *B2M* was used as endogenous control.

Preparation of PDOs for histological analysis

Organoids growing in Matrigel droplets were prepared for histology by removing media and adding 4% paraformaldehyde solution to each well to cover droplets and placed at room temperature for 60 minutes. Organoids were then pipetted into a 15 mL Falcon tube, washed

twice with PBS, collected by gentle centrifugation (150*g* for 5 minutes), resuspended in 50 - 100 µL of warmed HistoGel (Thermo Fisher Scientific) and placed on Parafilm laid over ice for 5 minutes until solidified. HistoGel droplets containing organoids were then paraffin-embedded for sectioning and haematoxylin and eosin (H&E) and/or immunohistochemical staining **(Supplementary Table 13)**.

Establishment of patient-derived BCLM *in vivo* models

All animal work was carried out under UK Home Office Project Licenses 70/7413, P6AB1448A and PP4856884 granted under the Animals (Scientific Procedures) Act 1986 (Establishment Licence, X702B0E74 70/2902) and was approved by the "Animal Welfare and Ethical Review Body" at The Institute of Cancer Research (ICR). All mice were housed in individually ventilated cages and kept at 21°C ± 2°C, humidity level between 45-65% and light/dark cycle of 12 hours. Mice were monitored daily by ICR Biological Services Unit staff and had food and water *ad libitum*. Mice were weighed at least two times per week. All experiment were performed with female 6-8 week NSG mice (Charles River). For PDOs KCL320, KCL450 and KCL566 sustained-release 17β-oestradiol pellets (0.36 mg/pellet, 90 day release, Innovative Research of America) were placed subcutaneously 7 days before PDO inoculation.

Mammary fat pad inoculation*:* Mice were anaesthetized using isoflurane, and placed supine on a heat pad. Forceps were used to gently lift the 4th mammary fat pad nipple, and undissociated PDOs in 100 µL total (50% PBS: 50% Matrigel) were injected into the fat pad underlying nipple. Tumour growth was measured twice weekly up to a maximum diameter of 18 mm. Tumour volume was calculated using the following formula: Volume = 0.5236 x $[(\text{width} + \text{length})/2]^3$. Maximal tumour size/burden was not exceeded.

Intraductal inoculation

Intraductal inoculation was performed as previously described[89]. In brief, dissociated PDO cells were inoculated bilaterally into the 3rd, 4th and 5th mammary glands (6 glands/mouse).

Tumour growth was measured twice weekly up to a maximum diameter of 15 mm if a single tumour was present and 12 mm when multiple tumours formed. Maximal tumour size/burden was not exceeded.

Intraperitoneal inoculation:

2 x 10^6 dissociated PDO cells in 100 µL PBS were injected intraperitoneally.

Intracardiac inoculation: Fur was removed from chest prior to procedure. Mice were anaesthetized with isoflurane and secured onto a VisualSonics Vevo-770 ultrasound imaging platform with continuous electrocardiogram monitoring. Ultrasound gel was applied over the chest. Dissociated PDO cells were passed through a 40 µm sieve. 5 x 10^5 cells in 100 µL sterile PBS were injected into the left ventricle under continuous ultrasound guidance. To confirm successful inoculation IVIS imaging was performed within 30 minutes of injection.

Intracerebroventricular (ICV) inoculation

Mice were anaesthetised with isoflurane, prior to depilation and midline incision of skin overlaying the skull and injected with 1 × 10^4 dissociated PDO cells in 5 µL PBS into the right lateral cerebral ventricle at a rate of 2.5 µL tumour cells/min using a stereotaxic frame with pre-defined co-ordinates relative to bregma (x = 0.7 mm, y = + 0.6 mm, z = +2.0 mm). Following injection, the skin incision was closed.

Fixation, staining and imaging of tissues

Organs removed at necroscopy were fixed for 24 hours in 10% neutral buffered formalin at room temperature for 24 hours and paraffin-embedded for sectioning. For bone histology, soft tissue was stripped off prior to fixation in 10% neutral buffered formalin for 72 hours at 4°C and decalcified in Hilleman and Lee EDTA solution (5.5% EDTA in 10% neutral buffered formalin) for 7 - 10 days. Once decalcification was complete, midline sagittal sections (whole head and spine) or coronal sections (whole head) were cut using a scalpel, and tissues were

paraffin-embedded for sectioning. Metastases were quantified from sections stained for human lamin A/C. Identification of dural and pituitary metastases was confirmed by pathology review (I.R.).

Statistics

All statistical tests were performed in GraphPad Prism 9 or R statistical environment (v3.6.0). Error bars indicate ± standard deviation (SD). All comparisons between two groups were unpaired, and data was tested for normality by Shapiro-Wilk test. All p-values are two sided. Correlation tests were performed in R using the test for association between paired samples, using Pearson's product moment correlation coefficient and two-sided p-values. Unless otherwise stated the proportions testing was carried out in R using Chi squared test. All measurements shown are from distinct samples.

Data availability

Raw data from whole exome sequencing have not been deposited due to ethical considerations, however, supplementary tables listing the processed data for (a) somatic variants identified, and (b) CNA standardised Log2 ratios for all samples are provided. The RNAseq data is deposited in Sequence Read Archive (SRA) - ID number, PRJNA988939 https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA988939. All other data generated in this study are provided in the Supplementary Information and Source Data files which accompany this manuscript. Correspondence and requests for further data or materials should be addressed to the corresponding author.

Graphical depictions were created with BioRender.com

References

1 Cagney, D. N. *et al.* Incidence and prognosis of patients with brain metastases at diagnosis of systemic malignancy: a population-based study. *Neuro Oncol* **19**, 1511-1521 (2017). https://doi.org:10.1093/neuonc/nox077

2 Leone, J. P. & Lin, N. U. Systemic Therapy of Central Nervous System Metastases of Breast Cancer. *Curr Oncol Rep* **21**, 49 (2019). https://doi.org:10.1007/s11912-019-0802-6

3 Clarke, J. L., Perez, H. R., Jacks, L. M., Panageas, K. S. & Deangelis, L. M. Leptomeningeal metastases in the MRI era. *Neurology* **74**, 1449-1454 (2010). https://doi.org:10.1212/WNL.0b013e3181dc1a69

4 Le Rhun, E. *et al.* Prognostic validation and clinical implications of the EANO ESMO classification of leptomeningeal metastasis from solid tumors. *Neuro Oncol* **23**, 1100-1112 (2021). https://doi.org:10.1093/neuonc/noaa298

5 Morikawa, A. *et al.* Characteristics and Outcomes of Patients With Breast Cancer With Leptomeningeal Metastasis. *Clin Breast Cancer* **17**, 23-28 (2017). https://doi.org:10.1016/j.clbc.2016.07.002

6 Kingston, B. *et al.* Treatment and prognosis of leptomeningeal disease secondary to metastatic breast cancer: A single-centre experience. *Breast* **36**, 54-59 (2017). https://doi.org:10.1016/j.breast.2017.07.015

7 Clatot, F. *et al.* Clinical improvement and survival in breast cancer leptomeningeal metastasis correlate with the cytologic response to intrathecal chemotherapy. *J Neurooncol* **95**, 421-426 (2009). https://doi.org:10.1007/s11060-009-9940-2

8 Le Rhun, E. *et al.* A retrospective case series of 103 consecutive patients with leptomeningeal metastasis and breast cancer. *J Neurooncol* **113**, 83-92 (2013). https://doi.org:10.1007/s11060-013-1092-8

9 Niwinska, A., Rudnicka, H. & Murawska, M. Breast cancer leptomeningeal metastasis: propensity of breast cancer subtypes for leptomeninges and the analysis of factors

influencing survival. *Med Oncol* **30**, 408 (2013). https://doi.org:10.1007/s12032-012-0408-4

10 Makki, J. Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance. *Clin Med Insights Pathol* **8**, 23-31 (2015). https://doi.org:10.4137/CPath.S31563

11 Chen, Z. *et al.* Invasive lobular carcinoma of the breast: A special histological type compared with invasive ductal carcinoma. *PLoS One* **12**, e0182397 (2017). https://doi.org:10.1371/journal.pone.0182397

12 Ciriello, G. *et al.* Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell* **163**, 506-519 (2015). https://doi.org:10.1016/j.cell.2015.09.033

13 Desmedt, C. *et al.* Genomic Characterization of Primary Invasive Lobular Breast Cancer. *J Clin Oncol* **34**, 1872-1881 (2016). https://doi.org:10.1200/JCO.2015.64.0334

14 Tan, P. H. *et al.* The 2019 World Health Organization classification of tumours of the breast. *Histopathology* **77**, 181-185 (2020). https://doi.org:10.1111/his.14091

15 Nguyen, B. *et al.* Genomic characterization of metastatic patterns from prospective clinical sequencing of 25,000 patients. *Cell* **185**, 563-575 e511 (2022). https://doi.org:10.1016/j.cell.2022.01.003

16 Razavi, P. *et al.* The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. *Cancer Cell* **34**, 427-438 e426 (2018). https://doi.org:10.1016/j.ccell.2018.08.008

17 Robinson, D. R. *et al.* Integrative clinical genomics of metastatic cancer. *Nature* **548**, 297-303 (2017). https://doi.org:10.1038/nature23306

18 Yates, L. R. *et al.* Genomic Evolution of Breast Cancer Metastasis and Relapse. *Cancer Cell* **32**, 169-184 e167 (2017). https://doi.org:10.1016/j.ccell.2017.07.005

19 Zehir, A. *et al.* Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat Med* **23**, 703-713 (2017). https://doi.org:10.1038/nm.4333

20 Angus, L. *et al.* The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. *Nature Genetics* **51**, 1450-1458 (2019). https://doi.org:10.1038/s41588-019-0507-7 PMID - 31570896

21 Fitzpatrick, A. *et al.* Assessing CSF ctDNA to Improve Diagnostic Accuracy and Therapeutic Monitoring in Breast Cancer Leptomeningeal Metastasis. *Clin Cancer Res* **28**, 1180-1191 (2022). https://doi.org:10.1158/1078-0432.CCR-21-3017

22 Mattos-Arruda, L. D. *et al.* Cerebrospinal fluid-derived circulating tumour DNA better represents the genomic alterations of brain tumours than plasma. *Nature communications* **6**, 8839 (2015). https://doi.org:10.1038/ncomms9839 PMID - 26554728

23 Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415-421 (2013). https://doi.org:10.1038/nature12477

24 Lockman, P. R. *et al.* Heterogeneous blood-tumor barrier permeability determines drug efficacy in experimental brain metastases of breast cancer. *Clin Cancer Res* **16**, 5664-5678 (2010). https://doi.org:10.1158/1078-0432.CCR-10-1564

25 Kingston, B. *et al.* Genomic profile of advanced breast cancer in circulating tumour DNA. *Nat Commun* **12**, 2423 (2021). https://doi.org:10.1038/s41467-021-22605-2

26 Razavi, P. *et al.* The Genomic Landscape of Endocrine-Resistant Advanced Breast Cancers. *Cancer Cell* **34**, 427 - 438.e426 (2018). https://doi.org:10.1016/j.ccell.2018.08.008

27 Lefebvre, C. *et al.* Mutational Profile of Metastatic Breast Cancers: A Retrospective Analysis. *PLoS Med* **13**, e1002201 (2016). https://doi.org:10.1371/journal.pmed.1002201

28 Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346-352 (2012). https://doi.org:10.1038/nature10983

29 Goh, J. Y. *et al.* Chromosome 1q21.3 amplification is a trackable biomarker and actionable target for breast cancer recurrence. *Nat Med* **23**, 1319-1330 (2017). https://doi.org:10.1038/nm.4405

30 Shen, H. *et al.* SCRIB Promotes Proliferation and Metastasis by Targeting Hippo/YAP Signalling in Colorectal Cancer. *Front Cell Dev Biol* **9**, 656359 (2021). https://doi.org:10.3389/fcell.2021.656359

31 Nola, S. *et al.* Scrib regulates PAK activity during the cell migration process. *Hum Mol Genet* **17**, 3552-3565 (2008). https://doi.org:10.1093/hmg/ddn248

32 Anastas, J. N. *et al.* A protein complex of SCRIB, NOS1AP and VANGL1 regulates cell polarity and migration, and is associated with breast cancer progression. *Oncogene* **31**, 3696-3708 (2012). https://doi.org:10.1038/onc.2011.528

33 Ghosh, S. K., McCormick, T. S. & Weinberg, A. Human Beta Defensins and Cancer: Contradictions and Common Ground. *Front Oncol* **9**, 341 (2019). https://doi.org:10.3389/fonc.2019.00341

34 Drees, F., Pokutta, S., Yamada, S., Nelson, W. J. & Weis, W. I. Alpha-catenin is a molecular switch that binds E-cadherin-beta-catenin and regulates actin-filament assembly. *Cell* **123**, 903-915 (2005). https://doi.org:10.1016/j.cell.2005.09.021

35 Mohl, M., Winkler, S., Wieland, T. & Lutz, S. Gef10--the third member of a Rho-specific guanine nucleotide exchange factor subfamily with unusual protein architecture. *Naunyn Schmiedebergs Arch Pharmacol* **373**, 333-341 (2006). https://doi.org:10.1007/s00210-006-0083-0

36 Schipper, K. *et al.* Rebalancing of actomyosin contractility enables mammary tumor formation upon loss of E-cadherin. *Nat Commun* **10**, 3800 (2019). https://doi.org:10.1038/s41467-019-11716-6

37 Lutchman, M. *et al.* Dematin interacts with the Ras-guanine nucleotide exchange factor Ras-GRF2 and modulates mitogen-activated protein kinase pathways. *Eur J Biochem* **269**, 638-649 (2002). https://doi.org:10.1046/j.0014-2956.2001.02694.x

38 Mohseni, M. & Chishti, A. H. The headpiece domain of dematin regulates cell shape, motility, and wound healing by modulating RhoA activation. *Mol Cell Biol* **28**, 4712-4718 (2008). https://doi.org:10.1128/MCB.00237-08

39 Ye, Y. P. *et al.* Hypermethylation of DMTN promotes the metastasis of colorectal cancer cells by regulating the actin cytoskeleton through Rac1 signaling activation. *J Exp Clin Cancer Res* **37**, 299 (2018). https://doi.org:10.1186/s13046-018-0958-1

40 Samstein, R. M. *et al.* Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet* **51**, 202-206 (2019). https://doi.org:10.1038/s41588-018-0312-8

41 Remsik, J. *et al.* Leptomeningeal metastatic cells adopt two phenotypic states. *Cancer Rep (Hoboken)* **5**, e1236 (2022). https://doi.org:10.1002/cnr2.1236

42 Schackmann, R. C. *et al.* Cytosolic p120-catenin regulates growth of metastatic lobular carcinoma through Rock1-mediated anoikis resistance. *J Clin Invest* **121**, 3176-3188 (2011). https://doi.org:10.1172/JCI41695

43 de Groot, J. S. *et al.* alphaE-catenin is a candidate tumor suppressor for the development of E-cadherin-expressing lobular-type breast cancer. *J Pathol* **245**, 456-467 (2018). https://doi.org:10.1002/path.5099

44 Jung, J. *et al.* Discordances in ER, PR, and HER2 between primary breast cancer and brain metastasis. *J Neurooncol* **137**, 295-302 (2018). https://doi.org:10.1007/s11060-017-2717-0

45 Li, B. D., Byskosh, A., Molteni, A. & Duda, R. B. Estrogen and progesterone receptor concordance between primary and recurrent breast cancer. *J Surg Oncol* **57**, 71-77 (1994). https://doi.org:10.1002/jso.2930570202

46 Amir, E. *et al.* Prospective study evaluating the impact of tissue confirmation of metastatic disease in patients with breast cancer. *J Clin Oncol* **30**, 587-592 (2012). https://doi.org:10.1200/JCO.2010.33.5232

47 Li, Z. *et al.* Mutual exclusivity of ESR1 and TP53 mutations in endocrine resistant metastatic breast cancer. *NPJ Breast Cancer* **8**, 62 (2022). https://doi.org:10.1038/s41523-022-00426-w

48 Boogerd, W. *et al.* The relevance of intraventricular chemotherapy for leptomeningeal metastasis in breast cancer: a randomised study. *Eur J Cancer* **40**, 2726-2733 (2004). https://doi.org:10.1016/j.ejca.2004.08.012

49 Boire, A. *et al.* Complement Component 3 Adapts the Cerebrospinal Fluid for Leptomeningeal Metastasis. *Cell* **168**, 1101-1113 e1113 (2017). https://doi.org:10.1016/j.cell.2017.02.025

50 Bumpers, H. L., Hassett, J. M., Jr., Penetrante, R. B., Hoover, E. L. & Holyoke, E. D. Endocrine organ metastases in subjects with lobular carcinoma of the breast. *Arch Surg* **128**, 1344-1347 (1993). https://doi.org:10.1001/archsurg.1993.01420240052009

51 Angus, L. *et al.* The genomic landscape of metastatic breast cancer highlights changes in mutation and signature frequencies. *Nat Genet* **51**, 1450-1458 (2019). https://doi.org:10.1038/s41588-019-0507-7

52 Brastianos, P. K. *et al.* Single-arm, open-label phase 2 trial of pembrolizumab in patients with leptomeningeal carcinomatosis. *Nature Medicine* **26**, 1280 - 1284 (2020).

53 Prakadan, S. M. *et al.* Genomic and transcriptomic correlates of immunotherapy response within the tumor microenvironment of leptomeningeal metastases. *Nat Commun* **12**, 5955 (2021). https://doi.org:10.1038/s41467-021-25860-5

54 Thomson, A. H. *et al.* Changing molecular profile of brain metastases compared with matched breast primary cancers and impact on clinical outcomes. *Br J Cancer* **114**, 793-800 (2016). https://doi.org:10.1038/bjc.2016.34

55 Priedigkeit, N. *et al.* Intrinsic Subtype Switching and Acquired ERBB2/HER2 Amplifications and Mutations in Breast Cancer Brain Metastases. *JAMA Oncol* **3**, 666-671 (2017). https://doi.org:10.1001/jamaoncol.2016.5630

56 Bartsch, R. *et al.* Trastuzumab deruxtecan in HER2-positive breast cancer with brain metastases: a single-arm, phase 2 trial. *Nat Med* **28**, 1840-1847 (2022). https://doi.org:10.1038/s41591-022-01935-8

57 Perez-Garcia, J. M. *et al.* Trastuzumab Deruxtecan in Patients with Central Nervous System Involvement from HER2-Positive Breast Cancer: The DEBBRAH Trial. *Neuro Oncol* (2022). https://doi.org:10.1093/neuonc/noac144

58 Christgen, M. & Derksen, P. Lobular breast cancer: molecular basis, mouse and cellular models. *Breast Cancer Res* **17**, 16 (2015). https://doi.org:10.1186/s13058-015-0517-z

59 Derksen, P. W. *et al.* Mammary-specific inactivation of E-cadherin and p53 impairs functional gland development and leads to pleomorphic invasive lobular carcinoma in mice. *Dis Model Mech* **4**, 347-358 (2011). https://doi.org:10.1242/dmm.006395

60 He, H., Gonzalez, A., Robinson, E. & Yang, W. T. Distant metastatic disease manifestations in infiltrating lobular carcinoma of the breast. *AJR Am J Roentgenol* **202**, 1140-1148 (2014). https://doi.org:10.2214/AJR.13.11156

61 Joseph, J. *et al.* Rho guanine nucleotide exchange factor ARHGEF10 is a putative tumor suppressor in pancreatic ductal adenocarcinoma. *Oncogene* **39**, 308-321 (2020). https://doi.org:10.1038/s41388-019-0985-1

62 Kircher, M., Sawyer, S. & Meyer, M. Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Res* **40**, e3 (2012). https://doi.org:10.1093/nar/gkr771

63 Di Tommaso, P. *et al.* Nextflow enables reproducible computational workflows. *Nat Biotechnol* **35**, 316-319 (2017). https://doi.org:10.1038/nbt.3820

64 Van der Auwera, G. A. *et al.* From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**, 11 10 11-11 10 33 (2013). https://doi.org:10.1002/0471250953.bi1110s43

65 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303 (2010). https://doi.org:10.1101/gr.107524.110

66 Lee, S. *et al.* NGSCheckMate: software for validating sample identity in next-generation sequencing studies within and across data types. *Nucleic Acids Res* **45**, e103 (2017). https://doi.org:10.1093/nar/gkx193

67 Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* **31**, 213-219 (2013). https://doi.org:10.1038/nbt.2514

68 Fan, Y. *et al.* MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. *Genome Biol* **17**, 178 (2016). https://doi.org:10.1186/s13059-016-1029-6

