In-depth analysis of the genomic landscape of 86 metastatic neuroendocrine neoplasms reveals subtype-heterogeneity and potential therapeutic targets

Job van Riet1,2,3, Harmen J. G. van de Werken1,2, Edwin Cuppen4,10, Ferry A. L. M. Eskens5, Margot Tesselaar5, Linde M. van Veenendaal5, Heinz-Josef Klümpen6, Marcus W. Dercksen7, Gerlof D. Valk8, Martijn P. Lolkema15, Stefan Sleijfer3,9, Bianca Mostert5†

1. Cancer Computational Biology Center, Erasmus MC Cancer Institute, University Medical Center, Wytemaweg 80, 3015 CN, Rotterdam, the Netherlands.
2. Department of Urology, Erasmus MC Cancer Institute, University Medical Center, Wytemaweg 80, 3015 CN, Rotterdam, the Netherlands.
3. Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, the Netherlands.
4. Center for Molecular Medicine and Oncode Institute, University Medical Center Utrecht, Heidelberglaan 100, 3584 CX, Utrecht, the Netherlands.
5. Department of Medical Oncology, Cancer Institute, University of Amsterdam, The Netherlands.
6. Department of Medical Oncology, Amsterdam University Medical Centers, Cancer Center Amsterdam, Amsterdam, The Netherlands.
7. Department of Internal Medicine, Maxima Medisch Centrum, Veldhoven, The Netherlands.
8. Department of Endocrine Oncology, University Medical Center Utrecht, Utrecht, The Netherlands.
9. Center for Personalized Cancer Treatment, Rotterdam, the Netherlands.
10. Hartwig Medical Foundation, Amsterdam, the Netherlands.

* These authors contributed equally.
† Corresponding author.

Address for correspondence:
Postbus 2040
3000 CA Rotterdam
b.mostert@erasmusmc.nl

Journal information:
Journal: Nature Communications
Number of text pages: 35
Number of tables: 1 supplementary table.
Number of figures: 5 figures, 10 supplementary figures.

Requirements: main text (excl. abstract, online Methods, refs and fig legends) is 2,000–4,000 words; abstract 100 words; max 8 figures/tables; Introduction (without heading), followed by results (incl subheadings), discussion, online methods (incl subheadings)

Keywords: neuroendocrine; neoplasm; neuroendocrine tumors; whole-genome sequencing; CPCT02; metastasis; therapy targets
Abstract

Metastatic neuroendocrine neoplasms (mNEN) form clinically and genetically heterogeneous malignancies, characterized by distinct prognoses based upon primary tumor localization, functionality, grade, proliferation index and diverse outcomes to treatment. Here, we report the mutational landscape of 86 whole-genome sequenced mNEN. This landscape revealed distinct genomic subpopulations of mNEN based on primary localization and differentiation grade; we observed relatively high tumor mutational burdens (TMB) in neuroendocrine carcinoma (5.45 somatic mutations per megabase) with TP53, KRAS, RB1, MYC and APC as major drivers versus an overall low TMB in neuroendocrine tumors (1.08). Furthermore, we observed distinct drivers which were enriched with somatic aberrations in pancreatic (MEN1, ATRX, DAXX, PCNT and SETD2) and midgut-derived neuroendocrine tumors (CDKN1B). Finally, 49% of mNEN patients revealed extensions of their treatment-repertoire based upon actionable (and responsive) somatic aberrations; potentially directing improvements in mNEN treatment strategies.
Introduction

Neuroendocrine neoplasms (NEN) is a heterogeneous and uncommon tumor type. It can arise from any of the neuroendocrine cells distributed widely throughout the body. Clinically, a distinction is made between the poorly differentiated neuroendocrine carcinomas (NEC) and the better differentiated neuroendocrine tumors (NET), the latter are further subdivided based on their primary site in pancreas (pNET), gastro-intestinal tract or lung NET. Further distinctions are made based upon grade (as assessed by Ki-67 or MIB-1 staining as a measure of proliferation index), differentiation, histology (small-cell vs. large-cell) and functionality (the presence or absence of hormone secretion resulting in typical clinical syndromes dependent upon the predominant hormone that is secreted). Tumor grade and differentiation are associated with prognosis, and all the aforementioned factors affect the choice of treatment. However, also in small subgroups of NEN, such as well-differentiated low-proliferating pNET, marked clinical and genetic heterogeneity occur, as well as, vastly different responses to treatment. Thus, the parameters by which NEN are currently classified do not sufficiently separate patients and tumors according to prognosis and response to therapy. Nonetheless, certain anti-tumor therapies (i.e., sunitinib and everolimus) have been registered for distinct NEN-subtypes. Hence, there is a high unmet need to better classify and understand these diverse tumors, ultimately leading to more tumor- or patient-tailored therapeutic strategies.

Thus far, limited whole genome sequencing data are available for NEN, probably reflecting the rarity of this disease. Currently, pNET have been characterized most extensively; 81 primary tumors were subjected to whole genome sequencing as part of the PCAWG project and another 104 primary pNET were described by Scarpa et al. Additionally, smaller series have been published containing lung and gastrointestinal NET. These studies have shown that NET have a relatively stable genome and only few driver mutations. However, these studies were all performed on primary tumor specimens, whilst a patient generally dies from the consequences of metastatic disease. Additionally, we know from other tumor types that marked heterogeneity can occur between primary and metastatic tumor cells, due to inherent genomic instability and/or the influence of targeted or cytotoxic treatment on the tumor genome. These discrepancies should be taken into account when assessing a patient’s prognosis and possible treatment options, and can be better understood through
thorough genomic characterization of metastases. To date, analysis of metastatic NET is
limited to two studies describing series of five patients with NET originating in the pancreas
and the small intestine (or midgut), respectively.\textsuperscript{10,11} These studies have shown focal
amplification of \textit{MYCN} concomitant with loss of \textit{APC} and \textit{TP53} in one sample as important
metastatic genetic aberrations. For NECs, only two series of whole genome sequencing of
the primary tumors of 1) five cervical and 2) 12 genitourinary NECs have been published.\textsuperscript{12,13}

Whole genome sequencing was performed on 86 NEN metastases (mNEN), one per distinct
patient, which were biopsied as part of the CPCT-02 study.\textsuperscript{14} We report here on the
presence of genomic alterations, mutational and rearrangement signatures for the whole
mNEN cohort and reveal genomic characteristics and alterations distinguishing mNEC from
mNET. Furthermore, we make a genomic distinction between pancreas- and midgut-derived
mNET. Additionally, we investigated the presence of actionable genetic alteration within
mNEN patients, which might render them eligible for off-label or experimental systemic
treatments to extend therapy options.

\textbf{Results}

\textit{Overview of included patients within the CPCT-02 mNEN cohort.}

A total of 109 patients, originally classified as having a neuroendocrine neoplasm, were
included in the CPCT-02 study and had a metastatic biopsy taken in parallel with a blood
control (\textbf{Figure 1}). Five patients were excluded because of missing or withdrawn informed
consent, and another five had non-evaluable biopsies due to low (<20\%) tumor cell
percentage or low DNA yield. Thirteen biopsies were excluded because of incomplete
clinical records, misclassifications of the tumor (based on additional checks of the medical
records), or were duplicate biopsies from the same patient. An overview of the mNEN
patient inclusion per participating Dutch center (\textit{n} = 13) can be found in \textit{Supplementary
Figure 1}.

The metastatic tumor biopsies and corresponding peripheral blood controls from the
remaining 86 distinct patients were whole genome sequenced using paired-end protocols,
to a median mean read coverage of 107x (Q(\text{uartile})\textsubscript{1}-Q\textsubscript{3}: 99x-115x) and 38x (Q\textsubscript{1}-Q\textsubscript{3}: 35-42x),
respectively to a median \textit{in silico} estimated tumor cell purity of 0.7 ($Q_1$-$Q_3$: 0.5-0.81) \textit{(Supplementary figure 2a-b)}.

The mNEN cohort is represented by 37 females and 49 males of a median age of 62 ($Q_1$-$Q_3$: 57-68) and 61 ($Q_1$-$Q_3$: 56-68), at time of biopsy respectively. In total, 70 NET and 16 NEC were included. The primary tumor location in the midgut was most common ($n = 42$, 49%), followed by pancreas ($n = 24$, 28%) and unknown ($n = 12$, 14%) \textit{(Figure 1)}. Most of the tumor biopsies were taken from liver metastases, and a minority from relapses at the primary site \textit{(Supplementary figure 2c-d)}.

The mutational landscape of metastatic neuroendocrine neoplasms reveals differences related to primary localization and degree of differentiation.

The overall mutational landscape of mNEN ($n = 86$; \textit{Figure 2}) reveals two strikingly distinct genomic populations of neuroendocrine neoplasms, i.e. the mNEC and mNET populations. The mNEC ($n = 16$) reveals diploid to triploid genomes and a median tumor mutational burden (TMB) of 5.45 somatic mutations per Mb ($Q_1$-$Q_3$: 3.8 - 8.85), which is in the mid-range of TMB known for human primary cancers\textsuperscript{15}. However, the mNET ($n = 70$) are hallmarked by a relatively stable diploid tumor genome with only few, but specific, chromosomal arm aberrations and harbors the lowest overall TMB of only 1.08 ($Q_1$-$Q_3$: 0.78 - 1.54) of all metastatic cohorts within the CPCT-02 study.\textsuperscript{14}

The somatically-acquired and whole-genomic mutational landscape of mNEC ($n = 16$) revealed a median of 13996 single-nucleotide variants (SNVs; $Q_1$-$Q_3$: 9465 - 22830), 1755 small insertions and deletions (InDels; $Q_1$-$Q_3$: 751 - 2244), 114 multiple-nucleotide variants (MNVs; $Q_1$-$Q_3$: 49 - 198), 150 structural variants (SVs; $Q_1$-$Q_3$: 82 - 264) and an overall diploid to triploid genome ($Q_1$-$Q_3$: 1.9 - 3.1; \textit{Supplementary figure 3}). Concordant with the lower TMB of the mNET ($n = 70$), the mNET revealed a median of 2822 SNVs ($Q_1$-$Q_3$: 1982 - 3892), 254 InDels ($Q_1$-$Q_3$: 185 - 329), 18 MNVs ($Q_1$-$Q_3$: 12 - 27), 17 Structural Variants (SVs; $Q_1$-$Q_3$: 7 - 56) and an overall diploid genome ($Q_1$-$Q_3$: 1.9 - 2.1). The discrepancy in mutational load between mNEC and mNET also held true when inspecting only the coding regions, in which mNEC revealed a higher number of SNVs, InDels, MNV compared to mNET \textit{(Supplementary figure 3)}.
Similarly, mNEC displayed elevated numbers of all SV classes (translocations, deletions, tandem duplications, insertions and inversions; Supplementary figure 3d).

The majority of somatic coding mutations for all mNEC and all mNET (n = 2730 and 3127; SNV, InDel and MNV) were found to be predicted missense variants (59% in mNEC vs. 55% in mNET), followed by synonymous variants (21% vs. 22%). The number of genes harboring somatic mutations within their coding regions differed between mNEC and mNET. Over the entire mNEC cohort (n = 16), 2841 distinct mutant genes were observed, versus 3134 distinct genes within the entire mNET cohort (n = 70). Per sample, a median of 149 (Q1-Q3: 127 - 270) versus 37 (median; Q1-Q3: 26 - 50) genes harboring mutations within coding regions were observed for mNEC and mNET samples respectively; revealing that mNEC harbor greater numbers of mutant genes compared to mNET.