69 Wilm, A. *et al.* LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res* **40**, 11189-11201 (2012). https://doi.org:10.1093/nar/gks918

70 Kim, S. *et al.* Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods* **15**, 591-594 (2018). https://doi.org:10.1038/s41592-018-0051-x

71 Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164 (2010). https://doi.org:10.1093/nar/gkq603

72 Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* **7**, 248-249 (2010). https://doi.org:10.1038/nmeth0410-248

73 Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* **4**, 1073-1081 (2009). https://doi.org:10.1038/nprot.2009.86

74 Riester, M. *et al.* PureCN: copy number calling and SNV classification using targeted short read sequencing. *Source Code Biol Med* **11**, 13 (2016). https://doi.org:10.1186/s13029-016-0060-z

75 Blokzijl, F., Janssen, R., van Boxtel, R. & Cuppen, E. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. *Genome Med* **10**, 33 (2018). https://doi.org:10.1186/s13073-018-0539-0

76 Alexandrov, L. B. *et al.* The repertoire of mutational signatures in human cancer. *Nature* **578**, 94-101 (2020). https://doi.org:10.1038/s41586-020-1943-3

77 Futreal, P. A. *et al.* A census of human cancer genes. *Nat Rev Cancer* **4**, 177-183 (2004). https://doi.org:10.1038/nrc1299

78 Pereira, B. *et al.* The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun* **7**, 11479 (2016). https://doi.org:10.1038/ncomms11479

79 Nik-Zainal, S. *et al.* Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**, 47-54 (2016). https://doi.org:10.1038/nature17676

80 Chakravarty, D. *et al.* OncoKB: A Precision Oncology Knowledge Base. *JCO Precis Oncol* **2017** (2017). https://doi.org:10.1200/PO.17.00011

81 Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput Biol* **12**, e1004873 (2016). https://doi.org:10.1371/journal.pcbi.1004873

82 Van Loo, P. *et al.* Allele-specific copy number analysis of tumors. *Proc Natl Acad Sci U S A* **107**, 16910-16915 (2010). https://doi.org:10.1073/pnas.1009843107

83 Roth, A. *et al.* PyClone: statistical inference of clonal population structure in cancer. *Nat Methods* **11**, 396-398 (2014). https://doi.org:10.1038/nmeth.2883

84 Dang, H. X. *et al.* ClonEvol: clonal ordering and visualization in cancer sequencing. *Ann Oncol* **28**, 3076-3082 (2017). https://doi.org:10.1093/annonc/mdx517

85 Wingett, S. W. & Andrews, S. FastQ Screen: A tool for multi-genome mapping and quality control. *F1000Res* **7**, 1338 (2018). https://doi.org:10.12688/f1000research.15931.2

86 Ewels, P., Magnusson, M., Lundin, S. & Kaller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047-3048 (2016). https://doi.org:10.1093/bioinformatics/btw354

87 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013). https://doi.org:10.1093/bioinformatics/bts635

88 Bajrami, I. *et al.* E-Cadherin/ROS1 Inhibitor Synthetic Lethality in Breast Cancer. *Cancer Discov* **8**, 498-515 (2018). https://doi.org:10.1158/2159-8290.CD-17-0603

89 Sflomos, G. *et al.* A Preclinical Model for ERalpha-Positive Breast Cancer Points to the Epithelial Microenvironment as Determinant of Luminal Phenotype and Hormone Response. *Cancer Cell* **29**, 407-422 (2016). https://doi.org:10.1016/j.ccell.2016.02.002

Acknowledgements

We would like to thank Naomi Guppy and her team in the Breast Cancer Now Toby Robins Research Centre Nina Barough Pathology Core Faculty for histopathology support and the following facilities at the ICR; Tumour Profiling Unit, Biological Services Unit, FACS and Light Microscopy Facility. We thank Eleanor Knight, Daniela Nova and Rebecca Marlow in the Breast Cancer Now Toby Robins Research Centre Patient Derived Models facility for their help in establishing the PDO culture methodologies and OcellO for providing the PDO primary organoid medium. We thank John Alexander, Erle Holgersen and Hui Xiao in Syed Haider's team for their contributions to data analysis. We thank Phil Bland (ICR) for mouse intraductal injections. For patient enrolment and sample collection at Kings College London and Guys and St Thomas's NHS Foundation Trust, we thank clinicians S. Irshad, H. Kristeleit, J. Mansi, M. Nathan, E. Sawyer, A. Swampillai, M. Harries and the study nurses and staff, and R. Liccardo, R. Marlow, and V. Shah for patient enrolment and sample preparation collection. For patient enrolment and sample collection in Royal Marsden NHS Foundation Trust, we thank clinicians S. Johnston, M. Parton, M. Robert, L. Ulrich and A. Okines and the study nurses and staff, A. Bambra, D. Kelly, S. Thompson.

Author Contributions

Conceptualization AF, ANJT, CMI

Supervision SH, ANJT, CMI

Funding acquisition AF, ANJT, CMI

Investigation AF, MI, AM, DV, TA, IR, NCT, SH, ANJT, CMI

Writing - original draft AF, MI, ANJT, CMI

Writing - review and editing AF, MI, AM, DV, TA, IR, NCT, SH, ANJT, CMI

Competing Interests statement

The authors have no conflicts of interest to declare.

Funding

This study was supported by: Medical Research Council Clinical Research Training Fellowship - MR/P001564/1 (AF, CI), NIHR Biomedical Research Centre (BRC) at the Royal Marsden and the ICR post-doctoral support funding - W94500 (AF), Breast Cancer Now Programme Funding to the Breast Cancer Now Toby Robins Research Centre - CTR-Q4-Y1-5, CTR-Q5-Y1 (CI, NT, ANJT) and to the Breast Cancer Now Research Unit at Kings College London - KCL-Q2-Y1-5, KCL-Q3-Y1 (ANJT), Cancer Research UK Centre Grant funding to the Kings College London - CRUK RE15196-9 (ANJT).

Table 1 | Clinicopathological characteristics of BCLM cohort

Study ID	Histological subtype	Time from pBC (years)	Time from mBC (years)	Parenchymal brain metastasis	Extracranial metastases	Extracranial disease status	Samples sequenced	BCLM survival (m)
KCL148	Lobular	4.7	4.7	None	Liver	Stable	CSF cfDNA, primary tumour	5.71
KCL320	Lobular	4.4	0.0	None	None	Absent	CSF cfDNA, primary tumour, PDO	5.50
KCL448	NST	4.3	1.0	Synchronous	Peritoneum, pericardium, distant nodes, soft tissue, bone	Progressive	CSF cfDNA, plasma cfDNA, primary tumour, skin metastasis	3.18
KCL449	Lobular	9.0	0.6	None	Bone	Stable	CSF cfDNA, primary tumour	27.43
KCL450	Lobular	14.0	1.0	None	Chest wall, peritoneal, pleura, bone	Progressive	CSF cfDNA, plasma cfDNA, primary tumour, PDO	1.39
KCL499	NST	0.7	0.0	Synchronous	None	Absent	CSF cfDNA, plasma cfDNA, primary tumour, second primary tumour	3.39
KCL523	Mixed	6.7	1.2	None	Peritoneum, bone, orbit	Stable	CSF cfDNA, plasma cfDNA, primary tumour	6.04
KCL553	Lobular	1.7	1.7	None	Peritoneum, pleura, ovaries, bone, orbit	Stable	CSF cfDNA, plasma cfDNA, primary tumour, ovary metastasis	18.39
KCL566	NST	6.7	1.1	Metachronous (stable)	Distant nodes, peritoneum, subcutaneous, bone	Progressive	CSF cfDNA, plasma cfDNA, primary tumour, nodal metastasis, PDO	0.86
KCL590	Micropapillary	2.6	2.6	None	Liver, spleen, bone	Progressive	CSF cfDNA, plasma cfDNA, primary tumour	1.29
KCL610	Lobular	6.5	6.5	None	Bone	Unknown	CSF cfDNA, plasma cfDNA, primary tumour	1.32

KCL616	Lobular	14.7	2.9	None	Bone, distant nodes	Unknown	CSF cfDNA, primary tumour	0.14
KCL617	Apocrine	1.0	0.0	None	Bone	Stable	CSF cfDNA, primary tumour	19.25
KCL622	Lobular	1.3	1.3	None	Bone, adrenal	Stable	CSF cfDNA, plasma cfDNA, primary tumour, PDO	3.64
KCL625	NST	11.1	0.0	None	None	Absent	CSF cfDNA, primary tumour, PDO	9.50
KCL650	NST	2.0	0.1	None	Liver, bone	Progressive	CSF cfDNA, plasma cfDNA	2.54
KCL658	NST	1.5	0.3	Metachronous (stable)	Skin	Progressive	CSF cfDNA, plasma cfDNA	9.75
KCL680	Lobular	0.7	0.7	None	Bone, peritoneum, ovaries	Stable	CSF cfDNA, primary tumour	25.61
RMH008	Lobular	4.3	2.0	Metachronous (stable)	Ovaries, cervix, vagina	Stable	CSF cfDNA, plasma cfDNA, ovary metastasis	0.93
RMH010	Lobular	10.6	4.0	Metachronous (stable)	Peritoneum, liver, bone, orbit	Progressive	CSF cfDNA, plasma cfDNA, primary tumour, liver metastasis	3.39
RMH011	NST	12.3	5.6	Metachronous (stable)	Lung, pleura, bone	Stable	CSF cfDNA, brain metastasis	18.64

BCLM, breast cancer leptomeningeal metastasis; cfDNA, cell-free DNA; CSF, cerebrospinal fluid; mBC, metastatic breast cancer; m, months; NST, no special type; PDO, patient-derived organoids; pBC, primary breast cancer. Metachronous, prior to BCLM diagnosis; Synchronous, diagnosed concurrenlty with BCLM.

Fig. 1. Molecular profiling of BCLM by whole exome sequencing. **a** Schematic showing the sites of material collected. **b** Flowchart of sample processing pipeline. cfDNA, cell-free DNA; DTC, ctDNA, circulating tumour DNA, DTC, disseminated tumour cell; PDO, patient-derived organoid. **c-e** Samples underwent WES to identify mutation (exonic variants excluding synonymous SNVs, plus splice variants) and copy number aberration (CNA) events. **c** Chemotherapy-related COSMIC Single Base Substitution (SBS) mutational signatures compared between sites. Signature contributions of the six chemotherapy-related SBS signatures (SBS11, SBS17b, SBS28, SBS31, SBS35, SBS86) were compared between CSF cfDNA, plasma cfDNA, primary and metastasis tumour samples. Stacked bar plots display the mean signature contribution for each SBS signature across all samples in each category. Significantly different total chemotherapy-related SBS signature contribution between CSF and plasma cfDNA is shown (Mann-Whitney two-tailed test). All other comparisons were ns. **d** WES of paired CSF cfDNA samples collected at BCLM diagnosis and primary tumour samples (*n* = 17 pairs). Bars show proportion of total variants unique to CSF (median 47.8%; IQR 43.0 - 56.4%), unique to primary tumour (median 20.4%; IQR 16.6 - 31.5%) and shared (median 23.0%; IQR 13.8 - 33.7%). **e** WES of paired CSF cfDNA and plasma cfDNA samples collected at BCLM diagnosis (*n* = 11 pairs). Bars show proportion of total variants unique to CSF (median 24.6%; IQR 16.0 - 38.6%), unique to plasma (median 16.8%; IQR 11.3 - 26.6%) and shared (median 43.4%; IQR 24.2 - 65.1%).

Fig. 2 BCLM clonal evolution modelling. ClonEvol was used to infer consensus clonal evolution models, using the variant clusters generated by PyClone. **a** KCL566, **b** RMH010, **c** KCL553. Remaining models are presented in Supplementary Fig. 4. Box plots show the variant clusters identified in each sample, and displays the variant allele frequency (VAF) of each mutation, with the median VAF depicted by black lines. Bell plots show the ClonEvol estimated cancer cell fraction (CCF) and evolution order of clones within each sample. Phylogenetic trees for each case depicts the evolution pattern. Each branch represents the evolution path of a clone from its parental clone, and each node marks the clone number of

the preceding branch. Clone colour and number match those in box and bell plots. Samples are annotated at clone nodes when no further subclone evolution occurs. Branch lengths are scaled by the square root of number of mutations per clone, and branch width is inversely proportional to branch length. Mutations per clone are shown to top left of the trees. Gene names and variant type are annotated on branches as follows; black font if appear in the list of BCLM frequently altered cancer driver genes (shown in Fig. 3); green font if variant is OncoKB annotated as potentially therapeutically actionable (Supplementary Table 10). Variant type abbreviations; fs, frameshift; nfs, non-frameshift; ns, nonsynonymous; sg, stop-gain; spl, splice-site variant.

Fig. 3 Cancer driver gene aberration in BCLM CSF cfDNA samples. Uppermost bar plot show tumour mutational burden (TMB) by sample. Stacked bar plot below shows mutational subtype for each sample as fraction of all variants (excluding synonymous). Banners indicate primary tumour histological subtype and ER and HER2 status. Main plot shows the 35 most frequently altered cancer driver genes (alterations include non-silent somatic mutation or high-level copy number change) in CSF cfDNA as follows; Cancer Gene Census (CGC) gene altered in at least 4 cases; breast cancer driver gene (defined in Methods) altered in at least 3 cases. Right side bar plot shows percentage of CSF cfDNA samples (n = 21) with gene alterations that are unique to the CSF sample (green) or shared with any other site per individual (grey). Left side bar plot shows genomic alteration rate in the MBC cohort. See Supplementary Fig. 5 for extended data on matched plasma cfDNA, primary tumour and metastasis samples, and Supplementary Table 7 for frequently altered cancer driver genes in plasma cfDNA. Gene names are highlighted in red if genomic alteration rate was significantly different ($p \leq 0.05$ by Chi-square test two-sided) in BCLM cohort compared to a publicly available dataset of 216 metastatic breast cancer (MBC cohort) samples (see Methods). Extended data in Supplementary Table 8 displays the significance values for this comparison.

Fig. 4 BCLM CNA landscape and gene set enrichment**. a** Genome wide purity-adjusted CNA plots of all samples (CSF cfDNA (*n* = 21), plasma cfDNA (*n* = 11), primary tumour (*n* = 18) and metastasis (*n* =8) showing frequency of copy number gain (red bars, positive y-axis) and copy number loss (blue bars, negative y-axis) at a gene level. **b** Volcano plots displaying comparison of CNA frequency per gene, between CSF cfDNA and primary tumours (upper panels) and between CSF cfDNA and plasma cfDNA (lower panels). Delta proportion (y-axis) is the proportional difference in CNA frequency between sample types. Significance (x-axis, -$\log_{10}$ p-value) obtained by Chi-squared proportion test (two-sided) compares CNA event proportion per gene between sample types (source data are provided as a Source Data file). Gene names are annotated if delta proportion is ≥±0.20 and -$\log_{10}$ p-value ≥1.5 (as shown in Supplementary Fig. 6). **c** Gene set enrichment analysis. Genes enriched for alterations (including non-silent somatic mutation or high-level copy number change) in CSF cfDNA (*n* = 257) and/or plasma cfDNA (*n* = 160) *vs.* matched primary tumour, underwent statistical overrepresentation test using GO biological process (PANTHER). Displayed terms were retained if -$\log_{10}$ p-value was ≥3.0 in either sample type (Fisher's exact test, adjusted for multiple testing using the Benjamini-Hochberg procedure with an overall false discovery rate of <0.05). **d** Focused alteration plot of genes from the indicated GO biological processes overrepresented in CSF cfDNA *vs*. matched primary tumour, but not in plasma cfDNA *vs*. primary tumour. Banners above indicate; histological subtype (primary tumour) and *CDH1* mutation status of CSF cfDNA. Left panel shows gene names and chromosomal location, grouped by GO biological processes. Bars to right show the frequency of gene alteration across CSF cfDNA samples (*n* = 21) with 'cfDNA only' indicating alterations unique to CSF cfDNA or CSF and plasma cfDNA at the time of BCLM diagnosis, and 'shared' occurring both in CSF cfDNA and primary tumour or metastasis sample. See Supplementary Fig. 7 for extended data on matched plasma cfDNA, primary tumour and metastasis samples.

Fig. 5. Generation of BCLM patient-derived organoids (PDOs) established from CSF DTCs. Histology and immunohistochemistry of BCLM PDOs and matched primary tumours. A

minimum of 2 biological replicates were performed for each PDO immunohistochemical staining (range 2 - 4 replicates) with representative images shown. Scale bars, 50 μm.

Fig. 6. Characterisation of BCLM PDOs. **a** Genomic comparison of CSF cfDNA and matched BCLM PDO (early passage) by WES. Top panel, correlation plots of copy number Z-score (one data point per gene) with regression line plotted (purple) and Pearson's r value plus two-sided significance values are shown. Lower panel, Venn diagrams of somatic mutations (excluding synonymous) present in CSF cfDNA and matched BCLM PDO, values shown are mutation count. **b** Comparison of Pearson r values from Fig. 6a and Supplementary Fig. 8 by Mann-Whitney, two-tailed test showing a higher level of correlation in CNA status between PDO (n = 5) and CSF cfDNA (n = 5) than between PDO and primary tumour DNA (n = 5) (left panel) and higher fraction of shared variants between CSF cfDNA and PDO than primary tumour and PDO (right panel), source data are provided as a Source Data file. **c** RNAseq of PDOs (fresh early passage) and matched primary tumours (FFPE). Expression levels of adherens junction components are shown as Z-scores of log2-TPM values (transcripts per million). **d** RTqPCR analysis of *CDH1* expression in PDOs as compared with cell line controls. n = 1 biological replicate, n = 3 technical replicates, source data are provided as a Source Data file. **e** BCLM PDOs and cell line control (DU4475) were cultured in the presence of methotrexate in 0.2% DMSO or 0.2% DMSO alone (vehicle alone). Cell viability was measured by CellTiter-Glo after 14 days. Data represents mean of n = 4 wells per datapoint ±SD, normalised to vehicle alone treatment, source data are provided as a Source Data file.

Fig. 7. Variable engraftment rate of BCLM PDOs in the orthotopic site. Female NSG mice were inoculated with mChLuc2-labelled PDOs as follows. **a** Non-dissociated BCLM PDOs (~1 x 10^6 cells) were injected transcutaneously into the 4th mammary fat pad (unilateral), n = 4 - 6 mice per PDO. Upper panel, tumour size was measured bi-weekly. Middle panel, mammary fat pad tumours stained for H&E, lamin A/C, ER and E-cadherin. Scale bar, 100 μm. Lower panel, bar charts show percentage of mice with metastatic dissemination in individual organs.

Pale bars indicate % tissues with single or small cell clusters (1 - 10 cells), dark bars indicate tissues with metastatic deposits (>10 cells), source data are provided as a Source Data file. **b** 4 x 10^4 dissociated KCL320 PDO cells per gland were inoculated via the mammary gland intraductal (MIND) route into the 3rd, 4th and 5th mammary glands, bilaterally (6 glands/mouse; *n* = 6 mice). Mice were IVIS imaged on day 0 and day 165 (endpoint). **c** 2 x 10^6 dissociated KCL320 PDO cells injected intraperitoneally (*n* = 4 mice) with IVIS imaging at day 0 and day 81 (endpoint).

Fig. 8. Modelling BCLM by intracardiac or intracerebroventricular injection of PDOs. BCLM PDOs were dissociated into single cells. **a** 0.5 - 1.0 x 10^6 cells inoculated into the left cardiac ventricle under ultrasound guidance. Upper panel, IVIS imaging immediately after cell injection (day 0) and at endpoint for each mouse (development of significant morbidity), day 40 - 210. Middle panel, metastatic site involvement assessed by lamin A/C staining of FFPE sections. Bar charts show percentage of mice with disseminated tumour cells in individual organs, with pale bars indicating with single or small clusters of cells (1 - 10 cells) and dark bars indicating tissues with metastatic deposits (>10 cells), source data are provided as a Source Data file. Bottom panel. IHC staining for human lamin A/C of whole cranium and spine FFPE sections showing tumour cells growing along the meninges (orange arrowheads) or in the brain parenchyma (blue arrowheads), in the dura (red arrowheads) and pituitary (green arrowhead). Scale bars; Left, 100 μm; other panels, 1 mm. **b** 1 x 10^5 cells inoculated intracerebroventricularly (ICV) (*n* = 4 mice). Left panel, IVIS imaging on day 0 and day 81 (endpoint) showing development of metastasis in two mice. Right panels, cranium and spine FFPE sections stained for human lamin A/C. Orange arrowheads indicate leptomeningeal involvement. Scale bars; 250 μm.

Figure 1

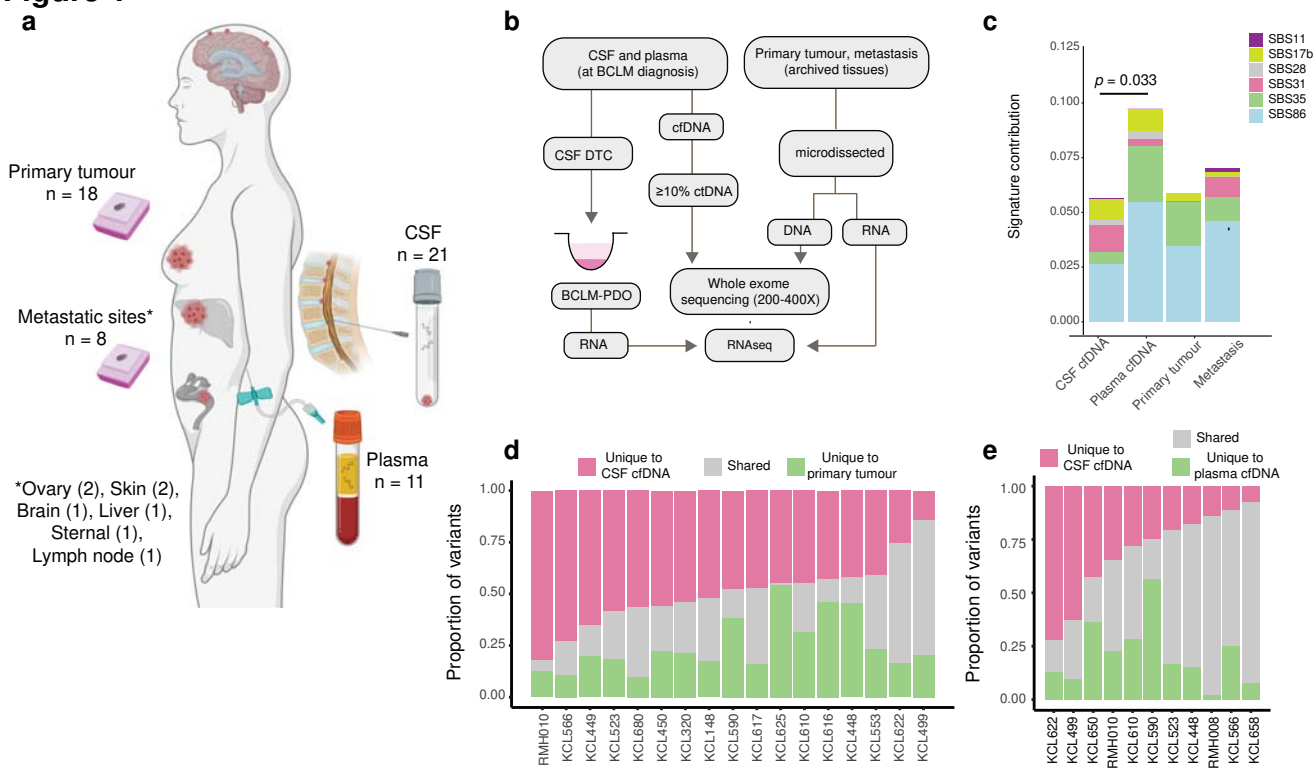

a
Primary tumour
n = 18
Metastatic sites*
n = 8
*Ovary (2), Skin (2),
Brain (1), Liver (1),
Sternal (1),
Lymph node (1)
CSF
n = 21
Plasma
n = 11
b
CSF and plasma
(at BCLM diagnosis)
Primary tumour, metastasis
(archived tissues)
cfDNA
CSF DTC
microdissected
≥10% ctDNA
DNA
RNA
BCLM-PDO
Whole exome
sequencing (200-400X)
RNA
RNAseq
c
SBS11
SBS17b
SBS28
SBS31
SBS35
SBS86
p = 0.033
Signature contribution
0.125
0.100
0.075
0.050
0.025
0.000
CSF cfDNA
Plasma cfDNA
Primary tumour
Metastasis
d
Unique to
CSF cfDNA
Shared
Unique to
primary tumour
Proportion of variants
1.00
0.75
0.50
0.25
0.00
RMH010
KCL566
KCL449
KCL523
KCL680
KCL450
KCL320
KCL148
KCL590
KCL617
KCL625
KCL610
KCL616
KCL448
KCL553
KCL622
KCL499
e
Unique to
CSF cfDNA
Shared
Unique to
plasma cfDNA
Proportion of variants
1.00
0.75
0.50
0.25
0.00
KCL622
KCL499
KCL650
RMH010
KCL610
KCL590
KCL523
KCL448
RMH008
KCL566
KCL658

Figure 2

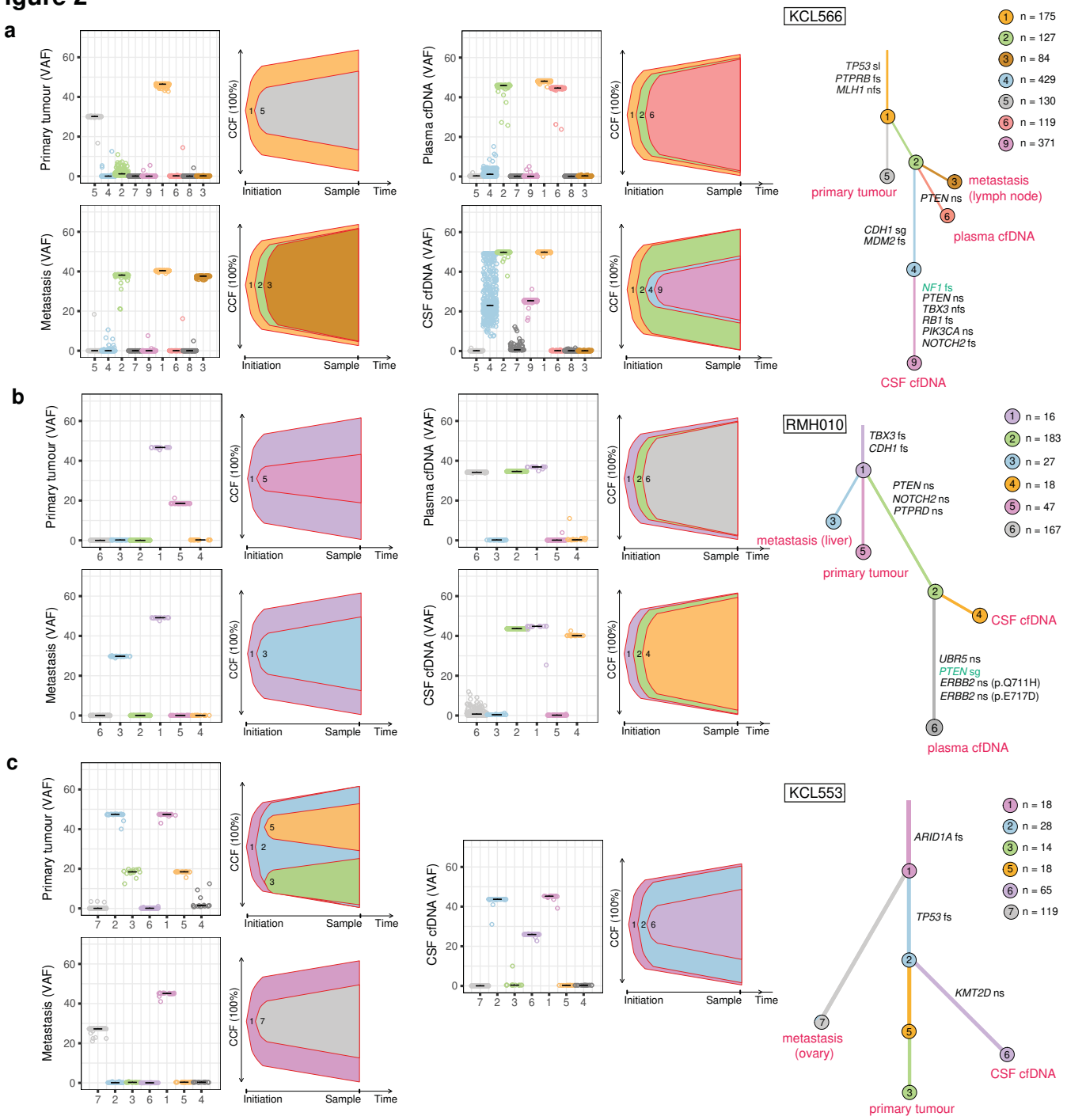

a
KCL566
Primary tumour (VAF)
Metastasis (VAF)
Plasma cfDNA (VAF)
CSF cfDNA (VAF)
CCF (100%)
Initiation
Sample
Time
n = 175
n = 127
n = 84
n = 429
n = 130
n = 119
n = 371
TP53 sl
PTPRB fs
MLH1 nfs
primary tumour
metastasis (lymph node)
PTEN ns
plasma cfDNA
CDH1 sg
MDM2 fs
NF1 fs
PTEN ns
TBX3 nfs
RB1 fs
PIK3CA ns
NOTCH2 fs
CSF cfDNA
b
RMH010
n = 16
n = 183
n = 27
n = 18
n = 47
n = 167
TBX3 fs
CDH1 fs
PTEN ns
NOTCH2 ns
PTPRD ns
metastasis (liver)
primary tumour
CSF cfDNA
UBR5 ns
PTEN sg
ERBB2 ns (p.Q711H)
ERBB2 ns (p.E717D)
plasma cfDNA
c
KCL553
n = 18
n = 28
n = 14
n = 18
n = 65
n = 119
ARID1A fs
TP53 fs
KMT2D ns
metastasis (ovary)
CSF cfDNA
primary tumour

Figure 3

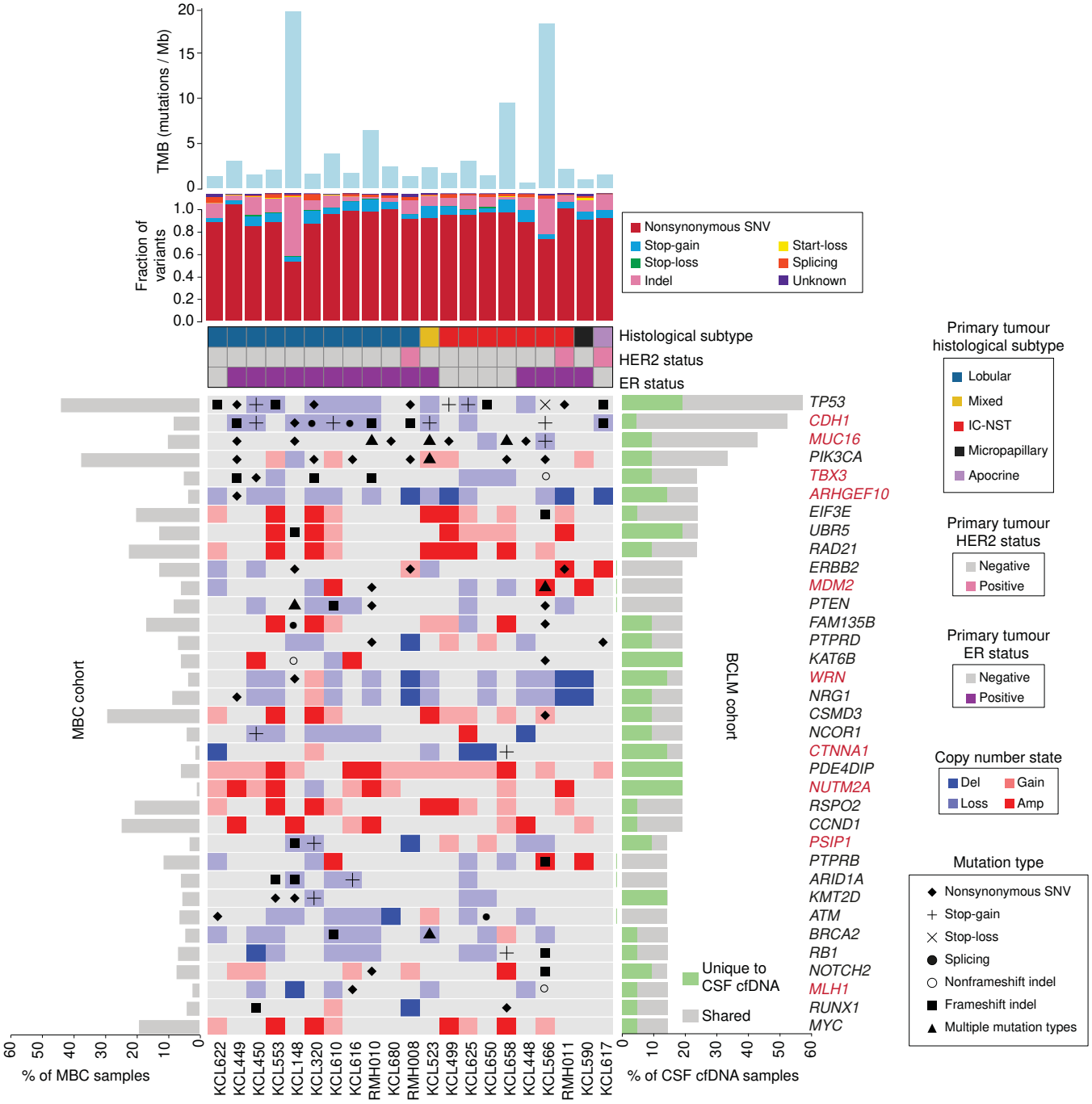

TMB (mutations / Mb)
Fraction of variants
Nonsynonymous SNV
Stop-gain
Stop-loss
Indel
Start-loss
Splicing
Unknown
Histological subtype
HER2 status
ER status
MBC cohort
BCLM cohort
% of MBC samples
% of CSF cfDNA samples
Unique to CSF cfDNA
Shared
TP53
CDH1
MUC16
PIK3CA
TBX3
ARHGEF10
EIF3E
UBR5
RAD21
ERBB2
MDM2
PTEN
FAM135B
PTPRD
KAT6B
WRN
NRG1
CSMD3
NCOR1
CTNNA1
PDE4DIP
NUTM2A
RSPO2
CCND1
PSIP1
PTPRB
ARID1A
KMT2D
ATM
BRCA2
RB1
NOTCH2
MLH1
RUNX1
MYC
KCL622
KCL449
KCL450
KCL553
KCL148
KCL320
KCL610
KCL616
RMH010
KCL680
RMH008
KCL523
KCL499
KCL625
KCL650
KCL658
KCL448
KCL566
RMH011
KCL590
KCL617
Primary tumour histological subtype
Lobular
Mixed
IC-NST
Micropapillary
Apocrine
Primary tumour HER2 status
Negative
Positive
Primary tumour ER status
Negative
Positive
Copy number state
Del
Gain
Loss
Amp
Mutation type
Nonsynonymous SNV
Stop-gain
Stop-loss
Splicing
Nonframeshift indel
Frameshift indel
Multiple mutation types

Figure 4

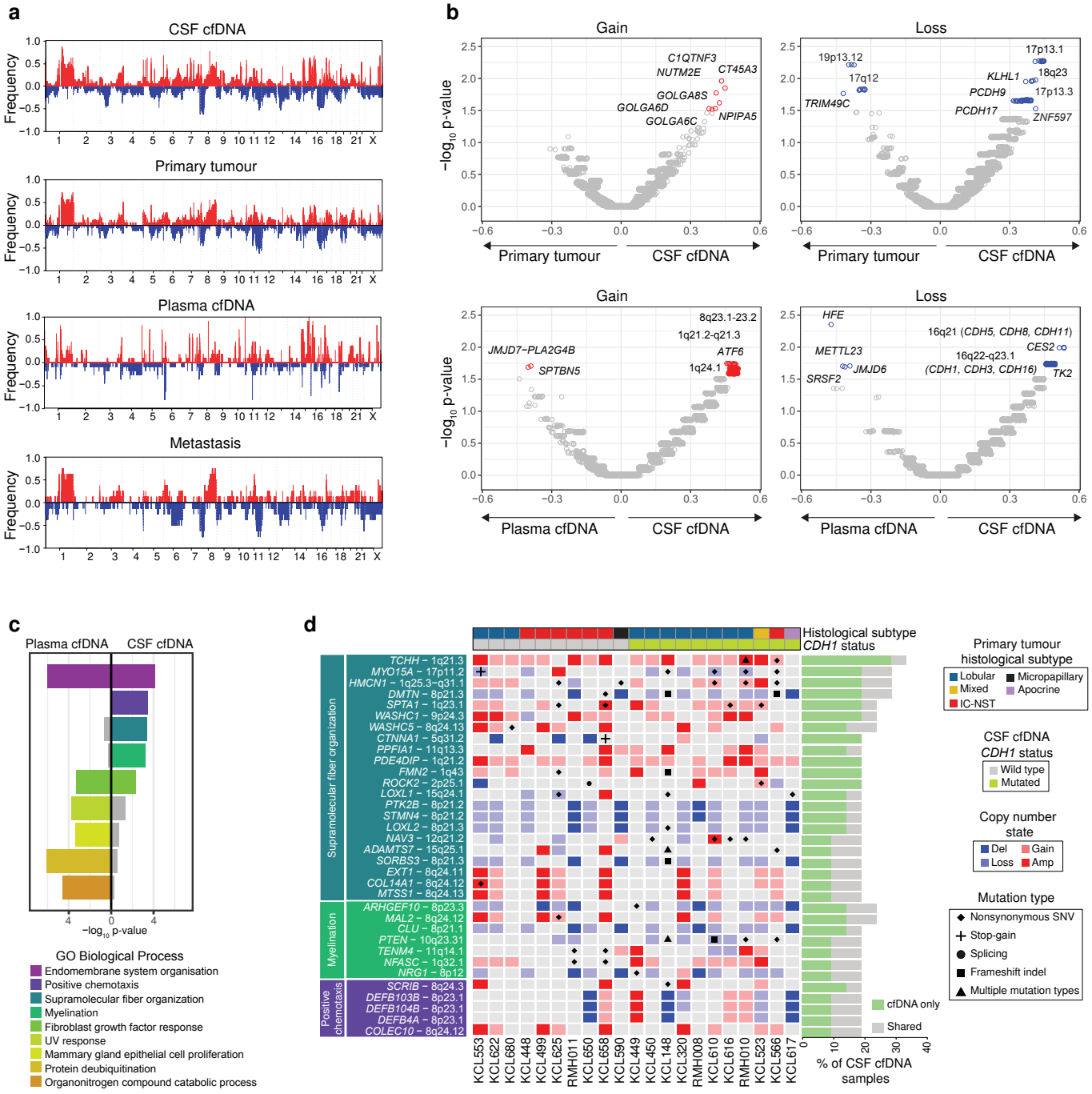

a
CSF cfDNA
Primary tumour
Plasma cfDNA
Metastasis
Frequency
b
Gain
Loss
−log10 p-value
C1QTNF3
NUTM2E
CT45A3
GOLGA8S
GOLGA6D
GOLGA6C
NPIPA5
19p13.12
17q12
TRIM49C
17p13.1
18q23
KLHL1
PCDH9
17p13.3
PCDH17
ZNF597
8q23.1-23.2
1q21.2-q21.3
JMJD7−PLA2G4B
ATF6
1q24.1
SPTBN5
HFE
16q21 (CDH5, CDH8, CDH11)
CES2
METTL23
16q22-q23.1
(CDH1, CDH3, CDH16)
JMJD6
SRSF2
TK2
Primary tumour
CSF cfDNA
Plasma cfDNA
c
Plasma cfDNA
CSF cfDNA
−log10 p-value
GO Biological Process
Endomembrane system organisation
Positive chemotaxis
Supramolecular fiber organization
Myelination
Fibroblast growth factor response
UV response
Mammary gland epithelial cell proliferation
Protein deubiquitination
Organonitrogen compound catabolic process
d
Histological subtype
CDH1 status
Supramolecular fiber organization
Myelination
Positive chemotaxis
TCHH − 1q21.3
MYO15A − 17p11.2
HMCN1 − 1q25.3−q31.1
DMTN − 8p21.3
SPTA1 − 1q23.1
WASHC1 − 9p24.3
WASHC5 − 8q24.13
CTNNA1 − 5q31.2
PPFIA1 − 11q13.3
PDE4DIP − 1q21.2
FMN2 − 1q43
ROCK2 − 2p25.1
LOXL1 − 15q24.1
PTK2B − 8p21.2
STMN4 − 8p21.2
LOXL2 − 8p21.3
NAV3 − 12q21.2
ADAMTS7 − 15q25.1
SORBS3 − 8p21.3
EXT1 − 8q24.11
COL14A1 − 8q24.12
MTSS1 − 8q24.13
ARHGEF10 − 8p23.3
MAL2 − 8q24.12
CLU − 8p21.1
PTEN − 10q23.31
TENM4 − 11q14.1
NFASC − 1q32.1
NRG1 − 8p12
SCRIB − 8q24.3
DEFB103B − 8p23.1
DEFB104B − 8p23.1
DEFB4A − 8p23.1
COLEC10 − 8q24.12
KCL553
KCL622
KCL680
KCL448
KCL499
KCL625
RMH011
KCL650
KCL658
KCL590
KCL449
KCL450
KCL148
KCL320
RMH008
KCL610
KCL616
RMH010
KCL523
KCL566
KCL617
cfDNA only
Shared
% of CSF cfDNA samples
Primary tumour histological subtype
Lobular
Micropapillary
Mixed
Apocrine
IC-NST
CSF cfDNA CDH1 status
Wild type
Mutated
Copy number state
Del
Gain
Loss
Amp
Mutation type
Nonsynonymous SNV
Stop-gain
Splicing
Frameshift indel
Multiple mutation types