The median genome-wide ratio of transitions (Ti; A↔G or T↔C) to transversions (Tv; C↔A, C↔G, T↔A or T↔G) within mNEC was found to be 0.78 Ti\Tv (Q1-Q3: 0.72 - 1.02) vs. 1.52 Ti\Tv (Q1-Q3: 1.12 – 2.20) in the coding regions. For mNET the median genome-wide and coding Ti\Tv were found to be 1.11 (Q1-Q3: 0.99 - 1.32) and 1.42 (Q1-Q3: 1 - 1.96), respectively (Supplementary figure 3f).

High TMB (≥10) are often associated with DNA repair deficiency and/or tumors with sensitivity for immune therapy, e.g. checkpoint inhibitors. Four mNEC samples, all from unknown origin, and a single pancreatic mNET showed this high-TMB genotype (Figure 2a). One mNET displayed signs of BRCA2-associated homologous recombination deficiency (HRD), as determined using the CHORD classifier which is mainly based on deletions with flanking microhomology and 1-100kb structural duplications (Figure 2j; supplementary figure 4). Further inspection revealed that this mNET harbored a somatic frameshift mutation within RAD51C, a known HRD-associated gene.16

Regional hypermutation (kataegis).

Regional hypermutation (kataegis) was detected in six mNEC; Figure 2i; supplementary Figure 5). Canonically, kataegis is associated with APOBEC-activity and indeed, four out of five (80%) of these kataegis events predominantly showed the canonical TpCpW context
associated with APOBEC alterations. Additionally, in the samples with kataegis \( n = 5 \), the absolute contribution of APOBEC SBS signatures (2 & 13) was significantly higher (median 45 vs. 533, \( p < 0.01 \), Wilcoxon rank-sum test) compared to mNEN without kataegis \( n = 81 \).

**Chromothripsis**

Multiple distinct mNEN (four mNEC and two mNET; 7%) revealed presence of chromothripsis, a catastrophic phenomenon of the shattering and interchromosomal recombination of one or more chromosomes (Figure 2h; Supplementary figure 6). Strikingly, four of the six observed chromothripsis events from distinct mNEN (two mNEC and two mNET) involved the same chromosome, namely chromosome 12. Within these four mNEN, we observed possible evidence for extrachromosomal DNA due to copy-number oscillations between one low (CN≤4) and one very high (CN≥10) states, consistent with the presence of double minutes. Several known oncogenes lie within these regions on chromosome 12, including MDM2 and possibly CCND2, and could be related to a phenomena which generates extrachromosomal DNA which has been described previously in primary tumors within the PCAWG cohort.

**Catalog of the cohort-wide mutational signatures provide biological insights into treatment effect.**

Different mutational processes, such as exposure to exogenous or endogenous mutagens and defective DNA repair mechanisms generate unique combinations of mutational trinucleotide contexts which are reflected in mutational signatures. To determine these mutational signatures within mNEN, we performed *de novo* mutational signature analysis and determined the contribution of previously described single base substitution (SBS) mutational signatures (COSMIC v3). The *de novo* mutational signature assessment revealed seven signatures, denoted as Sig. A to Sig. G, (supplementary figure 7d) which all strongly correlated to previously known mutational signatures (supplementary figure 7a-c). In particular, we observed samples with large relative contributions (>20%) of *de novo* signatures similar to the known signatures associated with aging (SBS1 & 5; Sig A and D), APOBEC activity (SBS2 & 13; Sig B.), tobacco smoking (SBS4; Sig F.), alkylating agents exposure (SBS11; Sig E.), 5-Fluorouracil exposure (SBS17a-b; Sig. C.) and MUTYH mutations (SBS36; Sig. G.).
Overall, the mutational signature profiles do not differ greatly within the mNEN cohort. SBS5 ($n = 48$; putative clock-like), SBS8 ($n = 45$; Unknown), SBS40 ($n = 22$; Unknown), SBS3 ($n = 16$; HRD-like), SBS1 ($n = 10$; clock-like), SBS39 ($n = 7$; Unknown) and SBS9 ($n = 5$; polymerase η ($POLH$) activity) were classified as dominant signatures (i.e., contributed at least 10% of total contribution within ≥5 mNEN; Figure 2e).

Two included mNEC of unknown primary localization are characterized by high-TMB (≥10) and SBS4, which is associated with smoking; likely due to tobacco mutagens. This could reflect that these metastases could be primary lung NEC, as a relation between mNEC on non-pulmonary origin and smoking is not known. However, as no somatic coding mutations in canonical lung cancer-associated genes were observed and the clinicopathological data of these patients did not point to any different primary tumor other than a NEC, it seems unlikely that these could be primary non-small cell lung cancers.

Strikingly, the only high-TMB (pancreatic) NET was strongly characterized by SBS11 which exhibits a mutational pattern resembling that of alkylating agents, with a strong enrichment for C/T (G>A) transitions. Previously, an association between treatment with the alkylating agent temozolomide and SBS11 mutations has been found. This same patient showed the highest tumor mutational burden with a TMB of 21.4 (median TMB of NET: 1.1) and was treated with a combination of 5-fluorouracil and streptozocin before undergoing a biopsy for the CPCT-02 study. Streptozocin is a capable of DNA alkylation and inhibition of DNA synthesis, and its mechanism of action closely resembles that of temozolomide.

One mNET was strongly characterized by SBS36, associated with base excision repair (BER) deficiency due to $MUTYH$ alterations, C>A mutations and previously also seen in pancreatic NET. Strikingly, this tumor did not harbor specific somatic alterations within $MUTYH$ but possessed a heterozygous germline pathogenic missense mutation within $MUTYH$ (c.527A>G / p.Tyr176Cys; rs34612342) coupled with a complete loss of a single chromosome 1, resulting in subsequent loss-of-heterozygosity.
Driver catalog of mNEN

Utilizing unbiased driver gene discovery by GISTIC2 and dN/dS on the entire mNEN cohort and separately on all mNET and mNEC samples, we detected seventeen focal deletion peaks and two focal copy-number amplifications peaks throughout the genome ($q \leq 0.1$) and nine genes enriched with non-synonymous mutations ($q \leq 0.1$; figure 3 and supplementary figure 8). Within these focal peaks, several oncogenes and tumor suppressors were present which could be the potential target of the copy-number alteration. These genes, which have been previously associated as driver genes in NET and/or pan-cancer cohorts$^{3,5,14}$, are shown in figure 3 for all mNEN with a distinction between mNEC and mNET. We detected several previously known tumor suppressors and oncogenes such as TP53, KRAS, MEN1, RB1, CDKN1B, DAXX and APC enriched with non-synonymous mutations ($q \leq 0.05$) as well two additional genes (SETD2 and CREBBP) just above the statistical threshold value ($q \leq 0.1$). By overlapping known drivers within the observed focal amplification and deletion peaks, we detect a plethora of putative drivers with copy-number alteration; such as deletions of TP53, CDKN2A, CDKN2B, CDKN1B, PTPRD, CBFAT3, PLCG2, ANKDR11, IRF8, LINCO1237, PRKN, ZNF407, common fragile sites such as DMD, FHIT and MACROD2, and amplifications of genes such as PCAT1/MYC and MDM2. Furthermore, focal deletions of additional genes such as CAMTA1, DLUE1/2, TRIM13, KCNNR1, FXD1 were found in $\leq 2$ samples (supplementary table 1). Large perturbations on chromosome 12q15 (MDM2) were observed within mNEN harboring chromothripsis (supplementary figure 6). Furthermore, we could detect a single in-frame fusion of the common fusion-partner EWSR1 seen in pNET$^3$. Moreover, we observed only two genes harboring hotspot coding mutations (on base-level) which were shared between three samples (ZNF829 and KRAS) and seven genes between two samples (UHRF1BP1L, CDKN1B, MEN1, LEKR1, OR5L1, CTNNB1 and GNAS; supplementary table 1).

We observed an overall heterogeneous pattern of putative drivers, the most frequently putative driver was found to be CDKN2A/B ($n = 18; 15$), followed by TP53 ($n = 17$), CDKN1B ($n = 12$), PTPRD ($n = 11$), KRAS ($n = 11$), MEN1 ($n = 11$) and RB1 ($n = 11$). Strikingly, a significant portion of the total mNEN cohort had no mutual putative driver(s) (9 out of 86; 10%) and only contained patient-specific putative drivers.
We next investigated whether any form of mutational enrichment, such as somatic alterations within certain genes (mutations and/or copy-number alterations) or evidence of large-scale events (kataegis and chromothripsis), could be related to one of our three major subgroups relating to subtype or primary localization; being mNEC \((n = 16)\), pancreas- \((n = 20)\) and midgut-derived mNET \((n = 40)\). Using a one-sided Fisher’s Exact Test (with Benjamini-Hochberg correction), we detected the enrichment of at least one such event(s) within these subgroups (supplementary figure 9e). Within mNEC, an enrichment of alterations within \(TP53\) (88% of mNEC), \(KRAS\) (50%), \(RB1\) (50%), \(MYC\) (31%), \(APC\) (31%), \(ZFHX4\) (31%), \(UBR5\) (25%) and presence of kataegis (31%) could be appreciated \((q \leq 0.05)\). Likewise, within pancreas-derived mNET, an enrichment of was seen for \(MEN1\) (40% of pancreas-derived mNET), \(ATRX\) (25%), \(DAXX\) (25%), \(SETD2\) (25%) and \(PCNT\) (20%) whilst midgut-derived mNET revealed enrichment of \(CDKN1B\) alterations (25% of midgut-derived mNET).

**Genomic differences relating to primary localization of mNET**

Due to distinct prognosis and previous genetic associations, we investigated genome-wide differences in regards to primary localization within the mNET population \((n = 70)\). We observed several genome-wide differences relating to primary localization (Figure 2, Supplementary figure 9), such as the median genome-wide TMB; ranging from 1.05 (mNET - Midgut; \(Q_1-Q_3\): 0.75 - 1.39) and 1.07 (mNET - Unknown; \(Q_1-Q_3\): 0.84 - 1.53) to 1.27 (mNET - Other; \(Q_1-Q_3\): 1.10 - 1.44) and 1.35 (mNET - Pancreas; \(Q_1-Q_3\): 0.9 - 2.12). A similar pattern was detected regarding the number of distinct genes with coding mutations. Midgut-derived mNET also presented a surprisingly low number of structural variants compared to the other mNET sub-populations.

Next, we investigated possible differences in putative drivers between our major mNET sub-populations, being midgut- \((n = 40)\) and pancreas-derived \((n = 20)\) mNET (Figure 4, supplementary figure 9). The copy-number profiles (GISTIC2) of both populations differed, in which midgut-derived mNET presented focal deletion peaks at 9p21 \((CDKN2A/B)\), 11q23 (131 genes), 12p13 \((CDKN1B)\), 13q14 (10 driver genes), 14q24 (17 genes) and 16q23 (234 genes; common fragile site) coupled with an overall flat diploid profile. Pancreas-derived
mNET presented a different profile harboring focal deletion peaks at 2q37 (LINC01237, LINC01880, LINC01238), 9p21 (CDKN2A/B) and Xp21 (DMD; common fragile site gene) couples with a more instable genomic profile, including several samples with large-scale chromosomal losses (supplementary figure 8, supplementary figure 9c). When investigating the statistically significant large-scale copy-number alterations of the chromosomal arms, we also detect striking differences between the major subgroups (supplementary figure 10). Within mNEC, we detected a large number of samples (69%) harboring a loss of 22q. Midgut-derived mNET revealed amplifications of chromosome 4p/q, 5p/q, 7p/q, 10p/q, 14p/q, 20p/q and loss of 9p/q in various samples (~30%) and a loss of 18p/q in 66% of samples. This re-confirms the high frequency of chromosome 18 loss in midgut-derived NET and the association with DDC26, as DCC is the most recurrently mutated gene on chromosome 18 in our cohort also (n = 5). Finally, over half of pancreas-derived mNET revealed amplifications of chromosome 5p/q, 7p/q, 9p, 12p/q, 13q, 14p/q, 17p/q, 18p/q, 19p/q, 20p/q and loss of 22q.

Unbiased driver gene analysis (dN/dS) on midgut-derived mNET presented CDKN1B whilst pancreas-derived mNET revealed MEN1, DAXX and SETD2. Several genes (present in ≥3 samples) were found only, or predominately, within midgut-derived mNET: CDKN1B, CDKN2A / CDKN2B, BIRC7, GNAS, PSIP1, KMT2A and PTPRD (figure 4). Conversely, MEN1, DAXX, SETD2, CREBBP, PCNT, KDR and TSC2 were found to be mutated only within pancreas-derived mNET. Moreover, MEN1 was found to be mutationally enriched within pancreas-derived mNET when compared to the entire mNEN cohort (q ≤ 0.05). Several midgut-derived mNET (n = 9; 23%) did not readily present a shared mutual driver and only harbored somatic mutations in private or as-of-yet unassociated cancer driver genes.

Clinically-actionable mutations.

We observed forty-two mNEN (49%) harboring one or more target-specific or general somatic aberrations which are known as possible (and responsive) druggable targets against currently-available (or under development) treatment agents are available. Twenty-one mNEN (24%) harbored somatic aberrations corresponding to a treatment that is currently registered for NEN or specifically for the NEN-subtype of that particular patient (figure 5, supplementary table 1). In addition, twelve patients (14%) could benefit from therapies that
are off-label, but are commonly considered best practice for NEN. Another eight patients (9%) could benefit from drugs which are registered for another indication but not currently administered in NEN treatment. Additionally, six tumors harbored an aberration rendering them sensitive to a drug that is still in development; including a single patient with no actionable alterations otherwise. We found RB1 (n = 11), KRAS (n = 11), MTAP (n = 5), high-TMB (≥10; n = 5), RICTOR (n = 4) and TP53 (n = 4) to be the most frequently observed (target-specific or general) somatic aberrations which granted eligibility to various possible treatment options. In total, ten midgut-derived mNET (25%) and eleven pancreas-derived mNET (55%) revealed potentially responsive alterations in various genes and most strikingly, almost all mNEN (94%) revealed potential responsive targets due to RB1 and/or KRAS mutations or towards checkpoint inhibitors due to high tumor mutational burden (≥10).

**Discussion**

Historically, NEN has long been known as a difficult malignancy to diagnose, monitor and treat due to presentation of an inherently wide spectrum of disease progression, cellular differentiation and low mutational burden, resulting in few targetable mutations and a relatively stable tumor genome. Indeed, mNET is characterized by the lowest TMB of all metastatic cohorts sequenced within the CPCT-02 study. This study is the first to have an in-depth look into the whole genome and mutations of a large cohort of 86 metastatic NEN from various primary localizations and differentiation grade. The relatively large number of unknown primary tumor localizations in this mNEN cohort (n = 12; 14%) reflects the difficulties in daily clinical practice to determine the site of origin for mNEN.

In our mNEN cohort, it is apparent that the molecular landscape of mNEC is markedly dissimilar from the more differentiated mNET, in terms of mutational burden (median TMB of 5.45 vs. 1.08, respectively), genomic stability, and distinct mutant (driver) genes. With respect to TMB, four mNEC and a single mNET presented a high-TMB genotype (TMB ≥10) which could render these patients eligible for immune-based therapies such as checkpoint inhibitors.27,28

The single high-TMB pancreas-derived mNET presented a striking contribution of the mutational signature associated with alkylating agents (temozolomide) and was previously
treated with a combination of 5-fluorouracil and the alkylating antineoplastic agent streptozocin. The mechanism of action for streptozocin closely resembles that of temozolomide as both react with DNA by undergoing substitution reactions forming a methylidiazonium ion, resulting in methylation of primarily $N^7$ guanine (67%). They both induce high levels of DNA methylation, and recognition and repair of this methylation results in single- and double-strand DNA breaks. To the best of our knowledge, no data have been published on a correlation between hypermutation and streptozocin treatment, but as streptozocin and temozolomide so closely resemble each other in their mechanism of action, one can hypothesize the same mechanism to occur in streptozocin-treated patients. It would be interesting to investigate whether prior treatment with streptozocin or temozolomide indeed induces high TMB in mNEN, and if so, whether pretreatment with streptozocin or temozolomide renders these tumors more sensitive to checkpoint inhibition.

Similarly, we observed a large contribution of the mutational signature associated with base excision repair deficiency due to MUTYH aberrations in the second highest-TMB mNET, and indeed this patient harbored a pathogenic germline MUTYH allele coupled with a complete somatic loss of the respective chromosomal arm. MUTYH abnormalities have also previously described to occur in pancreatic NET. A single mNET presented a BRCA2-genotype associated with homologous recombination deficiency but did not harbor (somatic) mutations within BRCA2. It did harbor a somatic mutation in RAD51C, a gene known to be involved with homologous recombination and repair of DNA.

Concerning genomic stability, we observed evidence of chromothripsis, a large-scale and catastrophic chromosomal rearrangement, within six mNEN (four mNEC, two mNET). Strikingly, four out of six chromothripsis events occurred on chromosome 12. In addition, we observe the first occurrence of localized hypermutation (kataegis) in six mNEC. Kataegis encompasses a pattern of localized hypermutations, which has been identified in various, but not all and to a varying degree, cancer types. These regions of kataegis often co-localize with regions of genetic rearrangements. Kataegis is thought to arise from frequent genomic C-to-U deamination events as a result of APOBEC-family enzyme activity, a DNA cytosine deaminase which was recently identified as an internal and thus far unrecognized source of DNA damage and mutagenesis in various cancer types. More recently, kataegis, rather than tumor mutational burden, microsatellite instability or mismatch repair
deficiency, was found to independently correlate with PD-L1/PD-L2 expression, and could thus be a marker in response to immune checkpoint inhibition. Using unbiased driver gene analysis (dN/dS and GISTIC2) on the mNEN cohort, and on mNEC/mNEC separately to explore putative driver genes, we (re-)discovered 9 genes to be enriched with non-synonymous mutations (TP53, CDKN1B, KRAS, MEN1, RB1, CREBBP, APC, DAXX and SETD2) and detected 17 focal deletion and 2 focal amplification peaks overlapping with a plethora of (driver) genes, including deletions of TP53, CDKN2A, CDKN2B, CDKN1B, PTPRD, CBFA2T3, PLCG2, ANKDR11, IRF8, LINC01237, PRKN, ZNF407 and fragile site genes FHIT, DMD and MACROD2, and amplifications of PCAT1/MYC and MDM2. Investigation of mutational enrichment within our major subgroups revealed that somatic alterations in TP53, KRAS, RB1, MYC, APC, ZFHX4 and UBR5, as well as presence of kataegis was enriched within mNEC. Within pancreas-derived mNET, we report the enriched presence of mutant MEN1, ATRX, DAXX, SETD2 and PCNT whilst midgut-derived mNET showed preference for CDKN1B alterations.

Of the genetic alterations named above, a number have been previously associated with midgut NET (CDKN1B)\textsuperscript{34,35}, lung NET (FHIT)\textsuperscript{36–38}, pancreas NET (TP53, MEN1, DAXX and SETD2)\textsuperscript{39} and NEC (TP53 and RB1)\textsuperscript{40,41}. Other genes are associated with various other malignancies (PTPRD, CBFA2T3, ANKRD11, IRF8, ZFHX4\textsuperscript{46}, and MDM2\textsuperscript{47}) or genomic instability (DMD and PRKN\textsuperscript{37, MACROD2}48). Deletion of PGLC2 has been linked to irregularities within the immune system.\textsuperscript{49} DAXX is a tumor suppressor gene, its protein functioning as a dimer together with ATRX involved in chromatin stabilization.\textsuperscript{50} Loss of function mutations in DAXX have been associated with an increased risk of liver metastases\textsuperscript{51} and with shortened overall survival\textsuperscript{25}. Conversely, in a series of 58 pNET, of which 17 had a somatic DAXX mutation, the presence of a DAXX mutation was associated with a favorable prognosis, as was the presence of a MEN1 mutation\textsuperscript{39}. In addition, mutations in TP53, APC, KRAS and RB1 has previously been associated with pancreatic or colorectal NEN and small cell lung carcinoma (SCLC)\textsuperscript{52–54}, and amplification of MYC has been found in SCLC with neuroendocrine differentiation\textsuperscript{55}. 
Currently, the choice of treatment in an individual mNEN patient is, besides factors such as comorbidity and patient preference, determined by primary tumor localization, proliferation index (as determined by Ki-67 or MIB-1 staining) and somatostatin expression. The distinction based on primary tumor localization stems from the different embryologic structures the tumor can originate from (foregut, midgut or hindgut). When we compared the various origins at a genomic level, we could observe an increasing TMB; ranging from 1.05 (mNET - Midgut; Q1-Q3: 0.75 - 1.39) and 1.07 (mNET - Unknown; Q1-Q3: 0.84 - 1.53) to 1.27 (mNET - Other; Q1-Q3: 1.10 - 1.44) and 1.35 (mNET - Pancreas; Q1-Q3: 0.9 - 2.12) to 5.45 (mNEC; Q1-Q3: 3.8 - 8.85). In addition, when we compared the two largest groups of mNET per primary localization (midgut and pancreas), we can readily distinguish between the two subtypes based on somatic mutation and copy-number profiles. Yet strikingly, many midgut-derived mNET (n = 9; 23%) did not present a mutual driver gene but each was characterized by distinct sets of mutated genes reflecting the heterogenous nature of the malignancy.