Figure 5

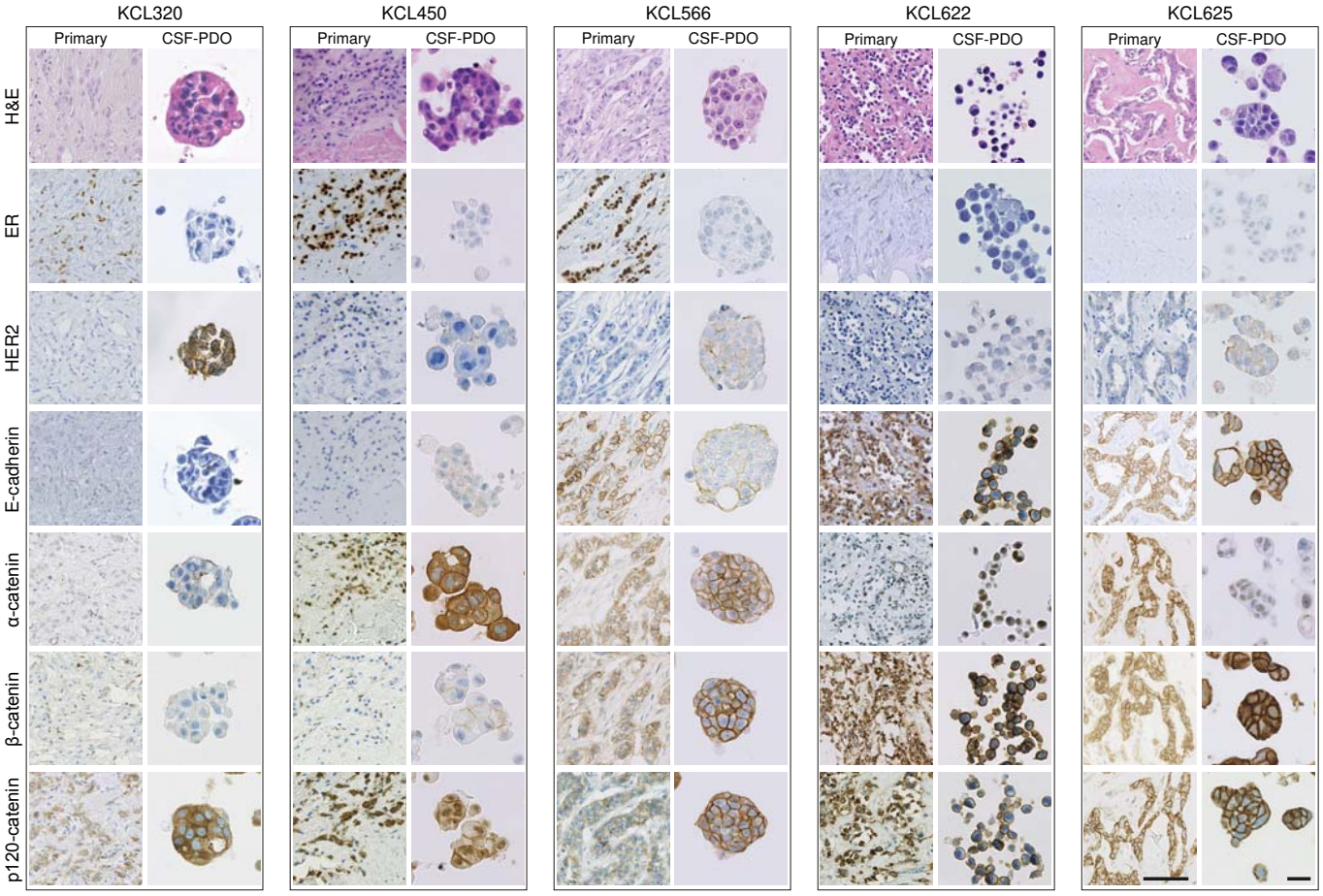

KCL320
KCL450
KCL566
KCL622
KCL625
Primary
CSF-PDO
H&E
ER
HER2
E-cadherin
α-catenin
β-catenin
p120-catenin

Figure 6

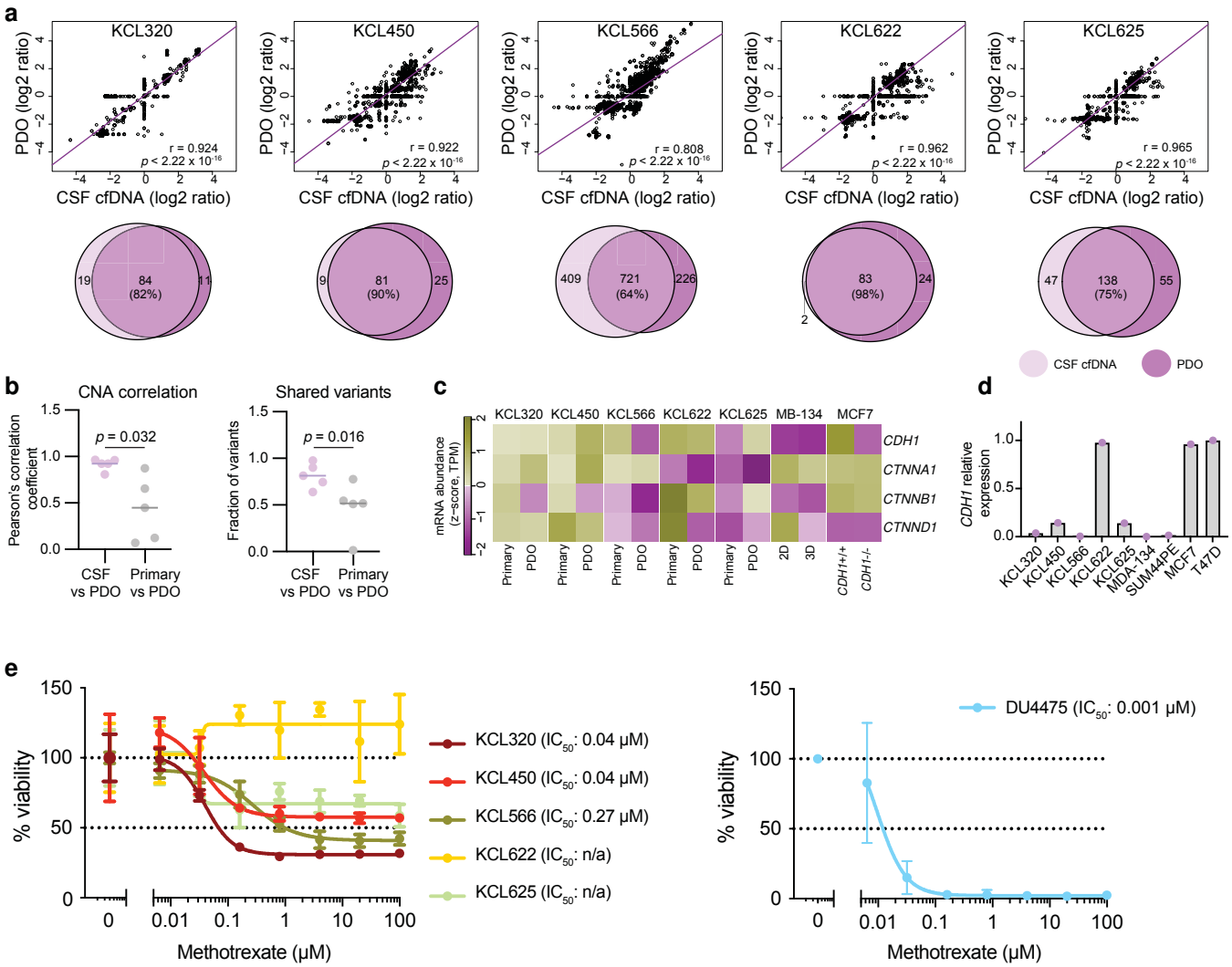
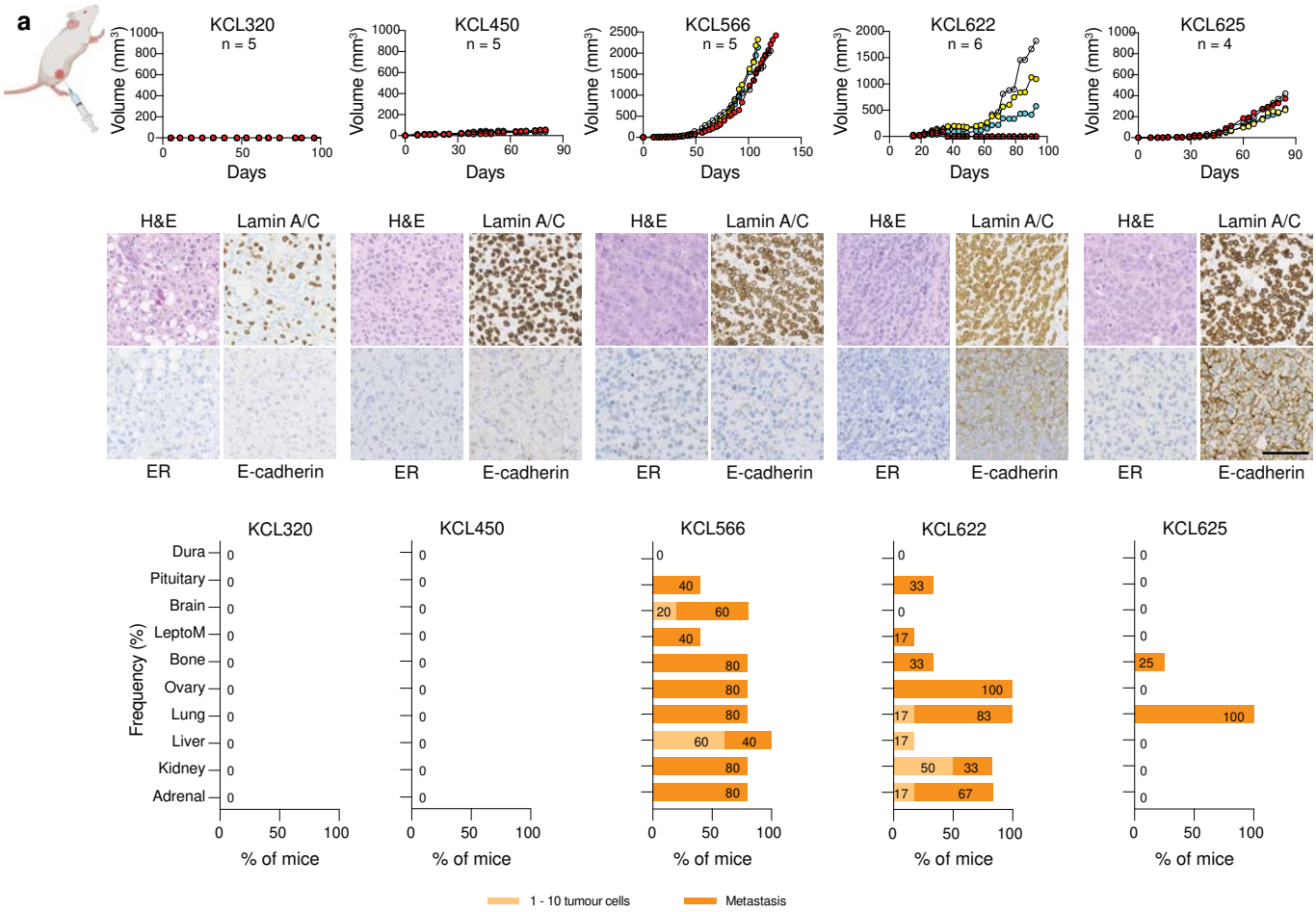

a
KCL320
KCL450
KCL566
KCL622
KCL625
PDO (log2 ratio)
CSF cfDNA (log2 ratio)
r = 0.924
r = 0.922
r = 0.808
r = 0.962
r = 0.965
p < 2.22 x 10-16
19
84 (82%)
11
9
81 (90%)
25
409
721 (64%)
226
83 (98%)
24
2
47
138 (75%)
55
CSF cfDNA
PDO
b
CNA correlation
Shared variants
Pearson's correlation coefficient
Fraction of variants
p = 0.032
p = 0.016
CSF vs PDO
Primary vs PDO
c
KCL320
KCL450
KCL566
KCL622
KCL625
MB-134
MCF7
mRNA abundance (z-score, TPM)
CDH1
CTNNA1
CTNNB1
CTNND1
Primary
PDO
2D
3D
CDH1+/+
CDH1-/-
d
CDH1 relative expression
KCL320
KCL450
KCL566
KCL622
KCL625
MDA-134
SUM44PE
MCF7
T47D
e
% viability
Methotrexate (μM)
KCL320 (IC50: 0.04 μM)
KCL450 (IC50: 0.04 μM)
KCL566 (IC50: 0.27 μM)
KCL622 (IC50: n/a)
KCL625 (IC50: n/a)
DU4475 (IC50: 0.001 μM)

Figure 7


a
KCL320
n = 5
KCL450
n = 5
KCL566
n = 5
KCL622
n = 6
KCL625
n = 4
Volume (mm³)
Days
H&E
Lamin A/C
ER
E-cadherin
Frequency (%)
Dura
Pituitary
Brain
LeptoM
Bone
Ovary
Lung
Liver
Kidney
Adrenal
% of mice
1 - 10 tumour cells
Metastasis

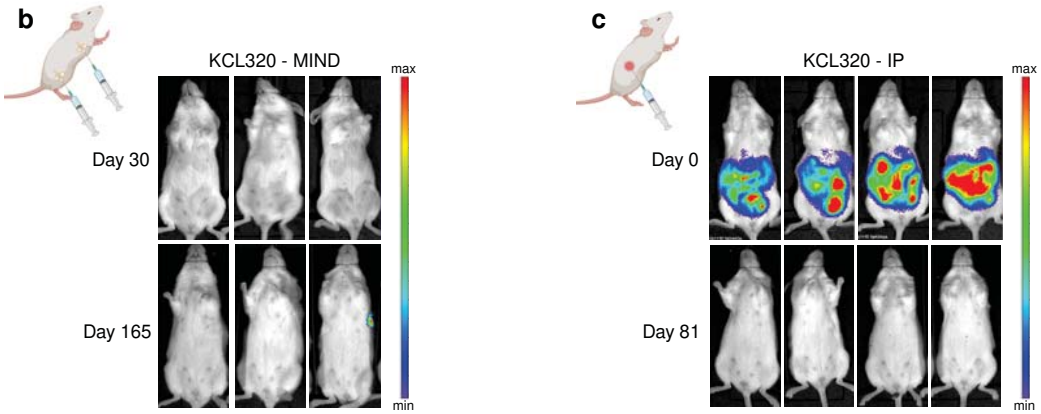
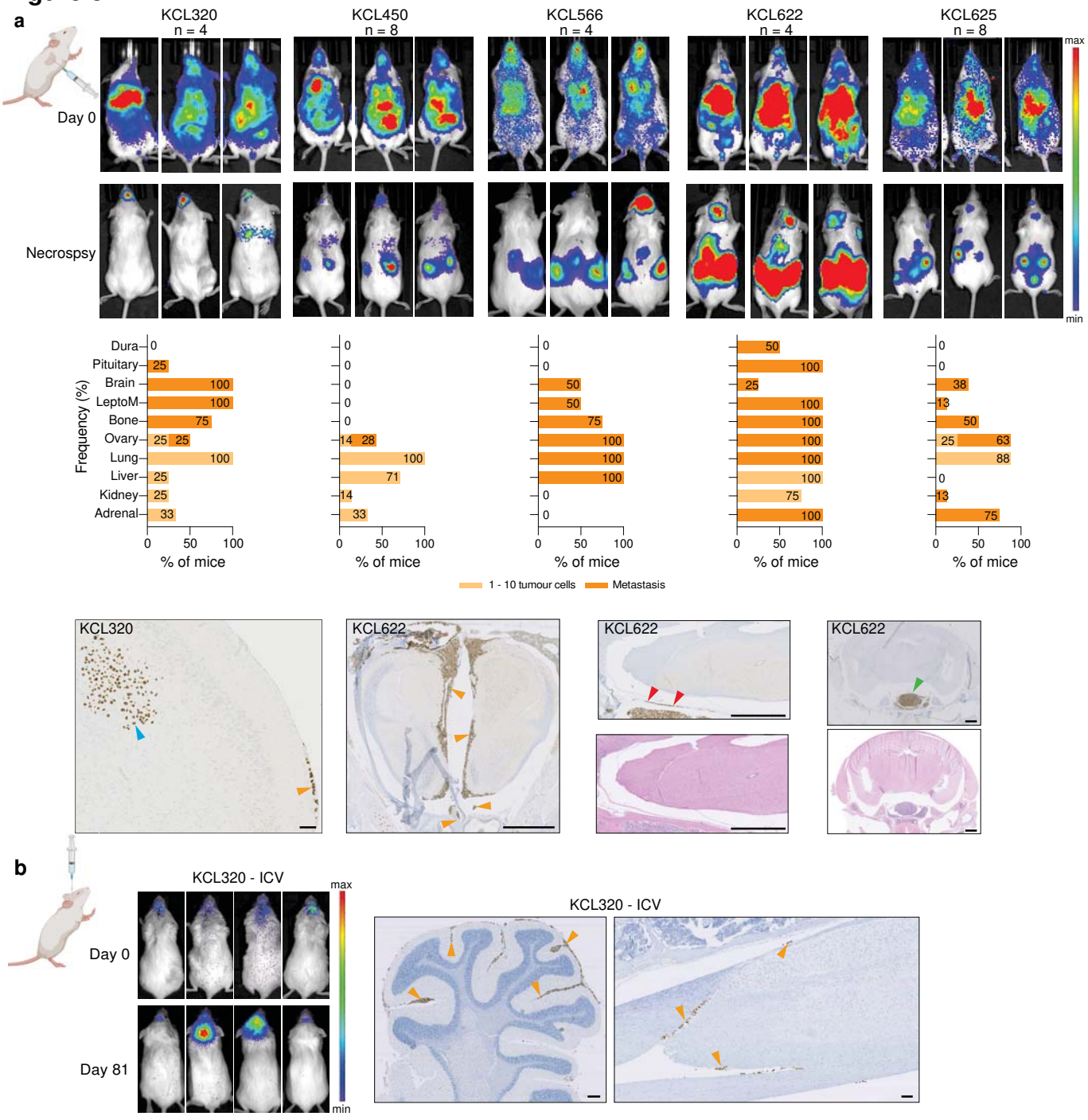

b
KCL320 - MIND
Day 30
Day 165
max
min
c
KCL320 - IP
Day 0
Day 81
max
min

Figure 8


a
KCL320
n = 4
KCL450
n = 8
KCL566
n = 4
KCL622
n = 4
KCL625
n = 8
Day 0
Necrospsy
max
min
Frequency (%)
Dura
Pituitary
Brain
LeptoM
Bone
Ovary
Lung
Liver
Kidney
Adrenal
% of mice
1 - 10 tumour cells
Metastasis
KCL320
KCL622
KCL622
KCL622
b
KCL320 - ICV
Day 0
Day 81
max
min
KCL320 - ICV

Genomic profiling and pre-clinical modelling of breast cancer leptomeningeal metastasis reveals acquisition of a lobular-like phenotype

Amanda Fitzpatrick, Marjan Iravani, Adam Mills, David Vicente, Thanussuyah Alaguthurai, Ioannis Roxanis, Nicholas C. Turner, Syed Haider, Andrew N.J. Tutt and Clare M. Isacke