Almost half of mNEN (n = 42; 49%) harbored a specific genomic alteration or genotype for which an FDA-approved drug is currently available, either on (registered for that indication) or off-label. Thus, whole genome sequencing of metastatic lesions revealed 49% of mNEN patients harboring clinically-relevant targetable somatic aberrations which could possibly extend their treatment-repertoire. Additionally, six patients’ tumors (also) harbored a specific genomic alteration which could render them sensitive to drugs that are currently under development in clinical trials. These drugs are currently not readily available for these patients, but could provide new treatment options in the future. When deciding upon a new line of systemic treatment, a metastatic biopsy could always be considered, preferably in the context of a study, as this could shed light upon additional and effective treatment options for these late-stage patients with otherwise few remaining treatment options. In the Netherlands, we have the DRUP study active, a study in which patients for whom no standard treatments are currently available and whom might be treated with anticancer treatments outside of their approved label based on the presence of actionable mutations in their tumors.56

In this current study exploring the largest whole-genome sequenced mNEN repository to date (n = 86), we focused on the genes that drive mNEN and analyzed additional aspects of...
genomic instability within mNEN, such as structural variants, kataegis, chromothripsis and homologous recombination deficiency. With increased knowledge and new algorithms on analyzing and interpreting large-scale WGS cohorts, our understanding of the complex molecular makeup of (m)NEN has been greatly increased and could be exploited for better distinction of tumor subgroups and new treatment options.
Online methods

Patient cohort and study procedures
Patients with mNEN were recruited under the study protocol (NCT01855477) of the Center for Personalized Cancer Treatment (CPCT) within the CPCT-02 study. This CPCT-02 protocol was approved by the medical ethical committee (METC) of the University Medical Center Utrecht. Patients were eligible for inclusion if the following criteria were met: 1) age ≥ 18 years; 2) locally advanced or metastatic solid tumor; 3) indication for new line of systemic treatment with registered anti-cancer agents; 4) safe biopsy according to the intervening physician. All patients provided written informed consent before any study procedure. The study procedures consisted of the collection of matched peripheral blood samples for reference DNA and image-guided percutaneous biopsy of the metastatic lesion. For the current study, patients were included for biopsy between May 10th 2016 and July 17th 2018 resulting in a cohort of 86 distinct patients from 13 Dutch hospitals (Supplementary table 1).

Collection, sequencing and processing of mNEN biopsies.
Blood samples were collected in CellSave preservative tubes (Menarini-Silicon Biosystems, Huntington Valley, PA, USA) and shipped by room temperature to the central sequencing facility at the Hartwig Medical Foundation. Tumor samples were fresh-frozen in liquid nitrogen directly after the procedure and send to a central pathology tissue facility. Tumor cellularity was estimated by assessing a hematoxylin-eosin (HE) stained 6 micron section. Subsequently, 25 sections of 20 micron were collected for DNA isolation. DNA was isolated with an automated workflow (QiaSymphony) using the DSP DNA Midi kit for blood and QIAsymphony DSP DNA Mini kit for tumor samples according to the manufacturer’s protocol (Qiagen). DNA concentration was measured by Qubit™ fluorometric quantitation (Invitrogen, Life Technologies, Carlsbad, CA, USA). DNA libraries for Illumina sequencing were generated from 50-100 ng of genomic DNA using standard protocols (Illumina, San Diego, CA, USA) and subsequently whole-genome sequenced in a HiSeq X Ten system using the paired-end sequencing protocol (2x150bp) for both the metastatic tumor and matched blood sample.
Subsequent alignment, somatic mutation detection and \textit{in silico} tumor cell percentage estimation were performed in a uniform manner as detailed by Priestley et al. (2019). Briefly, paired-end sequencing reads were aligned against the human reference genome (GRCh37) using BWA-mem (v0.7.5a). Duplicate reads were marked and small insertion and deletions (InDels) were realigned using GATK IndelRealigner (v3.4.46). Prior to somatic SNV and InDel variant calling, base qualities were recalibrated using GATK BQSR (v3.4.46). Somatic SNV, InDels and MNV were called by Strelka (v1.0.14) using the matched peripheral blood WGS sample for matched-normal variant calling.

Additional in-depth settings and optimizations of the HMF pipeline are described by Priestley et al. (2019) and tools are available at https://github.com/hartwigmedical/.

The somatic mutations (SNV, InDels and MNV) were further annotated with Ensembl Variant Effect Predictor\textsuperscript{50} (VEP, version 99, cache 99\_GRCh37) using GENCODE (v33) annotations in tandem with the dbNSFP\textsuperscript{61} plugin (version 3.5, hg19) for gnomAD\textsuperscript{62} population frequencies. SIFT\textsuperscript{63} and PolyPhen-2\textsuperscript{64} scoring was applied for additional functional effect prediction.

During downstream analysis, we only retained SNV, InDels and MNV which passed all of the following heuristic filters; default Strelka filters (PASS-only), gnomAD exome (ALL) allele frequency < 0.001, gnomAD genome (ALL) < 0.005, not present in ≥ 5 samples from the Hartwig Medical Foundation germline panel-of-normals (GATK Haplotyper) and not present in ≥ 3 samples from the Hartwig Medical Foundation Strelka-specific somatic blacklist.

Putative protein-altering (coding) or high-impact (e.g. splicing) mutations were aggregated per sample and gene by selecting the most deleterious annotated effect (from VEP) on any known overlapping gene-wise transcript (except those transcripts flagged as retained intron and nonsense mediated decay). In addition, structural variants with a Tumor Allele Frequency (TAF) ≥ 0.1 that overlapped only partly with the respective coding sequences (i.e. not all exons of the respective gene), were annotated as ‘Structural Variant’ mutations. Multiple coding mutations and/or SV per gene were annotated as 'multiple mutations'.
Discovery of somatic structural variants (SV), copy-number alterations and in-frame fusions of EWSR1 was performed using the GRIDDS, PURPLE and LINX suite. During the downstream analyses, we only retained somatic structural variants passing all default QC filters (PASS-only) and with an upstream and/or downstream TAF ≥ 0.1.

Mean read coverages of the reference and tumor samples were calculated using Picard Tools (v1.141; CollectWgsMetrics) based on GRCh37. Genomic and coding tumor mutational burden (TMB) was calculated as previously described by van Dessel/van Riet et al. (2019).

Determining purity-corrected allele frequencies for somatic alterations.

To calculate the Tumor Allele Frequencies (TAF) of somatic SNV, InDels, MNV and SV, representing the tumor purity-corrected variant frequencies, we followed a previously described approach by Stephens et al. (2012), implemented as:

\[
TAF_m = \frac{f_m}{p} [pC_t + (1 - p)C_h] \quad \text{(equation 1)}
\]

in which where \(f_m\) is the ratio of primary-aligned and non-duplicated reads observed for alternative allele \(m\) over the reference allele (VAF), \(p\) is the in silico estimated tumor purity fraction, \(C_t\) is the absolute copy-number of the segment overlapping \(m\) and \(C_h\) is the wild-type (healthy) copy number; \(C_h = 2\) for autosomes and allosomes in female samples and \(C_h = 1\) for allosomes in male samples.

Discovery of genes under evolutionary selection

We performed a dN/dS analysis on somatic mutations (SNV and InDels) using dndscv (v0.0.1.0) on respective genome sequences and transcript annotations using a custom transcript database based on ENSEMBL Genes (v99)/GENCODE (v33) annotations. We performed a dN/dS analysis over the entire NEN cohort (\(n = 86\)) and four separate dN/dS analysis on the major subgroups (mNEC; \(n = 16\), NET; \(n = 70\), mNET-midgut; \(n = 40\) and mNET-pancreas; \(n = 20\)). Genes-of-interest were selected based on the statistical significance, corrected for multiple hypothesis testing (Benjamini-Hochberg), which integrated all mutation types (missense, nonsense, essential splice-site mutations and InDels; \(q_{global.cv} \leq 0.1\)) and/or without InDels (\(q_{allsubs.cv} \leq 0.1\)).
Detection and annotation of recurrent copy-number alterations

To detect recurrent copy-number alterations, we performed a GISTIC2\(^{71}\) (v2.0.23) analysis over the entire mNEN cohort and, again, four separate GISTIC2 analysis on the major subgroups (mNEC, mNET and pancreas- and midgut-derived mNET).

The GISTIC2 was performed using the following settings:

Genes were annotated to GISTIC2 peaks (q ≤ 0.1) based on the following strategy:

1) GISTIC2 focal peaks (all_lesions.conf_95.txt) were overlapped to genes (from verified and manually annotated loci, no pseudogenes or read-throughs and from standard chromosomes; \(n = 36574\)) from GENCODE (GRCh37; v33), taking into consideration only the genes overlapping with at least 100 base pairs within the detected GISTIC2 peak.

2) If a GISTIC2 focal peak overlapped with multiple GENCODE genes, a combined database containing known drivers detected in a metastatic pan-cancer dataset (CPCT-02)\(^{14}\), COSMIC Cancer Gene Census (v85)\(^{72}\), OncoKB Cancer Gene Census (June 2019)\(^{73}\), Martincorena et al. (2017)\(^{69}\) and Priestley et al. (2019)\(^{14}\) were used to further pinpoint the possible target gene(s) (\(n = 1272\)), e.g. if a GISTIC2 peak overlapped both PTEN and near-adjacent non-driver gene, only PTEN would be chosen as possible target. The list of all overlapping GENCODE (v33) genes per GISTIC2 peak can be found in supplementary table 1.

3) If no overlapping genes were found, GISTIC2 peaks were annotated with the nearest GENCODE (v33) protein-coding gene (\(n = 19988\)).

Mutational signature analysis

Mutational signatures based on the trinucleotide contexts of SNVs was performed, mainly using the MutationalPatterns package (1.10.0)\(^{75}\) and as previously described.\(^{67}\) The 96 Single Base Substitution (SBS) mutational signatures (COSMIC v3) as established by Alexandrov et al. (2019)\(^{23}\), (matrix \(S_{ij}\); \(i = 96\); number of trinucleotide motifs; \(j = \) number of signatures) were downloaded from COSMIC (as deposited on May 2019). The proposed etiology of each SBS signature was derived from Alexandrov et al. (2019)\(^{23}\), Petljak et al. (2019)\(^{23}\), Angus et al. (2019)\(^{6}\) and Christensen et al. (2019)\(^{76}\).
In addition, *de novo* mutational signature analysis by MutationalPatterns was performed based on the max. number of relevant signatures as assessed using the NMF R package\(^77\) (v0.21.0) with 1000 iterations (supplementary figure 7d). By comparing the cophenetic correlation coefficient, residual sum of squares and silhouette, we opted to generate seven custom *de novo* signatures. Custom signatures were correlated to existing (COSMIC v3) mutational signatures using cosine similarity.

**Detection of chromothripsis**

Shatterseek\(^18\) (v0.4) using default parameters was used to detect chromothripsis-like events. As input, we used the rounded absolute copy numbers (as derived by PURPLE) and structural variants with an TAF ≥ 0.1 at either end of the breakpoint. The male sex chromosome (chrY) was excluded. The criteria for a chromothripsis-like event were based on the following criteria: a) total number of intra-chromosomal structural variants involved in the event ≥25; b) max. number of oscillating CN segments (2 states) ≥7 or max. number of oscillating CN segments (3 states) ≥14; c) total size of chromothripsis event ≥20 megabase pairs (Mbp); d) satisfying the test of equal distribution of SV types (\(p > 0.05\)); and e) satisfying the test of non-random SV distribution within the cluster region or chromosome (\(p \leq 0.05\)).

**Classification of homologous recombination deficiency genotypes**

To determine Homologous Recombination Deficiency (HRD) due to possible loss-of-function of *BRCA1* and/or *BRCA2* (amongst others), we utilized the Classifier for Homologous Recombination Deficiency with default settings (CHORD; v2.0). CHORD uses a random-forest approach to classify samples into HR-deficient / HR-proficient categories.\(^78\)

**Inventory of clinically-actionable somatic alterations and putative therapeutic targets**

Current clinical relevance of somatic alterations in relation to putative treatment options or resistance mechanisms and trial eligibility was determined based upon the following databases; CiViC\(^79\) (Nov. 2018), OncoKB\(^73\) (Nov. 2018), CGI\(^80\) (Nov. 2018) and the iClusion (Dutch) clinical trial database (Sept. 2019) from iClusion (Rotterdam, the Netherlands). The databases were aggregated and harmonized using the HMF knowledgebase-importer (v1.7). This list was manually corrected for discrepancies and subsequently, we curated the linked putative treatments for current on- and off-label mNEN and mNEN-subtype treatment
options, as defined within the Netherlands by the Dutch Medicines Evaluation Board ("College ter Beoordeling van Geneesmiddelen; CBG").

Data availability
WGS data and corresponding clinical data have been requested from Hartwig Medical Foundation and provided under data request number DR-036. Both WGS and clinical data is freely available for academic use from the Hartwig Medical Foundation through standardized procedures and request forms can be found at https://www.hartwigmedicalfoundation.nl. No additional data were used for this study.

Code availability
Analysis and visualization have been performed using the statistical platform language R (3.6.2), all utilized custom code and scripts can be freely requested and distributed by contacting the authors.
References

1. Bosman, F. T. WHO classification of tumours of the digestive system - NLM Catalog - NCBI. in WHO classification of tumours of the digestive system (2010).

2. Cheng, J. et al. Pan-cancer analysis of homozygous deletions in primary tumours uncovers rare tumour suppressors. Nat. Commun. (2017) doi:10.1038/s41467-017-01355-0.

3. Scarpa, A. et al. Whole-genome landscape of pancreatic neuroendocrine tumours. Nature (2017) doi:10.1038/nature21063.

4. Do, H. K. et al. Allelic alterations in well-differentiated neuroendocrine tumors (carcinoid tumors) identified by genome-wide single nucleotide polymorphism analysis and comparison with pancreatic endocrine tumors. Genes Chromosom. Cancer (2008) doi:10.1002/gcc.20510.

5. Banck, M. S. et al. The genomic landscape of small intestine neuroendocrine tumors. J. Clin. Invest. (2013) doi:10.1172/JCI67963.

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

7. Brown, D. et al. Phylogenetic analysis of metastatic progression in breast cancer using somatic mutations and copy number aberrations. Nat. Commun. (2017) doi:10.1038/ncomms14944.

8. Navin, N. et al. Tumour evolution inferred by single-cell sequencing. Nature (2011) doi:10.1038/nature09807.

9. Hunter, K. W., Amin, R., Deasy, S., Ha, N. H. & Wakefield, L. Genetic insights into the morass of metastatic heterogeneity. Nature Reviews Cancer (2018) doi:10.1038/nrc.2017.126.

10. Walter, D. et al. Genetic heterogeneity of primary lesion and metastasis in small intestine neuroendocrine tumors. Sci. Rep. (2018) doi:10.1038/s41598-018-22115-0.

11. Wong, H. L. et al. Molecular characterization of metastatic pancreatic neuroendocrine tumors (PNETs) using whole-genome and transcriptome sequencing. Cold Spring Harb. Mol. Case Stud. (2018) doi:10.1101/mcs.a002329.

12. Cho, S. Y. et al. Cervical small cell neuroendocrine tumor mutation profiles via whole exome sequencing. Oncotarget (2017) doi:10.18632/oncotarget.14098.

13. Shen, P. et al. Comprehensive genomic profiling of neuroendocrine bladder cancer pinpoints molecular origin and potential therapeutics. Oncogene (2018) doi:10.1038/s41388-018-0192-5.

14. Priestley, P. et al. Pan-cancer whole-genome analyses of metastatic solid tumours. Nature (2019) doi:10.1038/s41586-019-1689-y.

15. Alexandrov, L. B. et al. Signatures of mutational processes in human cancer. Nature 500, 415–421 (2013).

16. Prakash, R., Zhang, Y., Feng, W. & Jasin, M. Homologous recombination and human health: The roles of BRCA1, BRCA2, and associated proteins. Cold Spring Harb. Perspect. Biol. (2015) doi:10.1101/cshperspect.a016600.

17. Chan, K. et al. An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. Nat. Genet. 47, 1067–1072 (2015).
18. Cortés-ciriano, I., Lee, J., Xi, R., Jain, D. & Jung, Y. L. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. (2018).

19. Rausch, T. et al. Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. Cell (2012) doi:10.1016/j.cell.2011.12.013.

20. Sanborn, J. Z. et al. Double minute chromosomes in glioblastoma multiforme are revealed by precise reconstruction of oncogenic amplicons. Cancer Res. (2013) doi:10.1158/0008-5472.CAN-13-0186.

21. Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Campbell, P. J. & Stratton, M. R. Deciphering Signatures of Mutational Processes Operative in Human Cancer. Cell Rep. 3, 246–259 (2013).

22. Alexandrov, L. B. et al. The Repertoire of Mutational Signatures in Human Cancer. bioRxiv (2019) doi:10.1101/322859.

23. Petljak, M. et al. Characterizing Mutational Signatures in Human Cancer Cell Lines Reveals Episodic APOBEC Mutagenesis. Cell (2019) doi:10.1016/j.cell.2019.02.012.

24. David, S. S., O’Shea, V. L. & Kundu, S. Base-excision repair of oxidative DNA damage. Nature (2007) doi:10.1038/nature05978.

25. Pea, A. et al. Genetic Analysis of Small Well-differentiated Pancreatic Neuroendocrine Tumors Identifies Subgroups With Differing Risks of Liver Metastases. Ann. Surg. (2018) doi:10.1097/sla.0000000000003022.

26. Nieser, M. et al. Loss of Chromosome 18 in Neuroendocrine Tumors of the Small Intestine: The Enigma Remains. Neuroendocrinology (2017) doi:10.1159/000446917.

27. Samstein, R. M. et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nature Genetics (2019) doi:10.1038/s41588-018-0312-8.

28. Chan, T. A. et al. Development of tumor mutation burden as an immunotherapy biomarker: Utility for the oncology clinic. Annals of Oncology (2019) doi:10.1093/annonc/mdy495.

29. Li, G. M. Mechanisms and functions of DNA mismatch repair. Cell Research (2008) doi:10.1038/cr.2007.115.

30. Yousif, F. et al. The Origins and Consequences of Localized and Global Somatic Hypermutation. bioRxiv 287839 (2018) doi:10.1101/287839.

31. Nik-Zainal, S. et al. Mutational processes molding the genomes of 21 breast cancers. Cell 149, 979–993 (2012).

32. Harris, R. S. Molecular mechanism and clinical impact of APOBEC3B-catalyzed mutagenesis in breast cancer. Breast Cancer Research (2015) doi:10.1186/s13058-014-0498-3.

33. Boichard, A., Tsigelny, I. F. & Kurzrock, R. High expression of PD-1 ligands is associated with kataegis mutational signature and APOBEC3 alterations. Oncoimmunology 6, e1284719 (2017).

34. Francis, J. M. et al. Somatic mutation of CDKN1B in small intestine neuroendocrine tumors. Nat. Genet. (2013) doi:10.1038/ng.2821.

35. Alrezk, R., Hannah-Shmouni, F. & Stratakis, C. A. MEN4 and CDKN1B mutations: the latest of the MEN syndromes. Endocr. Relat. Cancer (2017) doi:10.1530/erc-17-0243.

36. Veeriah, S. et al. The tyrosine phosphatase PTPRD is a tumor suppressor that is frequently inactivated
and mutated in glioblastoma and other human cancers. Proc. Natl. Acad. Sci. U. S. A. (2009)
doi:10.1073/pnas.0900571106.

Mitsui, J. et al. Mechanisms of genomic instabilities underlying two common fragile-site-associated
Loci, PARK2 and DMD, in germ cell and cancer cell lines. Am. J. Hum. Genet. (2010)
doi:10.1016/j.ajhg.2010.06.006.

Duan, H. et al. PARK2 suppresses proliferation and tumorigenicity in non-small cell lung cancer. Front.
Oncol. (2019) doi:10.3389/fonc.2019.00790.

Zhang, J. et al. Current understanding of the molecular biology of pancreatic neuroendocrine tumors.
Journal of the National Cancer Institute (2013) doi:10.1093/jnci/djt135.

Tang, L. H. et al. Well-differentiated neuroendocrine tumors with a morphologically apparent high-
grade component: A pathway distinct from poorly differentiated neuroendocrine carcinomas. Clin.
Cancer Res. (2016) doi:10.1158/1078-0432.CCR-15-0548.

Jiao, Y. et al. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic
neuroendocrine tumors. Science (80-. ). (2011) doi:10.1126/science.1200609.

Kochetkova, M. et al. CBFA2T3 (MTG16) is a putative breast tumor suppressor gene from the breast
cancer loss of heterozygosity region at 16q24.3. Cancer Res. (2002).

Neilsen, P. M. et al. Identification of ANKRD11 as a p53 coactivator. J. Cell Sci. (2008)
doi:10.1242/jcs.026351.

Lim, S. P. et al. Specific-site methylation of tumour suppressor ANKRD11 in breast cancer. Eur. J. Cancer
(2012) doi:10.1016/j.ejca.2012.03.023.

Noll, J. E. et al. Mutant p53 drives multinucleation and invasion through a process that is suppressed by
ANKRD11. Oncogene (2012) doi:10.1038/onc.2011.456.

Qing, T. et al. Somatic mutations in ZFHX4 gene are associated with poor overall survival of Chinese
esophageal squamous cell carcinoma patients. Sci. Rep. (2017) doi:10.1038/s41598-017-04221-7.

Oliner, J. D., Saiki, A. Y. & Caenepeel, S. The role of MDM2 amplification and overexpression in
tumorigenesis. Cold Spring Harb. Perspect. Med. (2016) doi:10.1101/cshperspect.a026336.

Sakthianandeswaren, A. et al. MACROD2 haploinsufficiency impairs catalytic activity of PARP1 and
promotes chromosome instability and growth of intestinal tumors. Cancer Discov. (2018)
doi:10.1158/2159-8290.CD-17-0909.

Ombrello, M. J. et al. Cold urticaria, immunodeficiency, and autoimmunity related to PLCG2 deletions.
N. Engl. J. Med. (2012) doi:10.1056/NEJMo1102140.

Dyer, M. A., Qadeer, Z. A., Valle-Garcia, D. & Bernstein, E. ATRX and DAXX: Mechanisms and mutations.
Cold Spring Harb. Perspect. Med. (2017) doi:10.1101/cshperspect.a026567.

Ziv, E. et al. DAXX Mutation Status of Embolization-Treated Neuroendocrine Tumors Predicts Shorter
Time to Hepatic Progression. J. Vasc. Interv. Radiol. (2018) doi:10.1016/j.jvir.2018.05.023.

Konukiewitz, B. et al. Pancreatic neuroendocrine carcinomas reveal a closer relationship to ductal
adenocarcinomas than to neuroendocrine tumors G3. Hum. Pathol. (2018)
doi:10.1016/j.humpath.2018.03.018.
53. Takizawa, N. et al. Molecular characteristics of colorectal neuroendocrine carcinoma; Similarities with adenocarcinoma rather than neuroendocrine tumor. *Hum. Pathol.* (2015) doi:10.1016/j.humpath.2015.08.006.

54. Shamir, E. R. et al. Identification of high-risk human papillomavirus and Rb/E2F pathway genomic alterations in mutually exclusive subsets of colorectal neuroendocrine carcinoma. *Mod. Pathol.* (2019) doi:10.1038/s41379-018-0131-6.