Supplementary Material

Supplementary Table 1 | Clinicopathological characteristics of patients in BCLM cohort

Supplementary Table 2 | Whole exome sequencing and variant calling metrics

Supplementary Table 3 | Somatic variants identified by WES (excluding synonymous SNVs)

Supplementary Table 4 | Chemotherapy-related SBS mutational signatures

Supplementary Table 5 | SBS mutational signatures across sample types

Supplementary Table 6 | Characteristics of BCLM genomic evolution

Supplementary Table 7 | Frequently altered cancer driver genes in plasma cfDNA

Supplementary Table 8 | Genomic alterations in BCLM *vs*. MBC cohort

Supplementary Table 9 | Copy number alterations (standardised log2 ratio) determined by CNVkit analysis of WES data

Supplementary Table 10 | OncoKB actionable gene alterations in CSF cfDNA

Supplementary Table11 | Immunohistochemical scoring of primary tumours and matched BCLM PDOs

Supplementary Table 12 | STR authentication of cell lines and PDOs

Supplementary Table 13 | Antibodies used for immunohistochemistry

Supplementary Figure 1 | Metrics of samples undergoing WES

Supplementary Figure 2 | Trinucleotide context of single nucleotide substitutions

Supplementary Figure 3 | Correlation of CSF cfDNA unique variants with clinical history

Supplementary Figure 4 | BCLM clonal evolution modelling

Supplementary Figure 5 | Cancer-associated mutation and copy number landscape of all matched samples

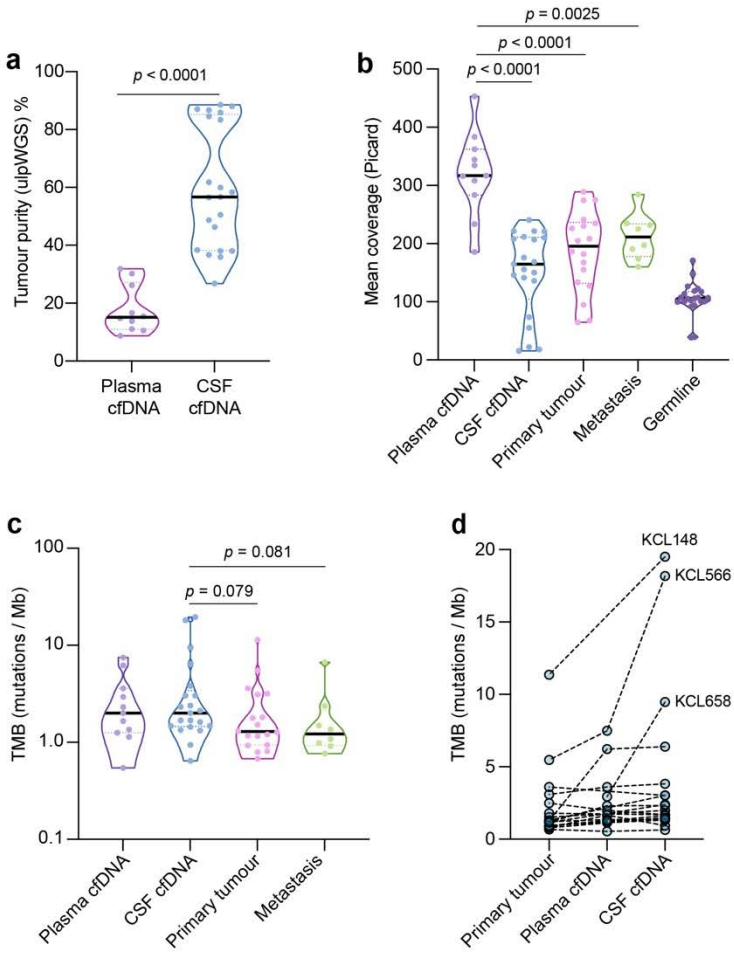
Supplementary Figure 6 | Comparison of CNA frequency between sample type

Supplementary Figure 7 | BCLM-enriched genomic alterations in biological processes

Supplementary Figure 8 | Comparison of genomic findings in BCLM PDOs and matched primary tumour, plasma cfDNA or metastatic tissue samples

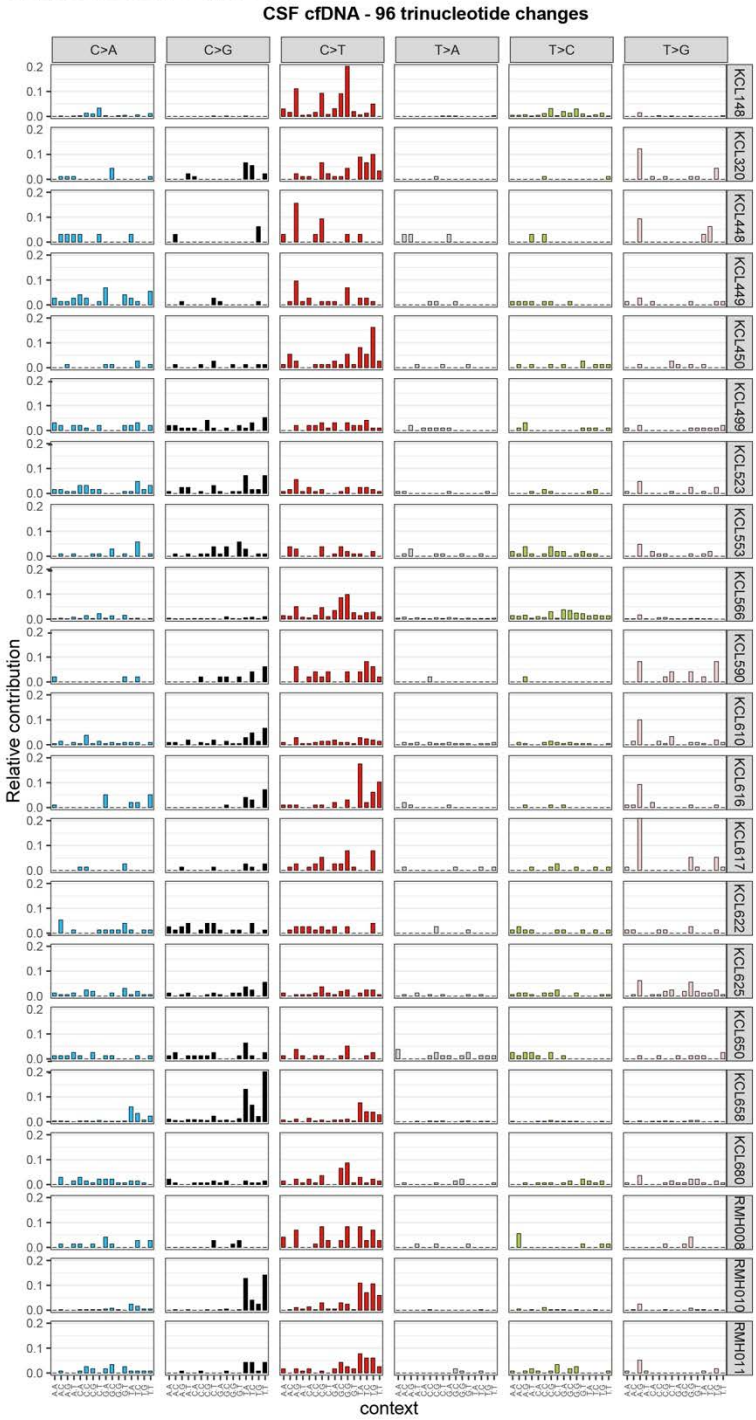
Supplementary Figure 9 | Transduction of BCLM PDOs to express mCherry and luciferase

Supplementary Figure 1



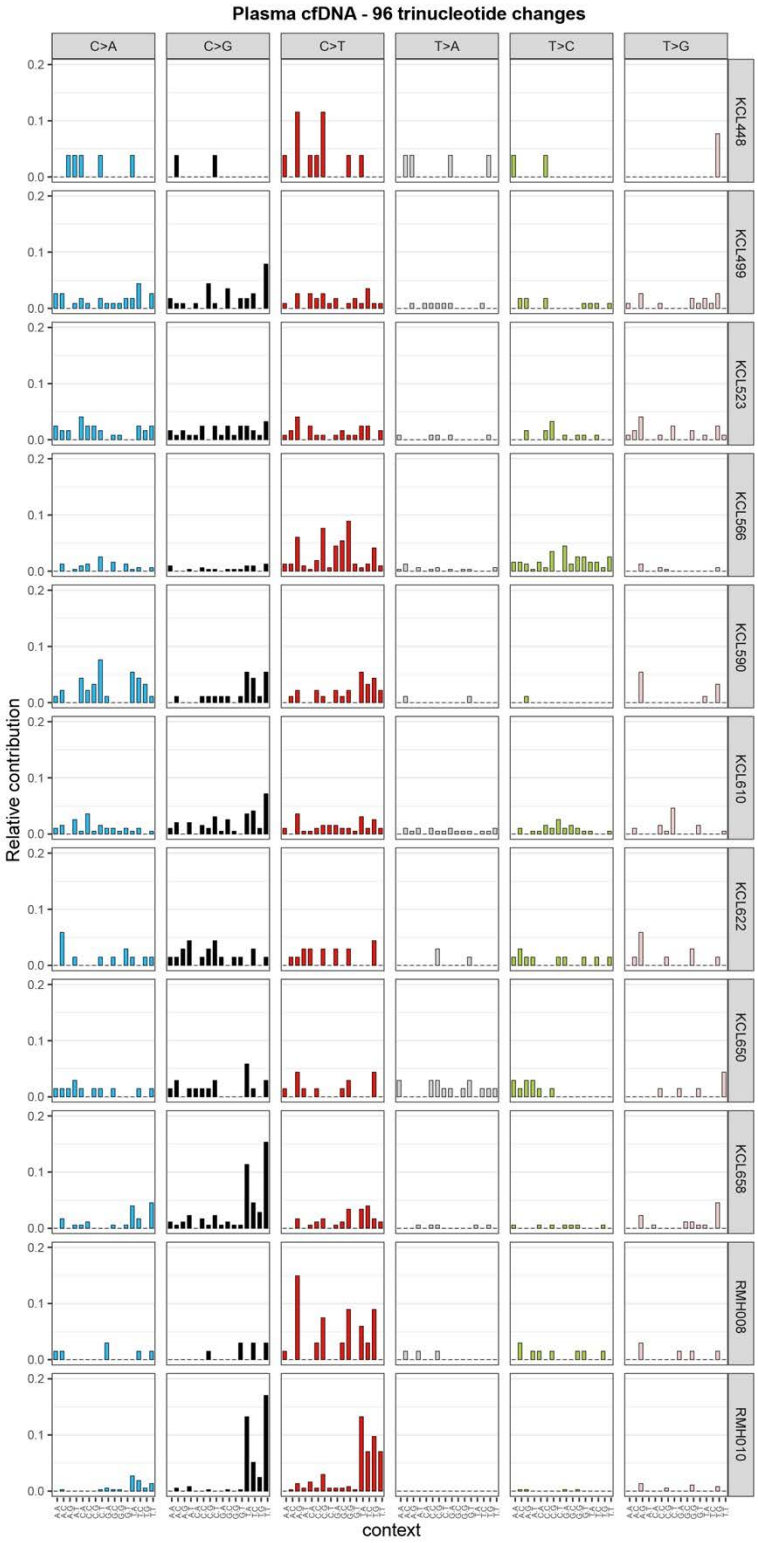
Supplementary Fig. 1. Metrics of samples undergoing WES**. a** Tumour purity of cfDNA assessed by ulpWGS (ultra-low-pass whole genome sequencing) within the concurrently collected CSF (n = 21) and plasma (n = 11) samples. The 10 plasma cfDNA samples which did not undergo WES (due to tumour content <10%) are not shown. Significance by Mann-Whitney test, two-tailed. **b** Sequencing coverage of all samples. Plasma cfDNA, n = 11; CSF cfDNA, n = 21; primary tumour, n = 18; metastasis, n = 8; germline, n = 21. Significance by Mann-Whitney test, two-tailed. **c** Tumour mutational burden (TMB) in samples shown in panel b. Significance by Mann-Whitney test, two-tailed, p-values <0.1 are shown. **d** Matched comparison of TMB in primary tumour, plasma cfDNA and CSF cfDNA samples collected from individual patients. Source data for all panels are provided as a Source Data file.

Supplementary Figure 2a



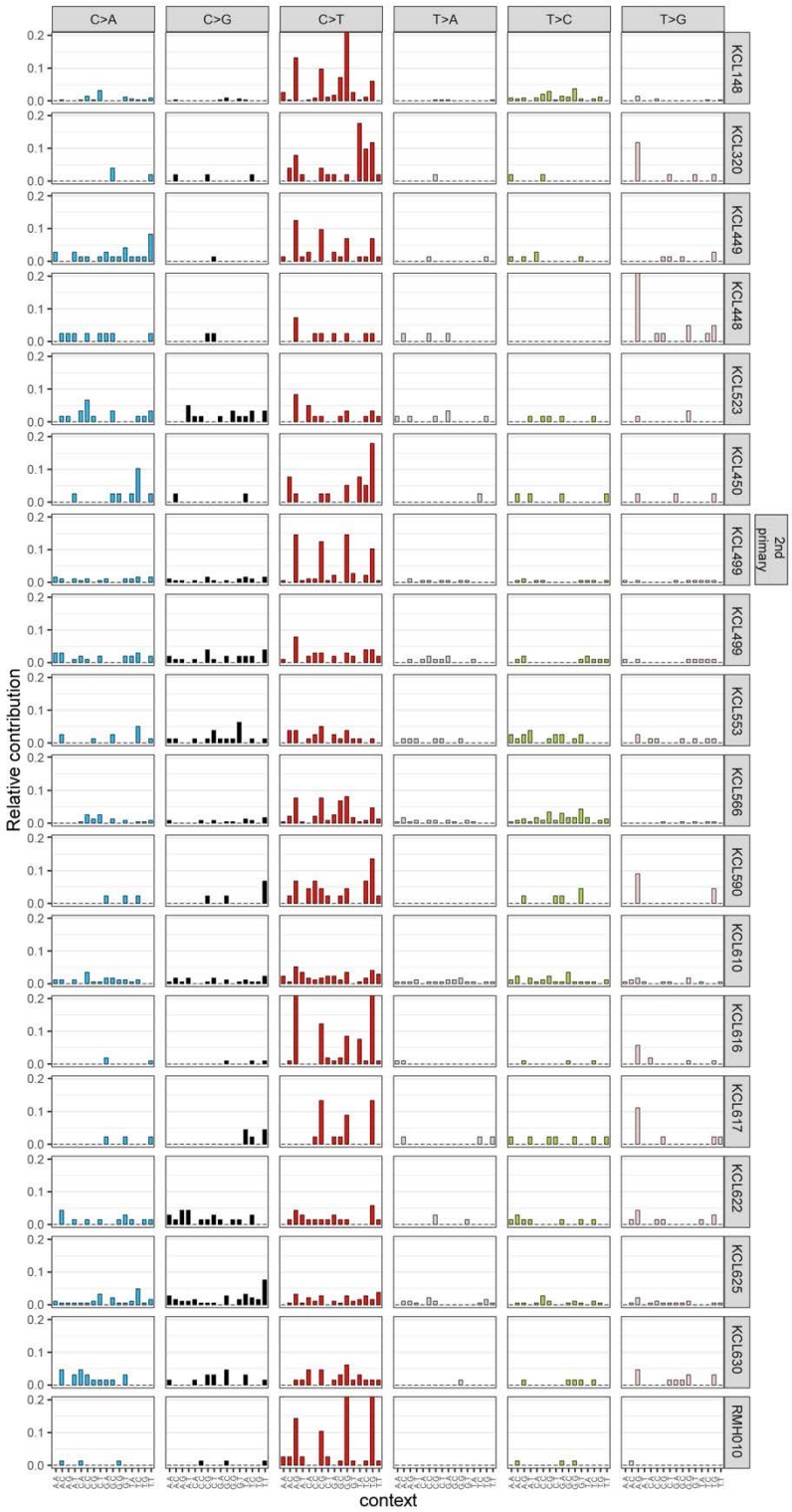
Supplementary Fig. 2. Trinucleotide context of single nucleotide substitutions. **a-d** Data associated with Figure 1c and Supplementary Tables 4 and 5. The x-axis indicates the 5′ and 3′ nucleotides for each of the top panel substitutions for the three base-pair motifs. The y-axis shows the single-base substitution (SBS) composition of each sample by the 96 trinucleotide sequence motifs, grouped by sample type; (**a**) CSF cfDNA, (**b**) plasma cfDNA, (**c**) primary tumour and (**d**) metastasis.

Supplementary Figure 2b

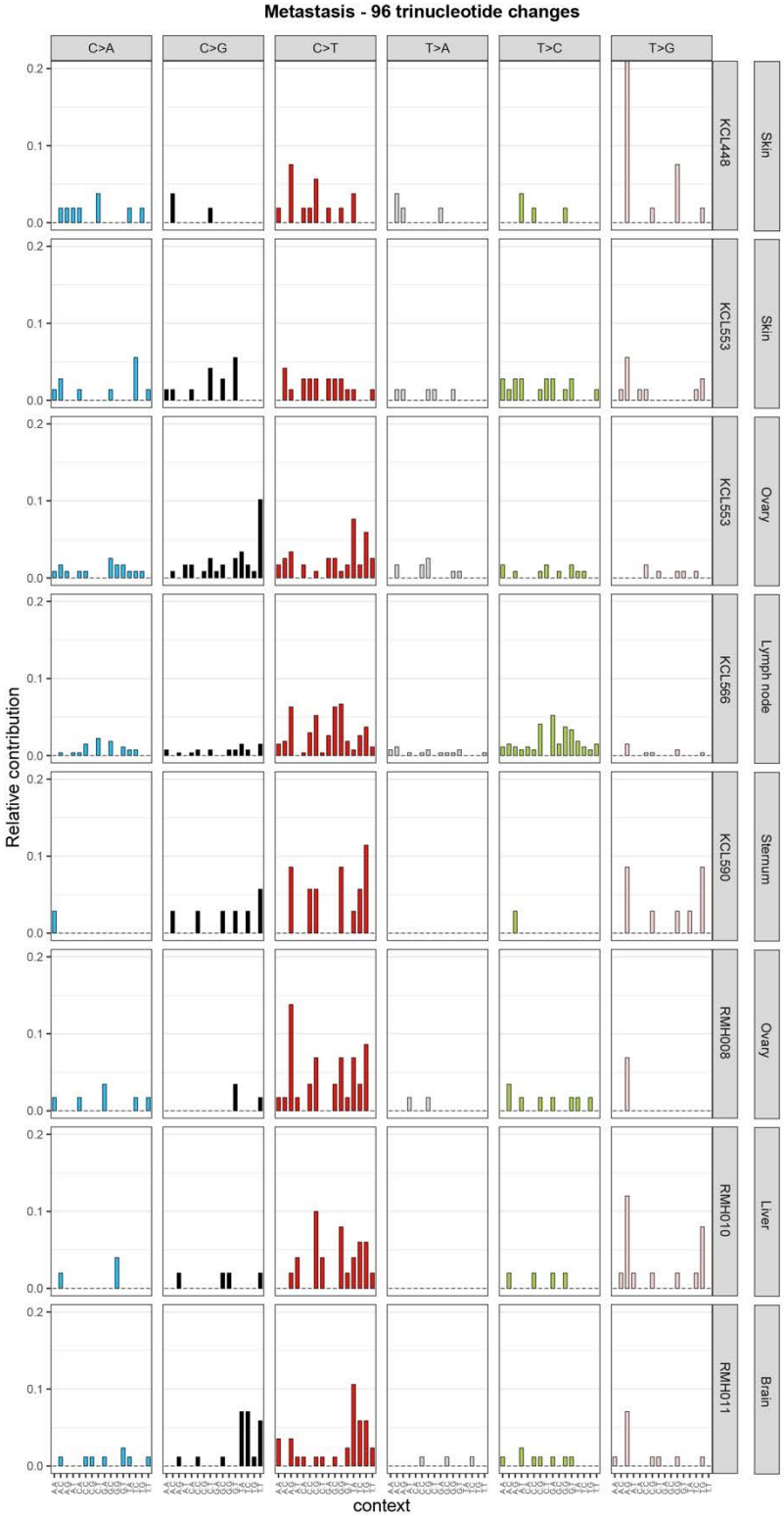


Supplementary Figure 2c

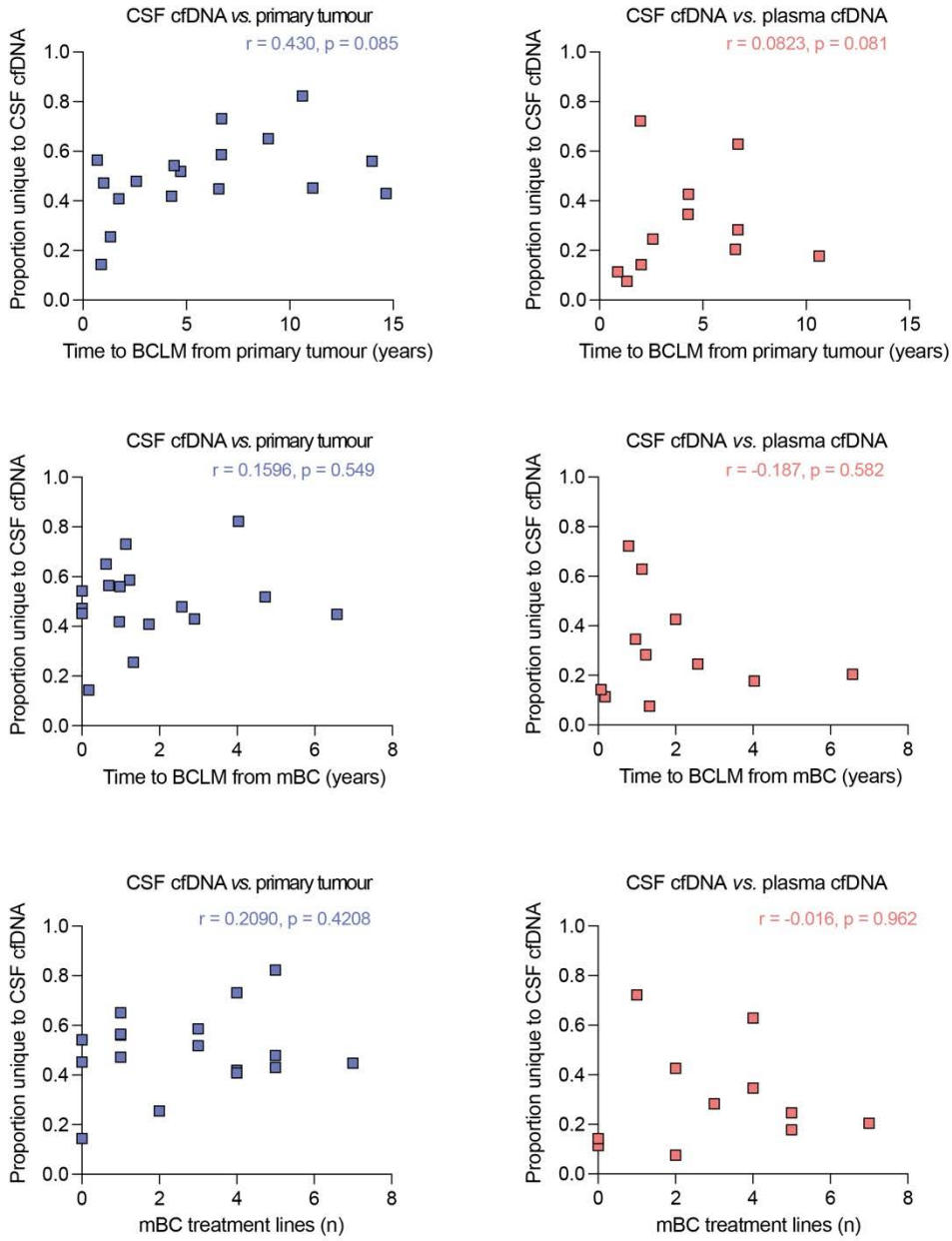
Primary tumour - 96 trinucleotide changes



Supplementary Figure 2d

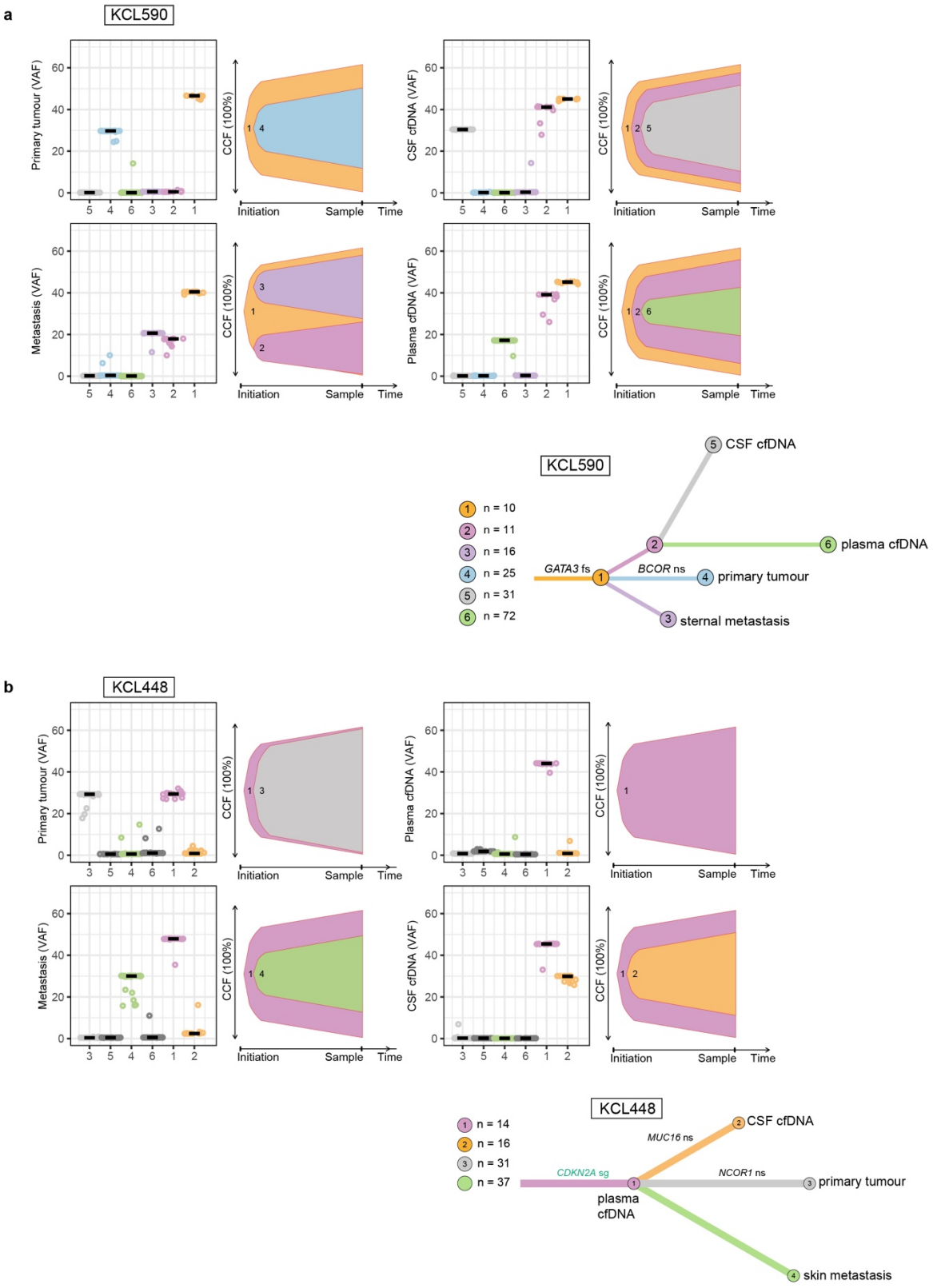


Supplementary Figure 3



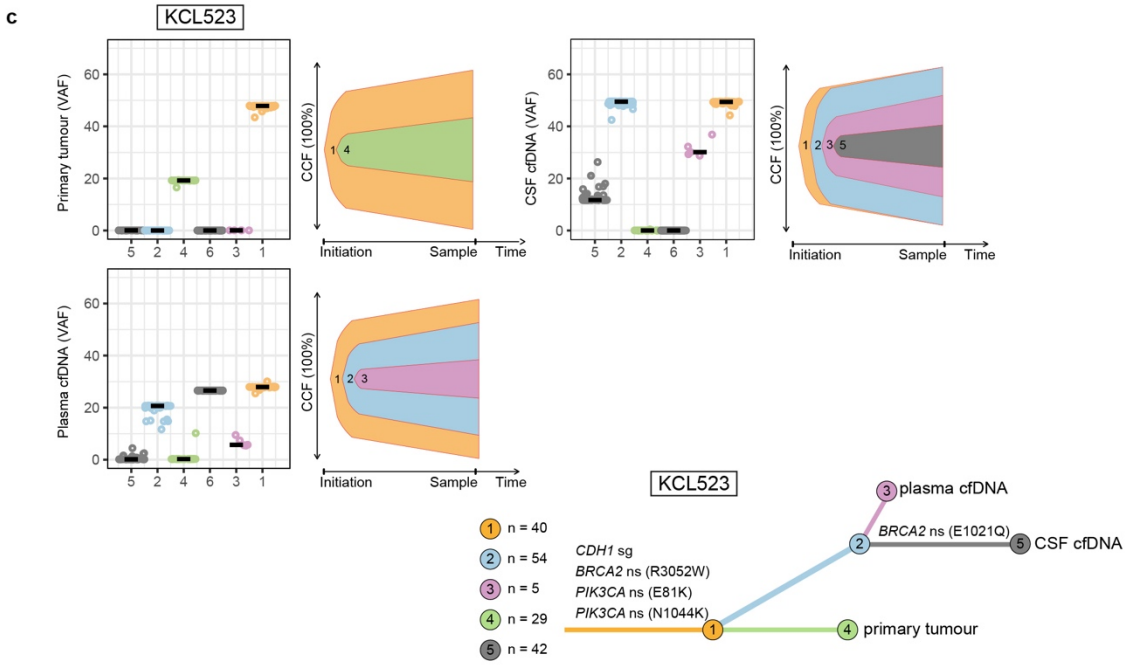
Supplementary Fig. 3. Correlation of CSF cfDNA unique variants with clinical history**.** Correlation of proportion of variants unique to CSF cfDNA *vs*. primary tumour (left panels) (*n* = 17 pairs) and CSF cfDNA *vs.* matched plasma cfDNA (*n* = 11 pairs) (right panels) with time to BCLM diagnosis from primary tumour diagnosis, time to BCLM diagnosis time from diagnosis of metastatic breast cancer, and to number of previous systemic therapies (treatment lines) completed. Pearson's r correlation, two-tailed p-values. Source data are provided as a Source Data file.

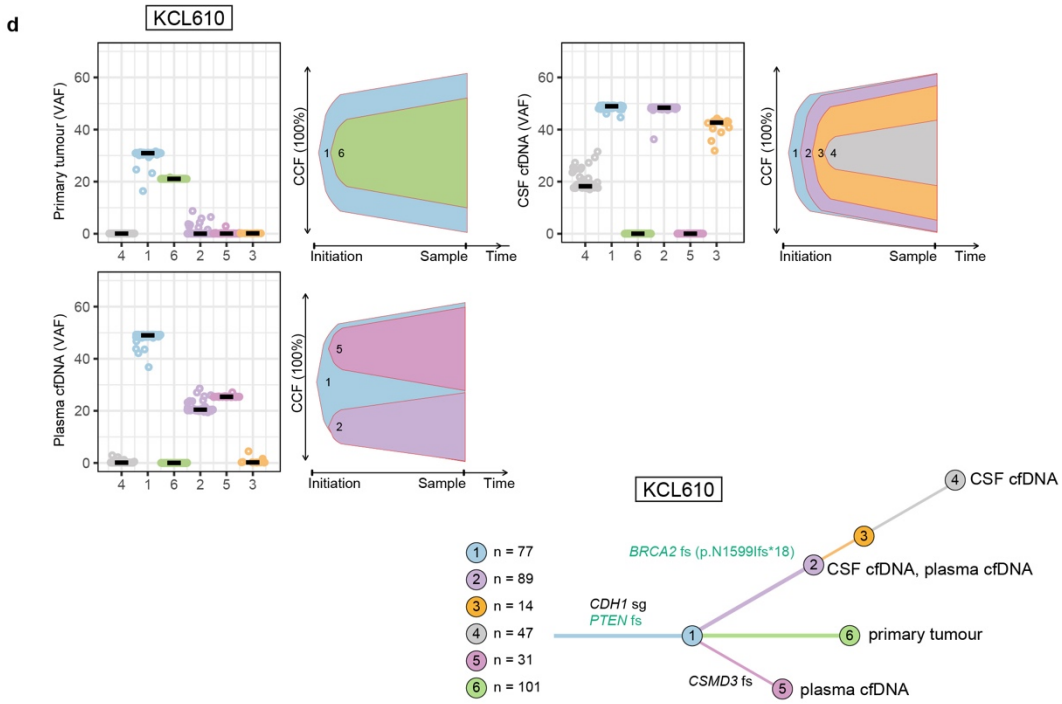
Supplementary Figure 4



(see Page 17 for Supplementary Figure 4 legend)

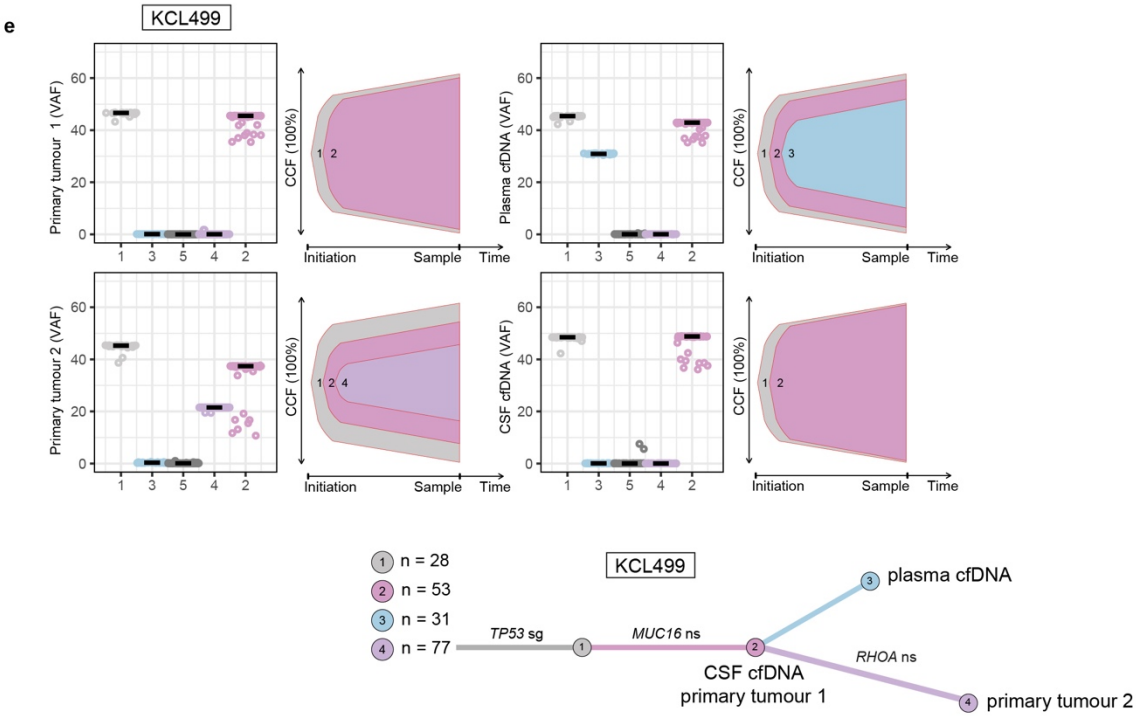
Supplementary Figure 4 (continued)

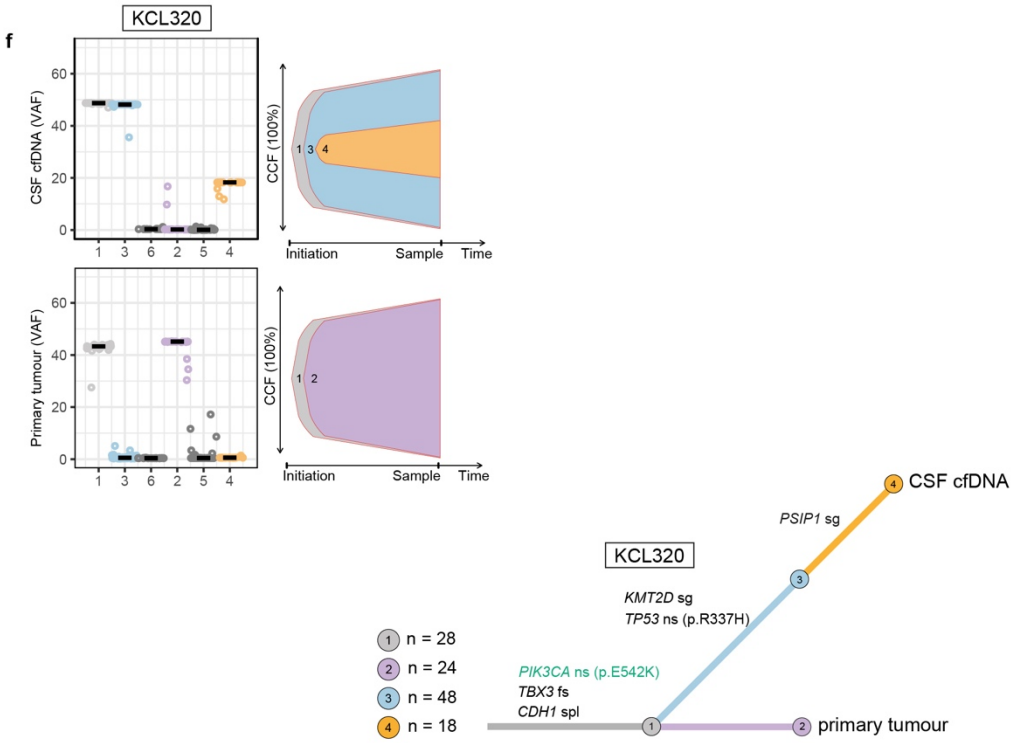




(see Page 17 for Supplementary Figure 4 legend)

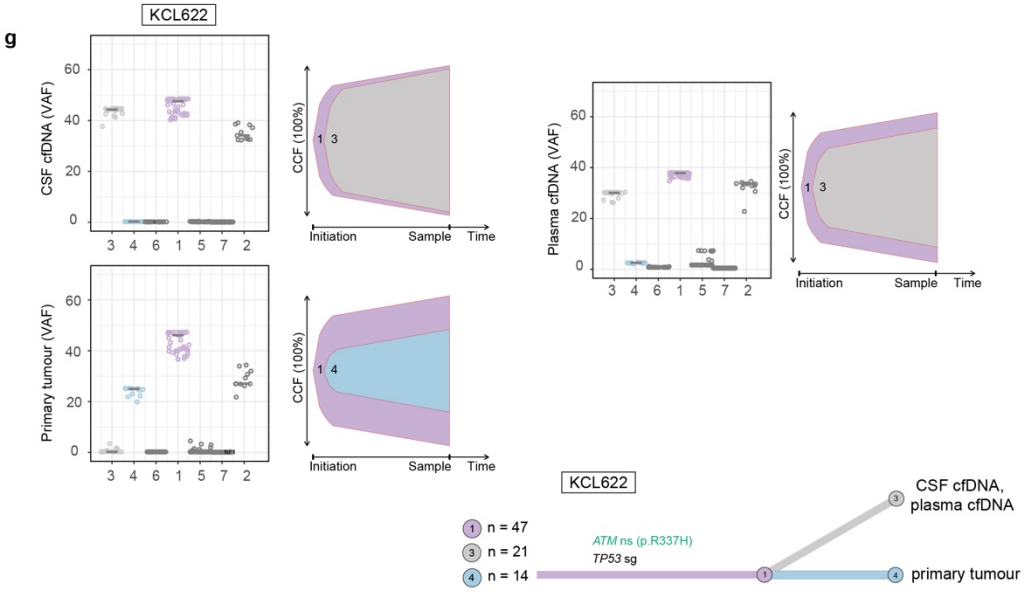
Supplementary Figure 4 (continued)