55. Mollaoglu, G. et al. MYC Drives Progression of Small Cell Lung Cancer to a Variant Neuroendocrine Subtype with Vulnerability to Aurora Kinase Inhibition. *Cancer Cell* (2017) doi:10.1016/j.ccell.2016.12.005.

56. van der Velden, D. L. et al. The Drug Rediscovery protocol facilitates the expanded use of existing anticancer drugs. *Nature* (2019) doi:10.1038/s41586-019-1600-x.

57. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760 (2009).

58. McKenna, A. et al. The genome analysis toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303 (2010).

59. Kim, S. et al. Strelka2: Fast and accurate variant calling for clinical sequencing applications. doi.org 192872 (2017) doi:10.1101/192872.

60. McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* (2016) doi:10.1186/s13059-016-0974-4.

61. Liu, X., Wu, C., Li, C. & Boerwinkle, E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Hum. Mutat.* (2016) doi:10.1002/humu.22932.

62. Lek, M. et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291 (2016).

63. Ng, P. C. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* (2003) doi:10.1093/nar/gkg509.

64. Adzhubei, I., Jordan, D. M. & Sunyaev, S. R. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr. Protoc. Hum. Genet.* (2013) doi:10.1002/0471142905.hg0720s76.

65. Cameron, D. L. et al. GRIDSS, PURPLE, LINX: Unscrambling the tumor genome via integrated analysis of structural variation and copy number. *bioRxiv* (2019) doi:10.1101/781013.

66. Broad Institute. Picard tools. https://broadinstitute.github.io/picard/ (2016).

67. van Dessel, L. F. et al. The genomic landscape of metastatic castration-resistant prostate cancers reveals multiple distinct genotypes with potential clinical impact. *Nat. Commun.* 10, 546051 (2019).

68. Stephens, P. J. et al. The landscape of cancer genes and mutational processes in breast cancer. *Nature* (2012) doi:10.1038/nature11017.

69. Martincorena, I. et al. Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* 171, 1029-1041.e21 (2017).
Zerbino, D. R. et al. Ensembl 2018. *Nucleic Acids Res.* (2018) doi:10.1093/nar/gkx1098.

Mermel, C. H. et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 12, R41 (2011).

Forbes, S. A. et al. COSMIC: Somatic cancer genetics at high-resolution. *Nucleic Acids Res.* 45, D777–D783 (2017).

Chakravarty, D. et al. OncoKB: A Precision Oncology Knowledge Base. *JCO Precis. Oncol.* 1–16 (2017) doi:10.1200/PO.17.00011.

Harrow, J. et al. GENCODE: The reference human genome annotation for the ENCODE project. *Genome Res.* 22, 1760–1774 (2012).

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

Christensen, S. et al. 5-Fluorouracil treatment induces characteristic T>G mutations in human cancer. *Nat. Commun.* (2019) doi:10.1038/s41467-019-12594-8.

Gaujoux, R. & Seoighe, C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* 11, (2010).

Nguyen, L., Martens, J., Hoeck, A. van & Cuppen, E. Pan-cancer landscape of homologous recombination deficiency. *bioRxiv* (2020) doi:10.1101/2020.01.13.905026.

Griffith, M. et al. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nat. Genet.* 49, 170–174 (2017).

Tamborero, D. et al. Cancer Genome Interpreter annotates the biological and clinical relevance of tumor alterations. *Genome Med.* 10, 25 (2018).

Pavel, M. et al. ENETS consensus guidelines update for the management of distant metastatic disease of intestinal, pancreatic, bronchial neuroendocrine neoplasms (NEN) and NEN of unknown primary site. in *Neuroendocrinology* (2016). doi:10.1159/000443167.

Acknowledgements

We would like to thank J. (Alberto) G. Nakauma Gonzalez for his assistance in implementing and describing the TAF calculations and iClusion for sharing their data on the association of genetic aberrations to actionable targets and clinical trials.

Grants

N/A.

Author contributions

JVR, BM and HJGVDW wrote the manuscript, which all authors critically reviewed. JVR and HJGVDW performed the bioinformatics analyses. BM managed clinical data assessment. FE,
MT, LMV, HK, WD and GV are clinical contributors. MJK is PI of the CPCT-02 study, SS is chair of the CPCT and both supervise the CPCT-02 study. EPJGC coordinated the sequencing of samples and contributed to the bioinformatics analyses.

**Competing interests**
All authors declare no competing interests.

**Additional information**
Supplementary information is available for this paper at https://doi.org/XXXXXXXXX.
Figures

Figure 1 - Overview of patient inclusion and sub-classification of biopsies.

a) Flowchart of patient inclusion. From the CPCT-02 cohort, single biopsies from 86 distinct patients with metastatic neuroendocrine neoplasms (mNEN) were selected. From the total pool of available whole-genome sequenced mNET samples. If multiple derived mNET biopsies from the same patient were available, we selected the mNET biopsy with the highest tumor cell purity.

b) Subclassification of mNEN based on primary localization. The 86 mNEN were subclassified, based on their primary localization, into six major categories; Foregut, Hindgut, Lung, Pancreas and Midgut; whilst samples with indeterminable localization were categorized as Unknown. The number of mNET (in blue) and mNEC (in red) are shown per category.

Figure 2 - Landscape of large-scale genomic alterations detected in mNEN, ordered by differentiation grade (NEC / NEC) and primary localization.

Overview of genome-wide characteristics of the mNEN cohort ordered by mNEC / mNET and primary localization on decreasing median tumor mutational burden. For each mNEN (n = 86), the following tracks are shown:

a) Number of genomic mutations per megabase over the entire genome (TMB); SNV, InDel and MNV are depicted in blue, orange and salmon respectively. Threshold for high-TMB (≥10) is shown by a horizontal red dotted line. Y-axis is shown in log_{10}-scale.

b) Total number of structural variants including deletions, tandem duplications, translocations, inversions and insertions as detected by GRIDSS. Y-axis is shown in log_{10}-scale.

c) Relative frequency of each of the structural variant categories; deletions in orange, tandem duplications in red, translocations in blue, inversions in light-blue and insertions in yellow.

d) Mean genome-wide ploidy, ranging from 0 (red) to 4 (green; tetraploid). Common diploid status is shown in white.

e) Relative contribution of the COSMIC single-base substitution mutational signatures (v3; n = 67). Signatures with less than 5 percent overall contribution within the entire
mNEN cohort were categorized under the “Filtered (<5%)” category. The proposed etiology of the signatures is denoted below.

f) Relative frequency of the pyrimidine mutations (SNV) in their six categories.
g) Relative frequency of Doublet Base Substitution (DBS) categories.
h) Presence of chromothripsis; mNEN with chromothripsis are shown in pink.
i) Presence of kataegis; mNEN with kataegis are shown in red.
j) Status of homologous recombination deficiency (HRD), as determined by CHORD; mNEN with BRCA1/2-associated HRD (p ≥ 0.5) are shown in pink, otherwise colored white.
k) Differentiation grade of the mNEN; NEC in red, NET in blue.
l) Primary localization of the mNEN.

Figure 3 - Putative drivers and NEN-associated genes within the mNEN cohort as detected by unbiased discovery (dN/dS, GISTIC2) and literature.

Overview of putative drivers harboring coding mutations within at least three mNEN. We show putative drivers as detected by dN/dS and/or GISTIC2 and supplemented this list with additional NEN-associated drivers. mNEN and genes are sorted based on mutually-exclusivity of the depicted putative drivers. Only GISTIC2 focal peaks with deep amplifications and deletions are shown.

a) Number of genomic mutations per megabase over the entire genome (TMB); SNV, InDel and MNV are depicted in blue, orange and salmon respectively. Threshold for high-TMB (≥10) is shown by a horizontal salmon dashed line. Y-axis is shown in log_{10} scale.
b) Mean genome-wide ploidy, ranging from 0 (red) to 4 (green; tetraploid). Diploidy is shown in white.
c) Relative contribution of the COSMIC single-base substitution mutational signatures (v3; n = 67). Signatures with less than 5 percent overall contribution within the entire mNEN cohort were categorized under the “Filtered (<5%)” category. The proposed etiology of the signatures is denoted below.
d) Overview of coding mutation(s) per mNEN, (light-)green or (light-)red backgrounds depict copy-number aberrations whilst the inner square depicts the type of (coding) mutation(s). The adjacent bar plots represent the relative proportions of mutational
categories (coding mutations (SNV, InDels and MNV), splicing mutations, SV, deep
gains (high-level amplifications resulting in many additional copies) and deep
deletions (high-level losses resulting in (near) homozygous losses) per gene. The
middle-outer barplot depicts the percentage of mNEC (in red) and mNET in blue
which harbored a mutation. In addition, dN/dS and/or GISTIC2 support are shown on
the outer-right bar plots for either the entire mNEN cohort or separate mNET/mNEC
analysis; GISTIC2 results are colored red if these genes were detected within a
recurrent focal deletion and green if detected within a recurrent focal gain.
e) Presence of chromothripsis; mNEN with chromothripsis are shown in pink.
f) Presence of kataegis; mNEN with kataegis are shown in red.
g) Status of homologous recombination deficiency (HRD), as determined by CHORD;
mNEN with BRCA1/2-associated HRD ($p \geq 0.5$) are shown in pink, otherwise colored
white.
h) Differentiation grade of the mNEN; mNEC in red, mNET in blue.
i) Primary localization of the mNEN.

Figure 4 - Putative drivers and NEN-associated genes within the pancreas- and midgut-
derived mNET as detected by unbiased discovery (dN/dS, GISTIC2) and literature.
Overview of putative drivers harboring coding mutations within at least two pancreas-
and/or midgut-derived mNET. We show putative drivers as detected by subgroup-specific
dN/dS and/or GISTIC2 and supplemented this list with additional NEN-associated drivers.
mNET and genes are sorted based on mutually-exclusivity of the depicted putative drivers.
Same layout as figure 3, except the adjacent middle-outer bar (in d) depicts the percentage
of pancreas-derived m(NET) in green and midgut-derived mNET in blue. In addition, e) represent chromosomal arm-aberrations on chr18p/q as detected by GISTIC2.

Figure 5 - Clinically-actionable somatic alterations observed within mNEN.
a) Overview of distinct mNEN harboring current clinically-actionable alterations for on-
and off-label NEN therapies. The highest NEN-therapy option (ranked as on-label
NEN subtype, on-label NEN, off-label for NEN, off-label for other cancer types but
currently-available and drugs in development) per distinct mNEN is shown.
b) mNEN harboring current clinically-actionable alterations, per gene. The highest NET-
therapy option per mNEN and gene is shown. Bottom track represents the
categorized primary localization of the mNEN whilst the right-hand side figure shown
the number of samples harboring a somatic alteration within the given gene and the
proposed level of therapy.
Supplemental data

Supp. Figure 1 - Overview of participating Dutch centers in the CPCT-02 mNEN cohort.

a) Absolute frequency of distinct included patients in the CPCT-02 mNEN cohort per participating center within the Netherlands, shown with their geographical position.

b) Absolute and relative frequencies of distinct included patients in the CPCT-02 mNEN cohort per participating center within the Netherlands.

Supp. Figure 2 - Overview of sequencing characteristics and selected biopsies.

a) Boxplot with individual data points of the estimated (in silico) tumor cell percentages based on the whole genome sequencing data with observed median displayed.

b) Boxplot with individual data points of the mean read-coverages (WGS) of the peripheral blood (reference; blue) and biopsy tissues (red) with observed median per variable displayed.

c) Age distribution stratified by gender of the mNEN cohort with observed median per variable displayed in a boxplot with individual data points.

d) Generalized location of the biopsy with the relative frequency shown between brackets.