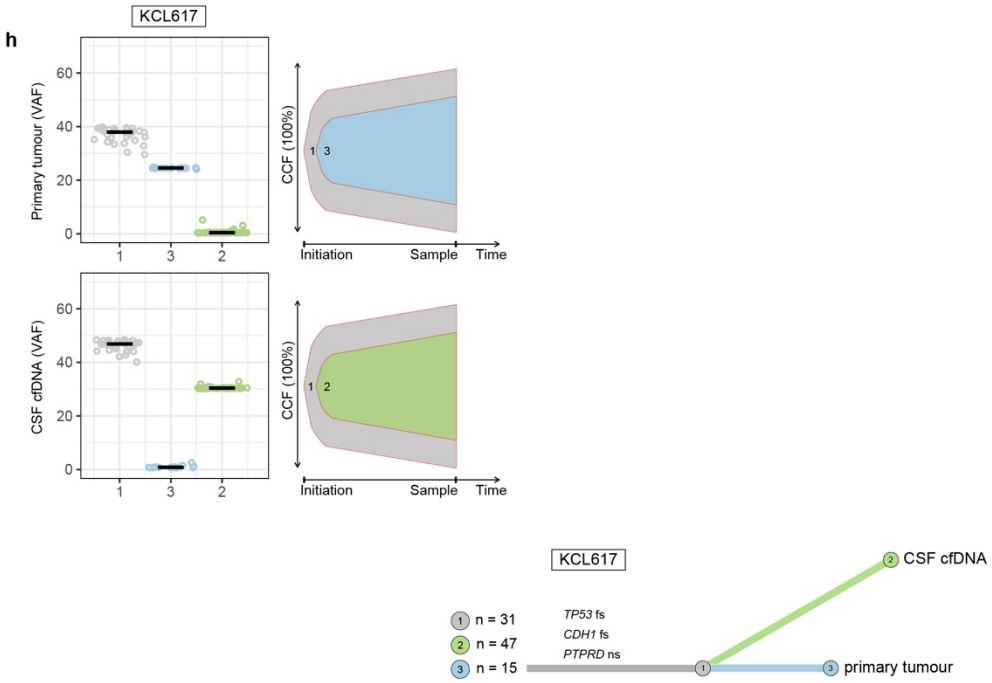




(see Page 17 for Supplementary Figure 4 legend)

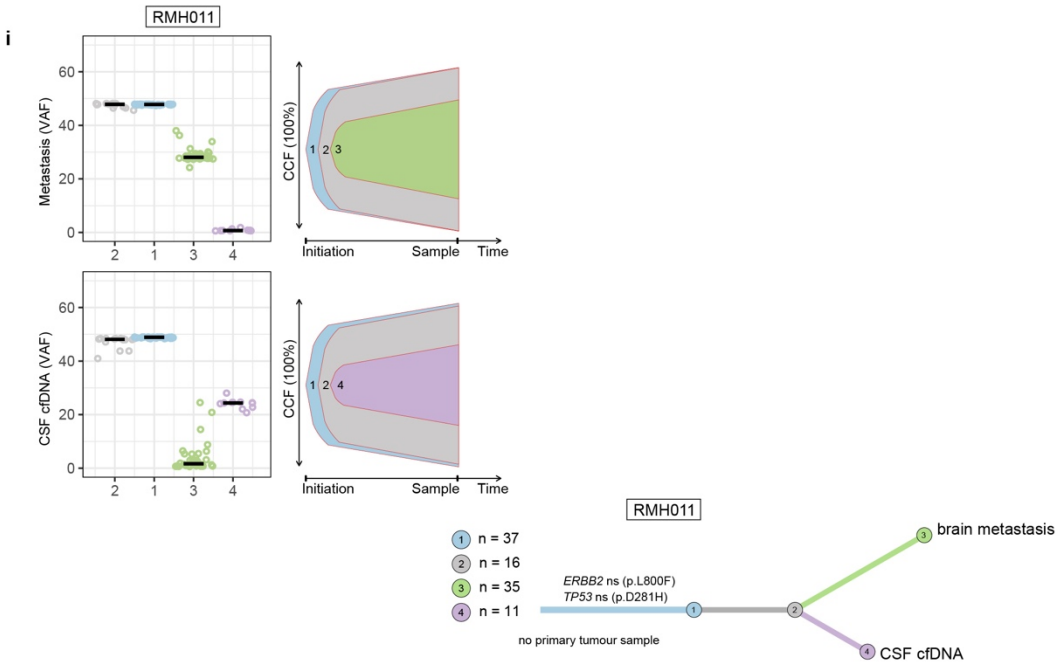
Supplementary Figure 4 (continued)

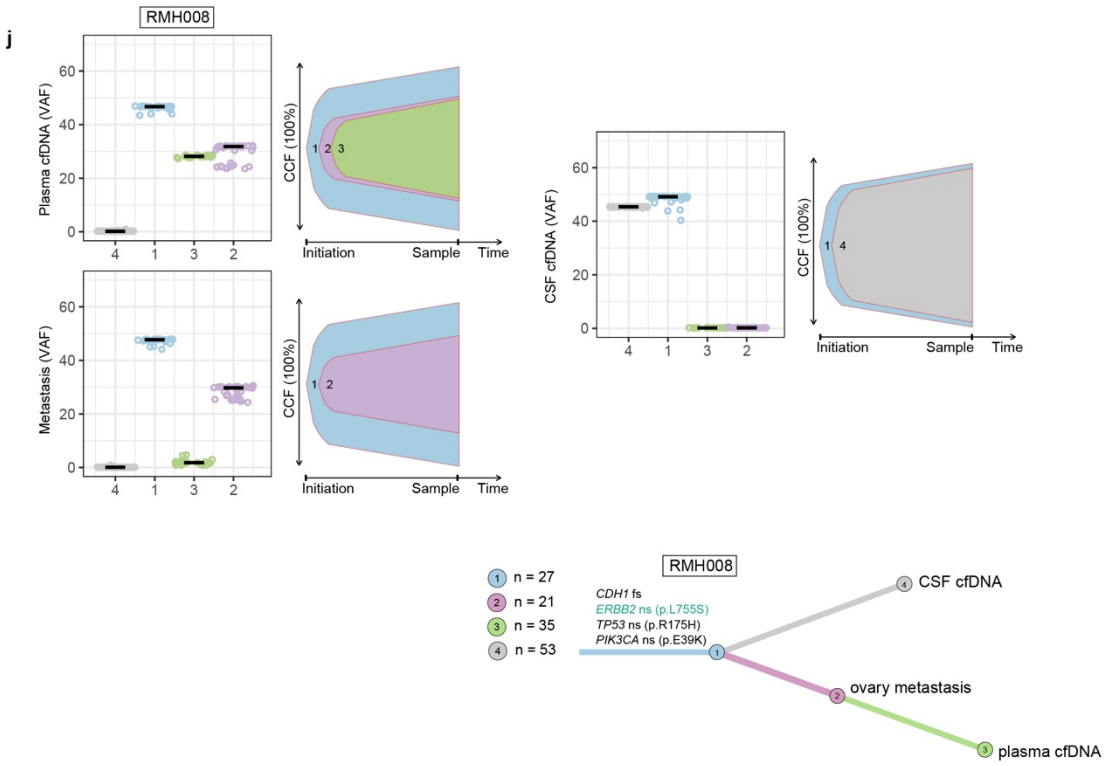




(see Page 17 for Supplementary Figure 4 legend)

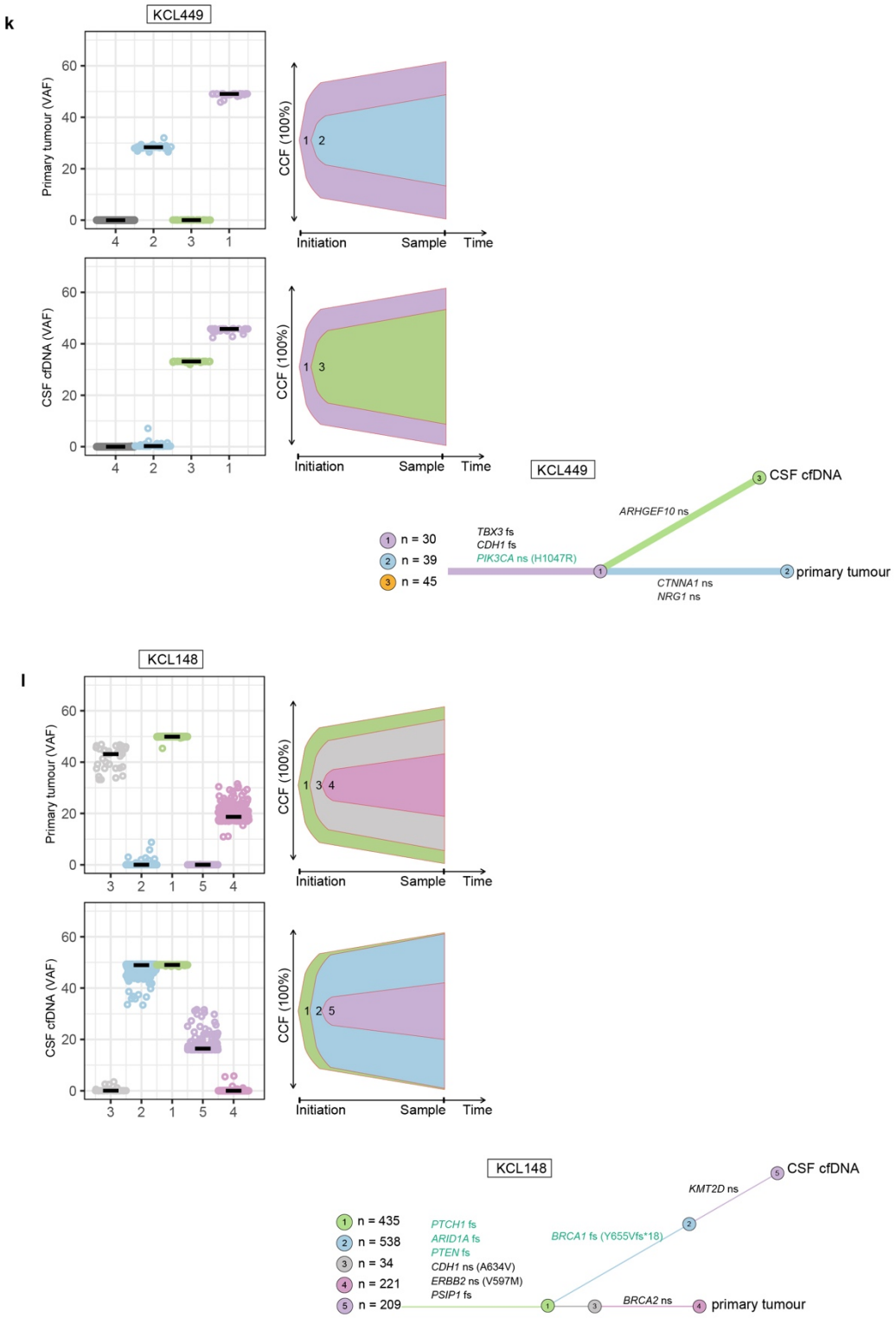
Supplementary Figure 4 (continued)





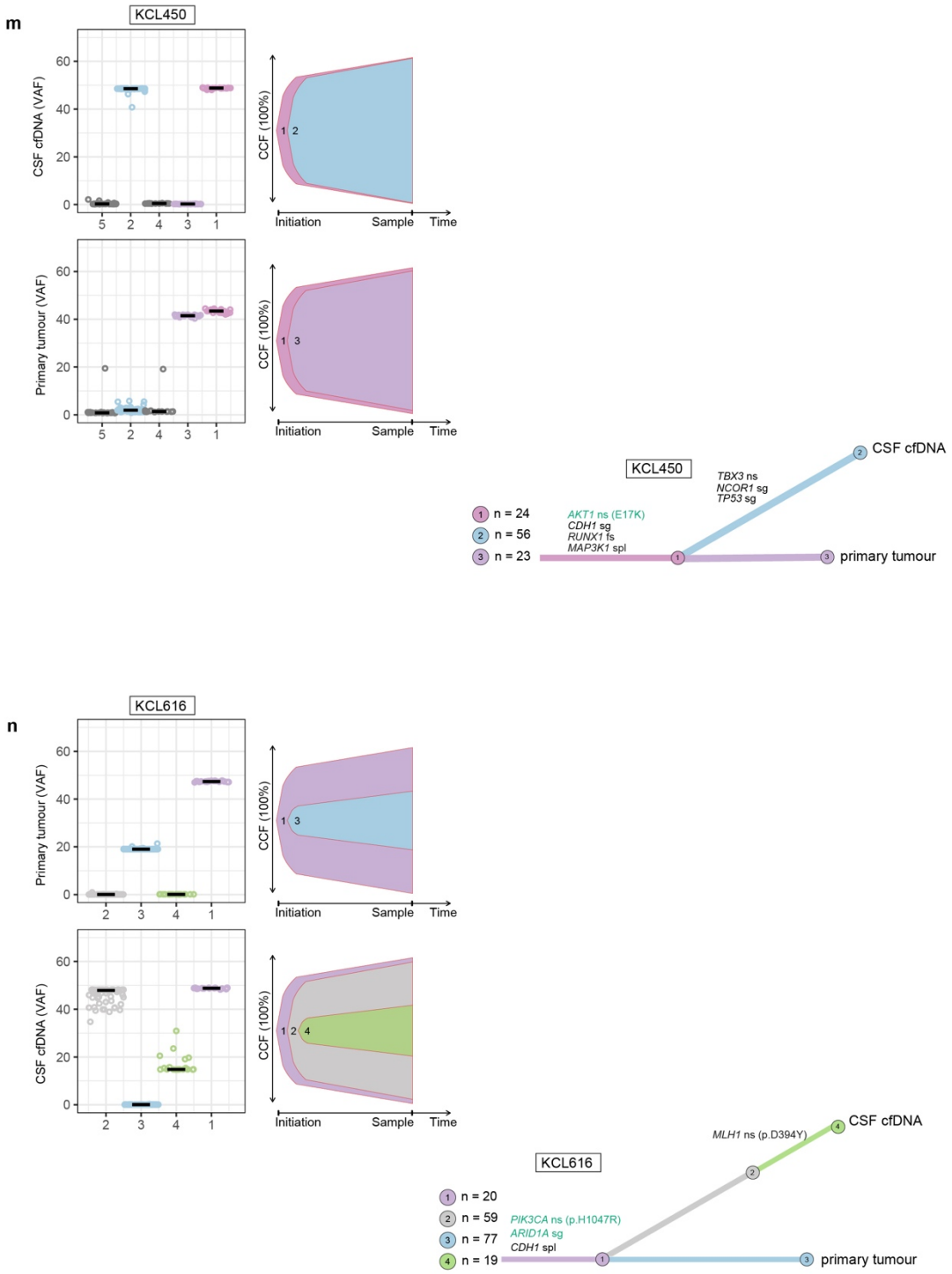
(see Page 17 for Supplementary Figure 4 legend)

Supplementary Figure 4 (continued)



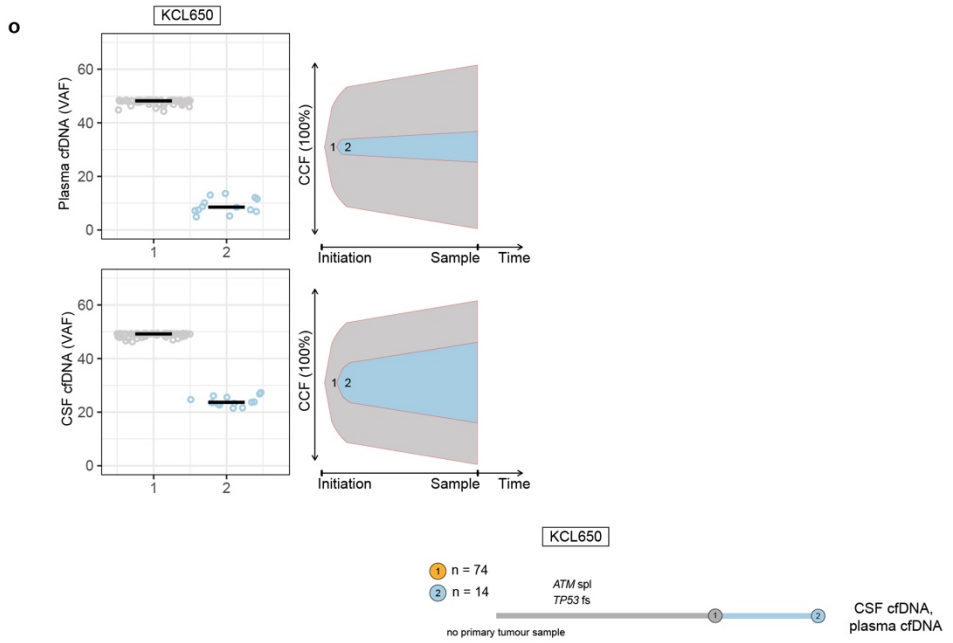
(see Page 17 for Supplementary Figure 4 legend)

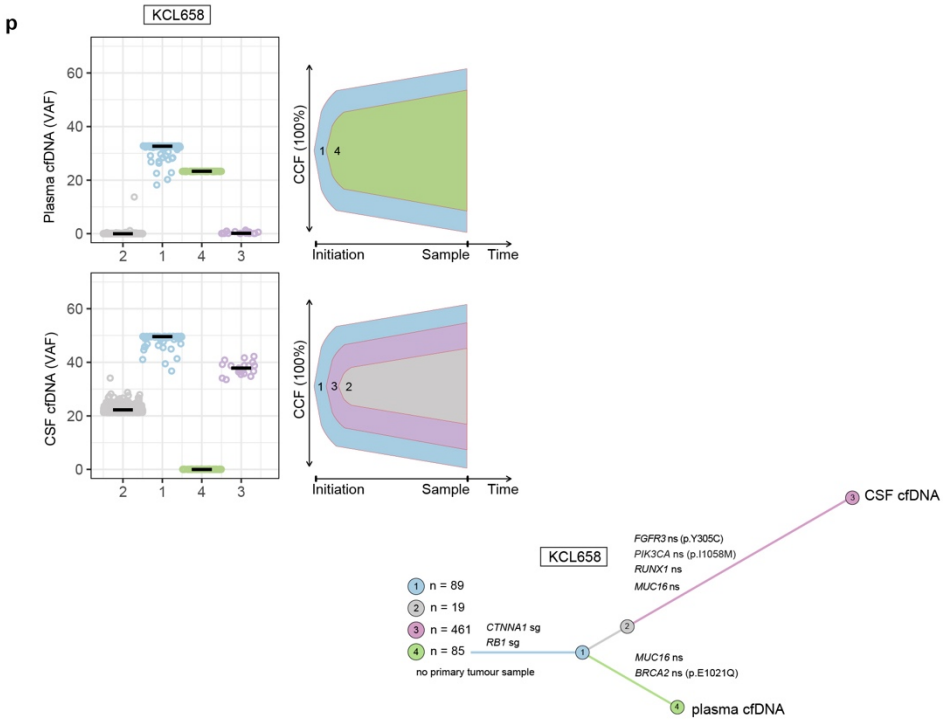
Supplementary Figure 4 (continued)



(see Page 17 for Supplementary Figure 4 legend)

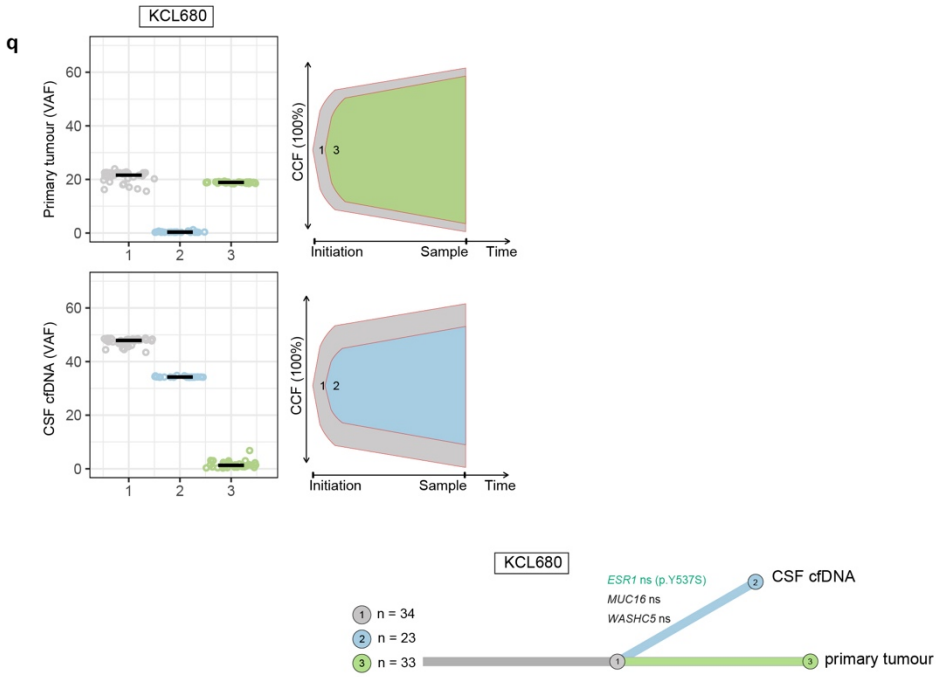
Supplementary Figure 4 (continued)