Supp. Figure 3 - Overview of mutational landscape, categorized per differentiation grade.

a) Number of SNV (gray), InDels (green), MNV (salmon) per whole-genome sequenced sample over three resolutions; genome-wide, within intragenic regions and within coding regions with observed median per variable displayed. Data is categorized on mNEC / mNET status.

b) Type of genome-wide SNVs. Transition (Ti) and transversion (Tv), with a special attention for C to T (Ti) in CpG context, are indicated per sample with observed median per variable displayed. Data is categorized on mNEC / mNET status.

c) Mean genome-wide tumor ploidy based on all autosomal chromosomes with observed median displayed. Data is categorized on mNEC / mNET status.

d) Frequency of Tandem Duplications, Insertions, Inversions, Deletions and interchromosomal translocations are indicated per sample with observed median per variable displayed. Data is categorized on mNEC / mNET status.
e) Mutational consequences of genomic variants overlapping genes using Ensembl Variant Effect Predictor (VEP). Data is categorized on mNEC / mNET status.
f) Genome-wide ratio of transitions (Ti) over transversion (Tv) with observed median displayed. Data is categorized on mNEC / mNET status.

Supp. Figure 4 - Overview of the distribution of somatically-acquired structural variants.
Overview of the genomic sizes and numbers of structural variants present in the mNEN cohort. Samples are sorted based on primary localization and decreasing number of total observed structural variants over all categories (deletions, tandem duplications, inversions, translocations and insertions).

a) Track displays boxplots representing the genomic width of deletions; Y-axis is in log10-scale. Lower track displays the total number of observed deletions.
b) Track displays boxplots representing the genomic width of tandem duplications; Y-axis is in log10-scale. Lower track displays the total number of observed tandem duplications.
c) Track displays boxplots representing the genomic width of inversions; Y-axis is in log10-scale. Lower track displays the total number of observed inversions.
d) Total number of observed translocations.
e) Total number of observed insertions.
f) Presence of chromothripsis, kataegis and/or predicted HRD (CHORD). Bottom track displays differentiation grade (NEC/NET) and primary localization of mNEN.

Supp. Figure 5 - Observed kataegis events within the mNET cohort.
a) Number of observed kataegis foci in the mNEN cohort (found in 6 distinct samples, blue bars) and the respective cumulative genomic width of all observed kataegis foci per sample (right y-axis; black points).
b) Relative frequency of SNV categories found in all observed kataegis foci per sample.
c) Relative frequency of SNV in observed kataegis foci with APOBEC-related TpCpW mutational context. W stands for T or A changes.
d) Genome-wide relative contribution to mutational signatures (COSMIC v3) for the respective mNET sample.
Representation of a single kataegis foci on chromosome 8 within a single respective sample (highlighted with * in a). SNV (colored on pyrimidine mutations) are shown with relative genomic distances (in log_{10}) to neighboring SNV. Observed kataegis focus on chromosome 8 is highlighted with a transparent red background.

Absolute mutational contribution of APOBEC COSMIC (v3) signatures (2 & 13) for samples without (n = 80) and with observed kataegis foci (n = 6). Statistical significance was tested with Wilcoxon rank-sum test and is denoted with * ≤ 0.05, ** ≤ 0.01 and *** ≤ 0.001.

Supp. Figure 6 - Genomic overview of mNEN displaying chromothripsis-like events.
Genomic representations of the chromothripsis-harboring mNEN (n = 6). The outer track displays the genomic ideogram, the second-outer track displays copy number profiles (amplification in light green; deep amplification beyond sample-specific threshold (GISTIC2) in dark green, deletions in blue; deep deletions beyond sample-specific threshold (GISTIC2) in dark blue). The third track displays TC%-corrected lower allele-frequency (LAF) values of individual copy number segments (LAF ≤ 0.33 in pink; LAF ≥ 0.33 in black). The fourth track displays the number of mutations per 5 Mbp, ranging from 0 to 60+; bins with ≥ 20 mutations are highlighted in blue. The fifth track highlights the regions harboring chromothripsis in a red line. The innermost track displays the breakpoints of the structural variants; interchromosomal translocations in dark blue, deletions in gray, insertions in yellow, inversion in light blue and tandem duplications in red. Samples are colored per NEC (in red) and NET (in blue) status.

Supp. Figure 7 - de novo mutational signatures assessment on mNEN.
Assessment and comparison of extracted de novo single base substitution mutational signatures (n = 7; Sig. A - I) using non-negative matrix factorization (NMF) within the mNEN cohort against the known COSMIC (v3; n = 67) signatures.

a) Overview of extracted de novo single base substitution mutational signatures (n = 7; Sig. A - I; upper track) vs. COSMIC signatures (v3; n = 67; lower track), per mNEN. mNEN are sorted based on unsupervised clustering (Ward.D; Euclidean distance; distances plotted in log_{10}-scale) of the relative contribution of the nine de novo mutational signatures.
b) Cosine similarity of the *de novo* mutational signatures against the known COSMIC v3 signatures (*n* = 67).

c) Trinucleotide mutational contexts of the nine extracted *de novo* signatures.

d) NMF quality metrics using between two to fifteen ranks over 1000 iterations.

**Supp. Figure 8 - Copy-number overview of mNEN cohort and subpopulations with re-occurring and focal amplifications and deletion highlighted (GISTIC2) and unbiased driver gene analysis.**

Circosplots with ideogram of recurrent copy-number aberrations as detected by GISTIC2 per sub-population (as shown above each circosplot). G-scores are depicted on the y-axis. Regions with amplifications (G-score > 0) are depicted in green and deletions (G-score < 0) in blue. Regions with significant (and recurring) copy-number aberrations (*q* ≤ 0.1) are denoted with a darker shade of green or blue, respective of amplification or deletion. Per region, the foci of maximal amplification or deletion (focal peaks; *q* ≤ 0.1) are denoted in the inner track; the peak identifiers with associated genes are also denoted and presented in supplementary table 1.

**Supp. Figure 9 - Genomic characteristics per differentiation grade (NEC/NET) and primary localization within mNET.**

a) Number of SNV, InDels, MNV and SV per whole-genome mNEN with observed median per variable displayed. Data is categorized on mNEC and distinct mNET subgroups based on primary localization.

b) Tumor mutational burdens (genome-wide; log<sub>10</sub>), with observed median per variable displayed. Data is categorized on mNEC and distinct mNET subgroups based on primary localization.

c) Mean genome-wide ploidy, with observed median per variable displayed. Data is categorized on mNEC and distinct mNET subgroups based on primary localization.

d) Number of genes harboring somatic coding mutations, with observed median per variable displayed. Data is categorized on mNEC and distinct mNET subgroups based on primary localization.

e) Mutational enrichment of mutant genes (mutations and copy-number alterations) and large-scale events (kataegis and chromothripsis) between our three major subgroups; mNEC, pancreas- and midgut-derived mNET. Statistical significance was
tested using a one-sided Fisher’s Exact Test with BH correction; significance is
denoted by *** ($q \leq 0.001$), ** ($q \leq 0.01$), * ($q \leq 0.05$) and . ($q \leq 0.1$).

**Supp. Figure 10 - Copy-number aberrations of chromosomal arms per differentiation
grade (NEC/NET) and primary localization within mNET.**

a) Unsupervised clustering (Euclidean distances, Ward.D2 method) of the mNET
samples based on the categorization of chromosomal arm copy-number aberrations
(based on GISTIC2 value per arm). Top color-bars depict the differentiation grade of
the mNEN (mNEC in red, mNET in blue) and the primary localization.

b) Overview of the relative frequency of samples with amplifications (green) and losses
(red) per arm within the given subgroup. Statistically significant ($q \leq 0.05$) arm-level
copy-number aberrations are depicted with an asterisk whilst the non-significant
events are shown as transparent.

**Supp. Table 1 - Overview of included patients and data presented in figures.**

Overview of all data presented and quantified in this manuscript.
**Figure 1 - Overview of patient inclusion and sub-classification of biopsies.**

a) Flowchart of patient inclusion. From the CPCT-02 cohort, single biopsies from 86 distinct patients with metastatic neuroendocrine neoplasms (mNEN) were selected. From the total pool of available whole-genome sequenced mNET samples. If multiple derived mNET biopsies from the same patient were available, we selected the mNET biopsy with the highest tumor cell purity.

b) Subclassification of mNEN based on primary localization. The 86 mNEN were subclassified, based on their primary localization, into six major categories; Foregut, Hindgut, Lung, Pancreas and Midgut; whilst samples with indeterminable localization were categorized as Unknown. The number of mNET (in blue) and mNEC (in red) are shown per category.
Figure 2 - Landscape of large-scale genomic alterations detected in mNEN, ordered by differentiation grade (NEC / NEC) and primary localization.

Overview of genome-wide characteristics of the mNEN cohort ordered by mNEC / mNET and primary localization on decreasing median tumor mutational burden. For each mNEN (n = 86), the following tracks are shown:

a) Number of genomic mutations per megabase over the entire genome (TMB); SNV, InDel and MNV are depicted in blue, orange and salmon respectively. Threshold for high-TMB (≥10) is shown by a horizontal red dotted line. Y-axis is shown in log10-scale.

b) Total number of structural variants including deletions, tandem duplications, translocations, inversions and insertions as detected by GRIDSS. Y-axis is shown in log10-scale.

c) Relative frequency of each of the structural variant categories; deletions in orange, tandem duplications in red, translocations in blue, inversions in light-blue and insertions in yellow.

d) Mean genome-wide ploidy, ranging from 0 (red) to 4 (green; tetraploid). Common diploid status is shown in white.

e) Relative contribution of the COSMIC single-base substitution mutational signatures (v3; n = 67). Signatures with less than 5 percent overall contribution within the entire mNEN cohort were categorized under the “Filtered (<5%)“ category. The proposed etiology of the signatures is denoted below.

f) Relative frequency of the pyrimidine mutations (SNV) in their six categories.

g) Relative frequency of Doublet Base Substitution (DBS) categories.

h) Presence of chromothripsis; mNEN with chromothripsis are shown in pink.

i) Presence of kataegis; mNEN with kataegis are shown in red.

j) Status of homologous recombination deficiency (HRD), as determined by CHORD; mNEN with BRCA2/2-associated HRD (p ≥ 0.5) are shown in pink, otherwise colored white.

k) Differentiation grade of the mNEN; NEC in red, NET in blue.

I) Primary localization of the mNEN.
Figure 3 - Putative drivers and NEN-associated genes within the mNEN cohort as detected by unbiased discovery (dN/dS, GISTIC2) and literature.

Overview of putative drivers harboring coding mutations within at least three mNEN. We show putative drivers as detected by dN/dS and/or GISTIC2 and supplemented this list with additional NEN-associated drivers. mNEN and genes are sorted based on mutational exclusivity of the depicted putative drivers. Only GISTIC2 focal peaks with deep amplifications and deletions are shown.

- a) Number of genomic mutations per megabase over the entire genome (TMB); SNV, Indel and MNV are depicted in blue, orange and salmon respectively. Threshold for high-TMB (≥10) is shown by a horizontal salmon dashed line. Y-axis is shown in log10-scale.
- b) Mean genome-wide ploidy, ranging from 0 (red) to 4 (green; tetraploid). Diploidy is shown in white.
- c) Relative contribution of the COSMIC single-base substitution mutational signatures (v3; n = 67). Signatures with less than 5 percent overall contribution within the entire mNEN cohort were categorized under the "Filtered (<5%)" category. The proposed etiology of the signatures is denoted below.
- d) Overview of coding mutation(s) per mNEN, (light-)green or (light-)red backgrounds depict copy-number aberrations whilst the inner square depicts the type of (coding) mutation(s). The adjacent bar plots represent the relative proportions of mutational categories (coding mutations (SNV, Indel and MNV), splicing mutations, SV, deep gains (high-level amplifications resulting in many additional copies) and deep deletions (high-level losses resulting in (near) homozygous losses) per gene. The middle-outter barplots depict the percentage of mNEN in red and mNEN in blue which harbored a mutation. In addition, dN/dS and/or GISTIC2 support are shown on the outer-right bar plots for either the entire mNEN cohort or separate mNET/mNEC analysis; GISTIC2 results are colored red if these genes were detected within a recurrent focal deletion and green if detected within a recurrent focal gain.
- e) Presence of chromothripsis; mNEN with chromothripsis are shown in pink.
- f) Presence of kataegis; mNEN with kataegis are shown in red.
- g) Status of homologous recombination deficiency (HRD), as determined by CHORD; mNEN with BRCA1/2-associated HRD (p ≤ 0.5) are shown in pink, otherwise colored white.
- h) Differentiation grade of the mNEN; mNEC in red, mNET in blue.
- i) Primary localization of the mNEN.
Overview of putative drivers harboring coding mutations within at least two pancreas- and/or midgut-derived mNET. We show putative drivers as detected by subgroup-specific dN/dS and/or GISTIC2 and supplemented this list with additional NEN-associated drivers. mNET and genes are sorted based on mutually-exclusivity of the depicted putative drivers. Same layout as figure 3, except the adjacent middle-outer bar (in d) depicts the percentage of pancreas-derived m(NET) in green and midgut-derived mNETH in blue. In addition, e) represent chromosomal arm- aberrations on chr18p/q as detected by GISTIC2.
Figure 5 - Clinically-actionable somatic alterations observed within mNEN.

a) Overview of distinct mNEN harboring current clinically-actionable alterations for on- and off-label NEN therapies. The highest NEN-therapy option (ranked as on-label NEN subtype, on-label NEN, off-label for NEN, off-label for other cancer types but currently-available and drugs in development) per distinct mNEN is shown.

b) mNEN harboring current clinically-actionable alterations, per gene. The highest NET-therapy option per mNEN and gene is shown. Bottom track represents the categorized primary localization of the mNEN whilst the right-hand side figure shown the number of samples harboring a somatic alteration within the given gene and the proposed level of therapy.
Supp. Figure 1 - Overview of participating Dutch centers in the CPT-02 mNEN cohort.

a) Absolute frequency of distinct included patients in the CPT-02 mNEN cohort per participating center within the Netherlands, shown with their geographical position.

b) Absolute and relative frequencies of distinct included patients in the CPT-02 mNEN cohort per participating center within the Netherlands.
**Supp. Figure 2 - Overview of sequencing characteristics and selected biopsies.**

a) Boxplot with individual data points of the estimated (*in silico*) tumor cell percentages based on the whole genome sequencing data with observed median displayed.

b) Boxplot with individual data points of the mean read-coverages (WGS) of the peripheral blood (reference; blue) and biopsy tissues (red) with observed median per variable displayed.

c) Age distribution stratified by gender of the mNEN cohort with observed median per variable displayed in a boxplot with individual data points.

d) Generalized location of the biopsy with the relative frequency shown between brackets.
Supp. Figure 3 - Overview of mutational landscape, categorized per differentiation grade.

a) Number of SNV (gray), InDels (green), MNV (salmon) per whole-genome sequenced sample over three resolutions; genome-wide, within intragenic regions and within coding regions with observed median per variable displayed. Data is categorized on mNEC / mNET status.

b) Type of genome-wide SNVs. Transition (Ti) and transversion (Tv), with a special attention for C to T (Ti) in CpG context, are indicated per sample with observed median per variable displayed. Data is categorized on mNEC / mNET status.

c) Mean genome-wide tumor ploidy based on all autosomal chromosomes with observed median displayed. Data is categorized on mNEC / mNET status.

d) Frequency of Tandem Duplications, Insertions, Inversions, Deletions and interchromosomal translocations are indicated per sample with observed median per variable displayed. Data is categorized on mNEC / mNET status.

e) Mutational consequences of genomic variants overlapping genes using Ensembl Variant Effect Predictor (VEP). Data is categorized on mNEC / mNET status.

f) Genome-wide ratio of transitions (Ti) over transversion (Tv) with observed median displayed. Data is categorized on mNEC / mNET status.
Supp. Figure 4 - Overview of the distribution of somatically-acquired structural variants.

Overview of the genomic sizes and numbers of structural variants present in the mNEN cohort. Samples are sorted based on primary localization and decreasing number of total observed structural variants over all categories (deletions, tandem duplications, inversions, translocations and insertions).

a) Track displays boxplots representing the genomic width of deletions; Y-axis is in log10-scale. Lower track displays the total number of observed deletions.

b) Track displays boxplots representing the genomic width of tandem duplications; Y-axis is in log10-scale. Lower track displays the total number of observed tandem duplications.

c) Track displays boxplots representing the genomic width of inversions; Y-axis is in log10-scale. Lower track displays the total number of observed inversions.

d) Total number of observed translocations.

e) Total number of observed insertions.

f) Presence of chromothripsis, kataegis and/or predicted HRD (CHORD). Bottom track displays differentiation grade (NEC/NET) and primary localization of mNEN.
Supp. Figure 5 - Observed kataegis events within the mNET cohort.

a) Number of observed kataegis foci in the mNEN cohort (found in 6 distinct samples, blue bars) and the respective cumulative genomic width of all observed kataegis foci per sample (right y-axis; black points).

b) Relative frequency of SNV categories found in all observed kataegis foci per sample.

c) Relative frequency of SNV in observed kataegis foci with APOBEC-related TpCpW mutational context. W stands for T or A changes.

d) Genome-wide relative contribution to mutational signatures (COSMIC v3) for the respective mNET sample.

e) Representation of a single kataegis foci on chromosome 8 within a single respective sample (highlighted with * in a). SNV (colored on pyrimidine mutations) are shown with relative genomic distances (in log10) to neighboring SNV. Observed kataegis focus on chromosome 8 is highlighted with a transparent red background.

f) Absolute mutational contribution of APOBEC COSMIC (v3) signatures (2 & 13) for samples without (n = 80) and with observed kataegis foci (n = 6). Statistical significance was tested with Wilcoxon rank-sum test and is denoted with * ≤ 0.05, ** ≤ 0.01 and *** ≤ 0.001.
Supp. Figure 6 - Genomic overview of mNEN displaying chromothripsis-like events.

Genomic representations of the chromothripsis-harboring mNEN (n = 6). The outer track displays the genomic ideogram, the second-outer track displays copy number profiles (amplification in light green; deep amplification beyond sample-specific threshold (GISTIC2) in dark green, deletions in blue; deep deletions beyond sample-specific threshold (GISTIC2) in dark blue). The third track displays TC%-corrected lower allele-frequency (LAF) values of individual copy number segments (LAF ≤ 0.33 in pink; LAF ≥ 0.33 in black). The fourth track displays the number of mutations per 5 Mbp, ranging from 0 to 60+; bins with ≥ 20 mutations are highlighted in blue. The fifth track highlights the regions harboring chromothripsis in a red line. The innermost track displays the breakpoints of the structural variants; interchromosomal translocations in dark blue, deletions in gray, insertions in yellow, inversion in light blue and tandem duplications in red. Samples are colored per NEC (in red) and NET (in blue) status.
Supp. Figure 7 - de novo mutational signatures assessment on mNEN. Assessment and comparison of extracted de novo single base substitution mutational signatures \((n = 7; \text{Sig. A - t})\) using non-negative matrix factorization (NMF) within the mNEN cohort against the known COSMIC \((v3; n = 67)\) signatures.

a) Overview of extracted de novo single base substitution mutational signatures \((n = 7; \text{Sig. A - t}; \text{upper track})\) vs. COSMIC signatures \((v3; n = 67; \text{lower track})\) per mNEN. mNEN are sorted based on unsupervised clustering (Ward.D; Euclidean distance; distances plotted in log10-scale) of the relative contribution of the nine de novo mutational signatures.

b) Cosine similarity of the de novo mutational signatures against the known COSMIC \(v3\) signatures \((n = 67)\).

c) Trinucleotide mutational contexts of the nine extracted de novo signatures.

d) NMF quality metrics using between two to fifteen ranks over 1000 iterations.
Supp. Figure 8 - Copy-number overview of mNEN cohort and subpopulations with re-occurring and focal amplifications and deletion highlighted (GISTIC2) and unbiased driver gene analysis. Circosplots with ideogram of recurrent copy-number aberrations as detected by GISTIC2 per sub-population (as shown above each circosplot). G-scores are depicted on the y-axis. Regions with amplifications (G-score > 0) are depicted in green and deletions (G-score < 0) in blue. Regions with significant (and recurring) copy-number aberrations (q ≤ 0.1) are denoted with a darker shade of green or blue, respective of amplification or deletion. Per region, the foci of maximal amplification or deletion (focal peaks; q ≤ 0.1) are denoted in the inner track; the peak identifiers with associated genes are also denoted and presented in supplementary table 1.
Supp. Figure 9 - Genomic characteristics per differentiation grade (NEC/NET) and primary localization within mNET.

**a** Number of SNV, InDels, MNV and SV per whole-genome mNEN with observed median per variable displayed. Data is categorized on mNEC and distinct mNET subgroups based on primary localization.

**b** Tumor mutational burdens (genome-wide; log10), with observed median per variable displayed. Data is categorized on mNEC and distinct mNET subgroups based on primary localization.

**c** Mean genome-wide ploidy, with observed median per variable displayed. Data is categorized on mNEC and distinct mNET subgroups based on primary localization.

**d** Number of genes harboring somatic coding mutations, with observed median per variable displayed. Data is categorized on mNEC and distinct mNET subgroups based on primary localization.

**e** Mutational enrichment of mutant genes (mutations and copy-number alterations) and large-scale events (kataegis and chromothripsis) between our three major subgroups; mNEC, pancreas- and midgut-derived mNET. Statistical significance was tested using a one-sided Fisher’s Exact Test with BH correction; significance is denoted by *** (q ≤ 0.001), ** (q ≤ 0.01), * (q ≤ 0.05) and . (q ≤ 0.1).
Supp. Figure 10 - Copy-number aberrations of chromosomal arms per differentiation grade (NEC/NET) and primary localization within mNET.

a) Unsupervised clustering (Euclidean distances, Ward.D2 method) of the mNET samples based on the categorization of chromosomal arm copy-number aberrations (based on GISTIC2 value per arm). Top color-bars depict the differentiation grade of the mNEN (mNEN in red, mNET in blue) and the primary localization.

b) Overview of the relative frequency of samples with amplifications (green) and losses (red) per arm within the given subgroup. Statistically significant ($q \leq 0.05$) arm-level copy-number aberrations are depicted with an asterisk whilst the non-significant events are shown as transparent.