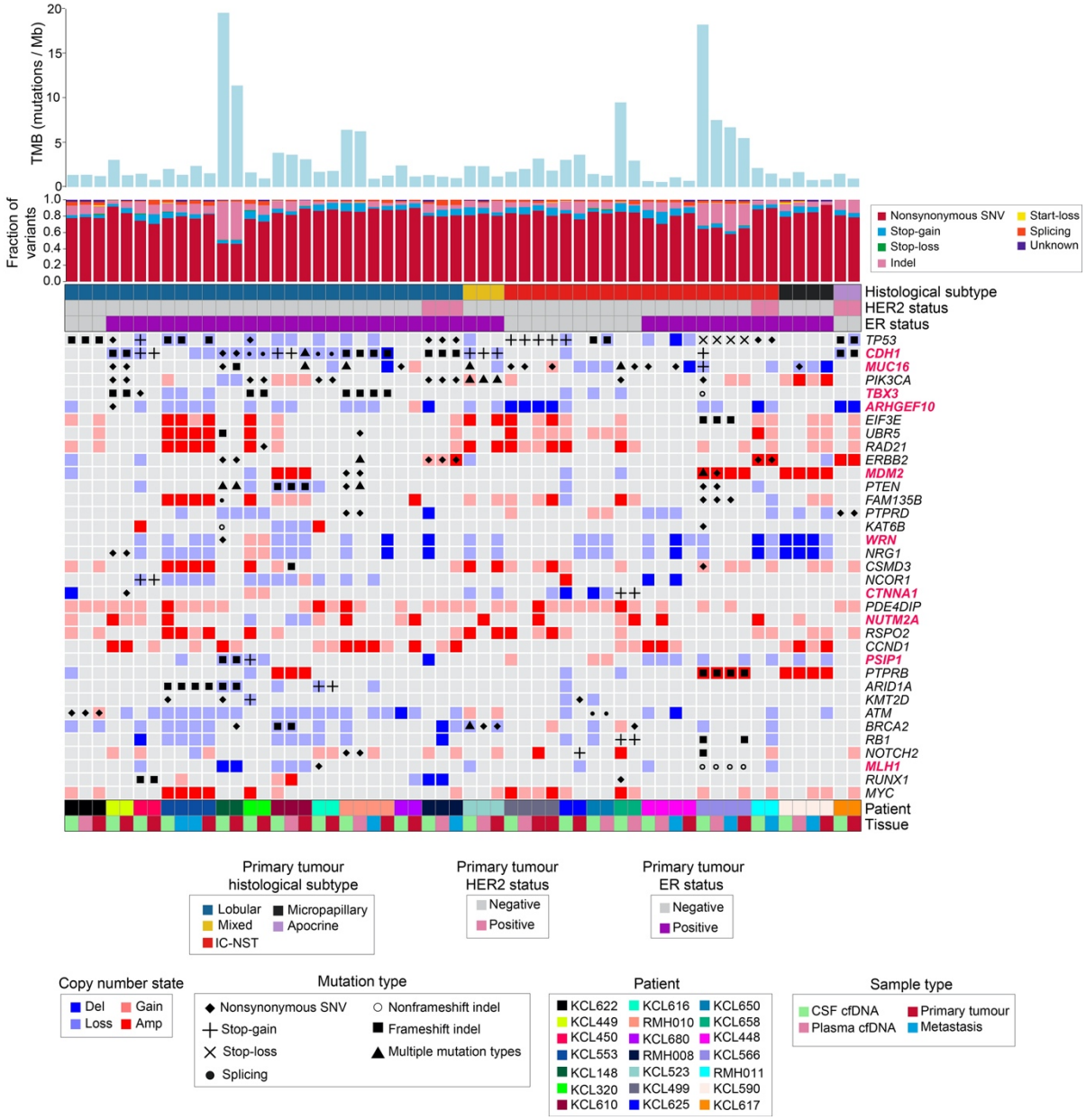
(see Page 17 for Supplementary Figure 4 legend)

Supplementary Figure 4 (continued)



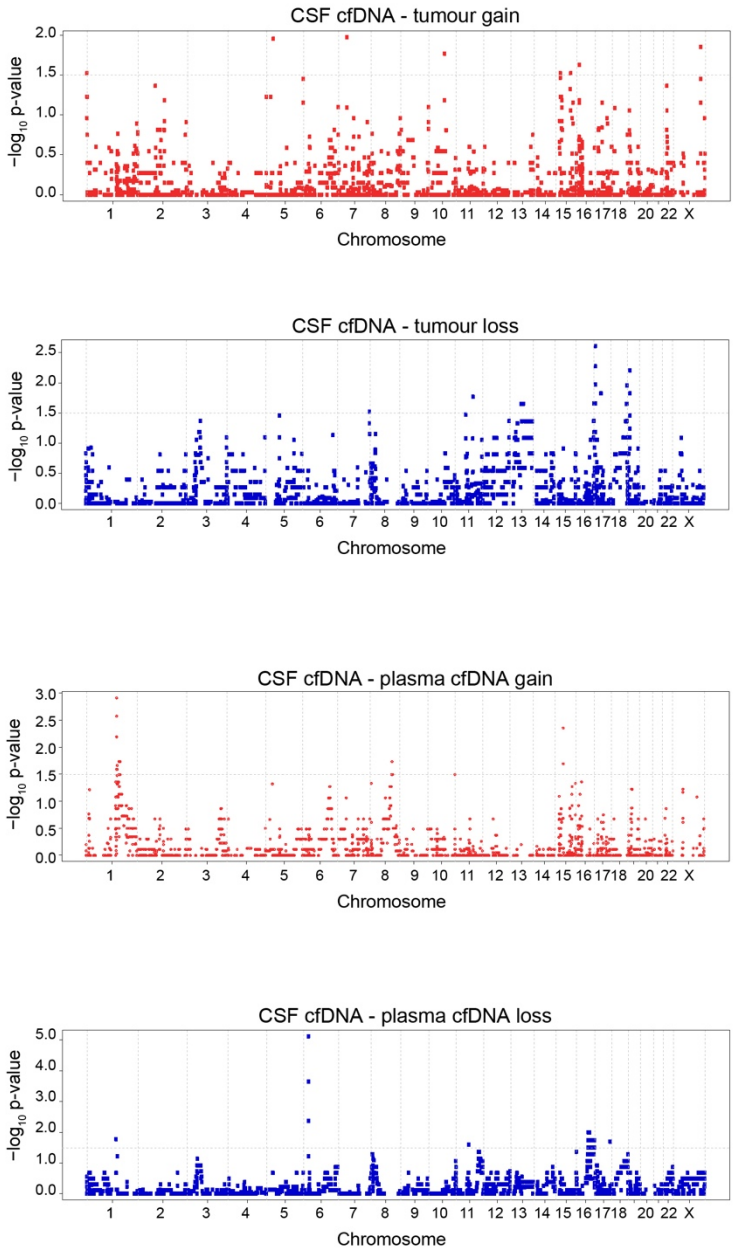
Supplementary Fig. 4. BCLM clonal evolution modelling. Associated with Fig. 2. ClonEvol was used to infer consensus clonal evolution models, using the variant clusters generated by PyClone. Box plots show the variant clusters identified in each sample, and displays the variant allele frequency (VAF) of each mutation, with the median VAF depicted by black lines. Bell plots show the ClonEvol estimated cancer cell fraction (CCF) and evolution order of clones within each sample. Phylogenetic trees for each case depicts the evolution pattern. Each branch represents the evolution path of a clone from its parental clone, and each node marks the clone number of the preceding branch. Clone colour and number match those in box and bell plots. Samples are annotated at clone nodes when no further subclone evolution occurs. Branch lengths are scaled by the square root of number of mutations per clone, and branch width is inversely proportional to branch length. Mutations per clone are shown to top left of the trees. Gene names and variant type are annotated on branches as follows; black font if appear in the list of BCLM frequently altered cancer driver genes (shown in Fig. 3); green font if variant is OncoKB annotated as potentially therapeutically actionable (Supplementary Table 10). Variant type abbreviations; fs, frameshift; nfs, non-frameshift; ns, nonsynonymous; sg, stop-gain; spl, splice-site variant.

Supplementary Figure 5



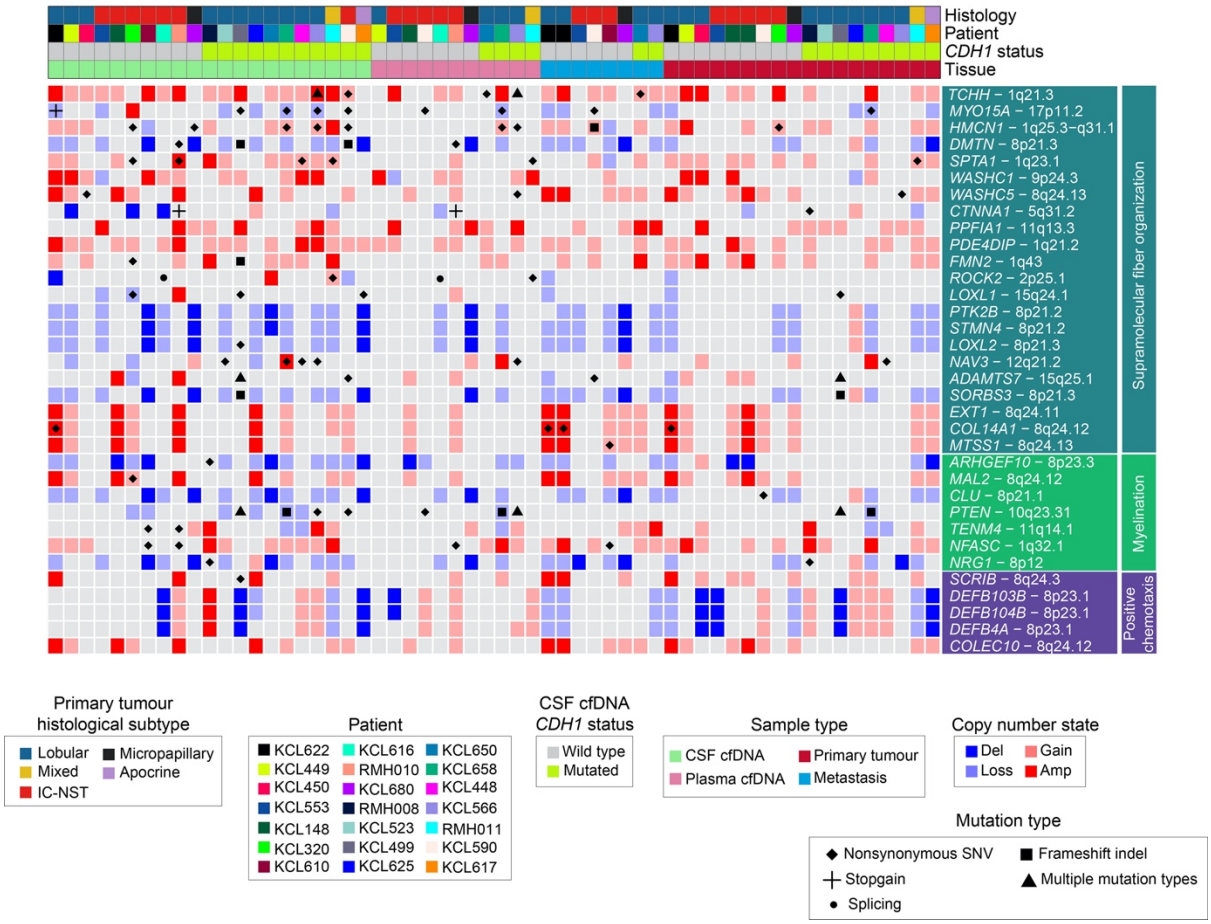
Supplementary Fig. 5. Cancer-associated mutation and copy number landscape of all matched samples. Associated with Fig. 3. Uppermost bar plot show tumour mutational burden (TMB) by sample. Stacked bar chart below shows mutational subtype for each sample as percentage of all variants (excluding synonymous). Banners indicate primary tumour histological subtype and HER2 and ER status. Main plot shows the 36 most frequently cancer-associated altered genes (alterations include mutation and high-level copy number events) in CSF cfDNA as follows; Cancer Gene Census (CGC) gene altered in at least 4 cases; breast cancer driver gene (definition in Methods) altered in at least 3 cases. Gene names are highlighted in red if genomic alteration rate was significantly different ($p \leq 0.05$ by Chi-square test two-sided) in BCLM cohort compared to a publicly available dataset of 216 metastatic breast cancer (MBC cohort) samples (see Methods). Extended data in Supplementary Table 8 displays the significance values for this comparison.

Supplementary Figure 6



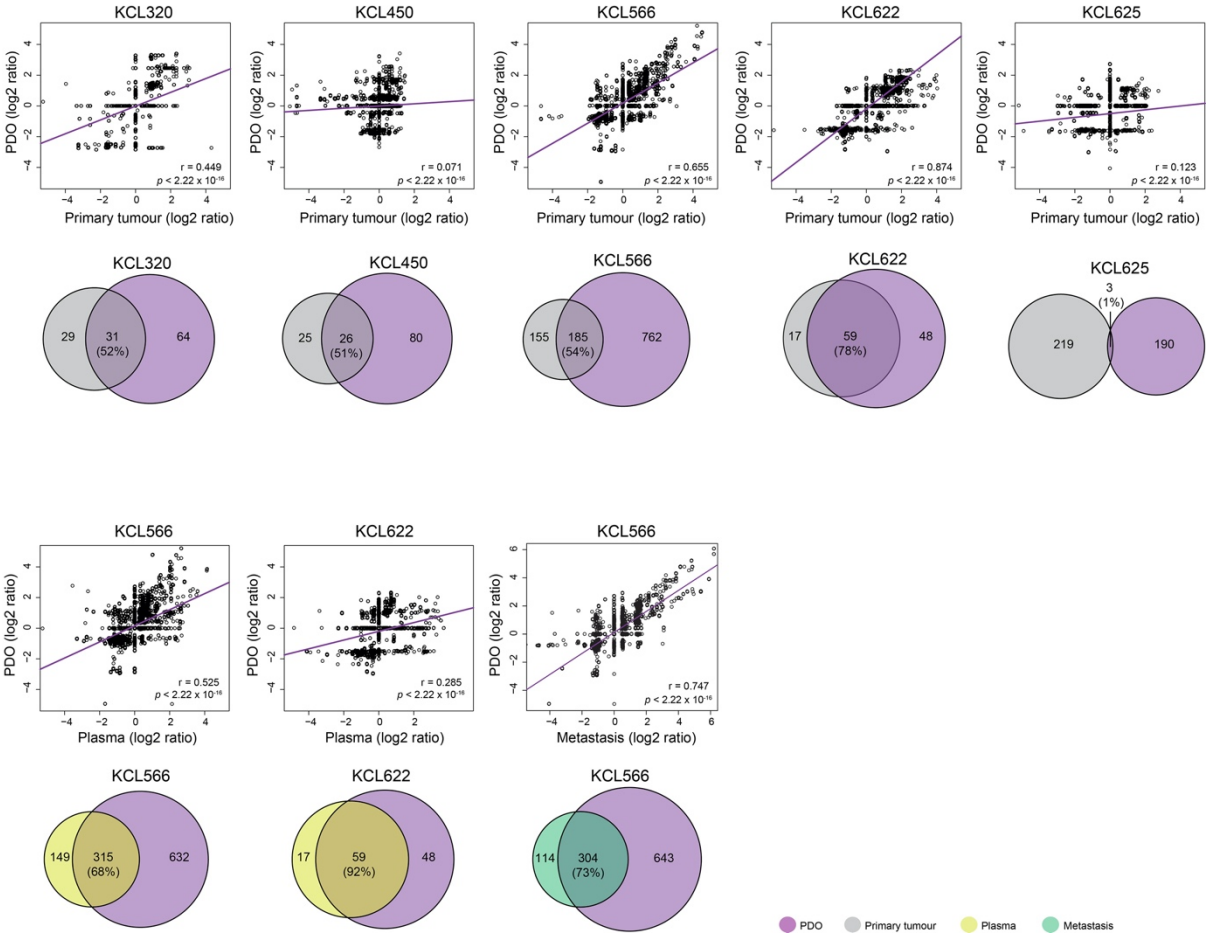
Supplementary Fig. 6. Comparison of CNA frequency between sample type. Associated with Fig. 4a,b. Copy number alteration (CNA) across the genome was determined using WES and adjusted for sample purity and ploidy. Differences in CNA frequency between sample types were analysed with two-sided tests of equal proportions, comparing CSF cfDNA and tumour (n = 17 pairs), and CSF cfDNA and plasma cfDNA (n = 11). Genome-wide plots show the inverse p-value (-log10 p-value) of the significance of CNA frequency difference, where each datapoint represents a gene.

Supplementary Figure 7



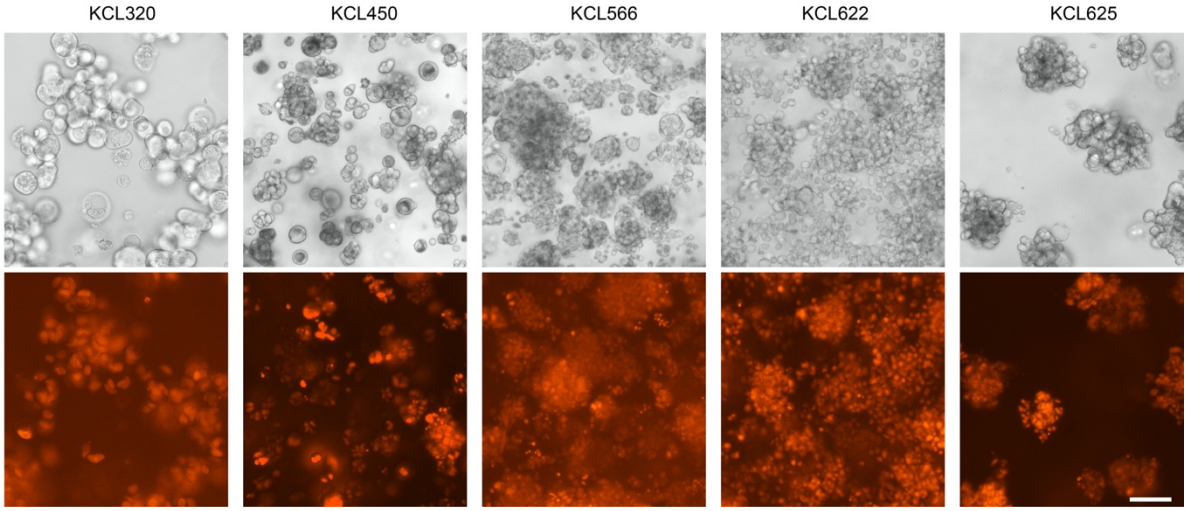
Supplementary Fig. 7. BCLM-enriched genomic alterations in biological processes. Associated with Fig. 4d. Genomic alterations in GO biological processes with -$\log_{10}$ p-value ≥3.0 for enrichment in CSF cfDNA *vs.* primary tumour. Genomic alterations are shown for all sample types in 21 individuals. Banners above indicate; primary tumour histological subtype, patients, *CDH1* mutation status of CSF cfDNA and sample type.

Supplementary Figure 8



Supplementary Fig. 8. Comparison of genomic findings in BCLM PDOs and matched primary tumour, plasma cfDNA or metastatic tissue samples. Associated with Fig. 6a. Top panel, correlation plots of CNA (standardised log2 ratio, one data point per gene) with regression line plotted (purple) and Pearson's r value plus two-sided significance values are shown. Lower panel, shared variants between primary tumour, metastasis or plasma cfDNA and BCLM PDO. Numbers within circles show mutation count, with the percentage showing variants in sample which also detected in BCLM PDO.

Supplementary Figure 9



Supplementary Fig. 9. Transduction of BCLM PDOs to express mCherry and luciferase. PDOs were transduced with a lentivirus encoding luciferase and mCherry (mChLuc2) and imaged on an EVOS microscope during 3D culture in bright field and fluorescence mode. Scale bar, 100 μm.