Merkel cell carcinoma (MCC) is an aggressive neuroendocrine carcinoma of the skin with two distinct etiologies. Clonal integration of Merkel cell polyomavirus DNA into the tumor genome with persistent expression of viral T antigens causes at least 60% of all MCC. UV damage leading to highly mutated genomes causes a nonviral form of MCC. Despite these distinct etiologies, both forms of MCC are similar in presentation, prognosis, and response to therapy. At least three oncogenic transcriptional programs feature prominently in both forms of MCC driven by the virus or by mutation. Both forms of MCC have a high proliferative growth rate with increased levels of cell cycle–dependent genes due to inactivation of the tumor suppressors RB and p53, a strong MYC signature due to MYCL activation by the virus or gene amplification, and an attenuated neuroendocrine differentiation program driven by the ATOH1 transcription factor.
MCC: Merkel cell carcinoma

Atonal bHLH transcription factor 1 (ATOH1): a transcription factor required for Merkel cell development

MCC

Research into the development, physiology, and molecular features of normal Merkel cells has helped guide our understanding of Merkel cell carcinoma (MCC) tumor biology. Merkel cells were first described in 1875 by the German anatomist Friedrich Sigmund Merkel as tastzellen or touch cells (1). He found these cells to be in close association with nerve endings and proposed that they function as mechanoreceptors. Merkel cells are typically located in the basal layer of the skin epithelium and surround hair follicles. Clusters of Merkel cells also form touch domes, specialized structures present in areas of highly sensitive skin. Merkel cells make synaptic connections to slowly adapting type I afferent sensory nerves (1). Lineage tracing in mice identified a series of transcription factors essential for the development of normal Merkel cells. The ATOH1 (atonal bHLH transcription factor 1) gene is required for the development of Merkel cells (2). Conditional knockout of ATOH1 using a Cre deleter strain driven by the epidermal cytokeratin 14 (KRT14, CK14) promoter eliminated the development of Merkel cells. In contrast, Cre driven by the neural crest–specific WNT1 promoter had no impact on the number or location of Merkel cells (3). This important experiment indicated that normal Merkel cells are likely derived from an epidermal skin precursor and not from the neural crest. SOX2 and ISL1 (Islet1) form a heterodimeric transcription factor that cooperates to sustain ATOH1 expression (4, 5).

Merkel cells express cytokeratins CK8, CK18, CK19, and CK20, characteristic of epidermal cells. They also express several markers common to neuroendocrine cells, including neuron-specific enolase, synaptophysin, chromogranin A, CD56, and INSM1 (Insulinoma1) (1, 6). In addition to these epithelial and neuroendocrine markers, Merkel cells express PIEZO2, which functions as a mechanically activated ion channel that can detect and convert mechanical stimuli into electrical signals (7). PIEZO2 is required for the Merkel cell touch sensation (8). Recent structural analysis of the homologous PIEZO1 revealed a homotrimer that assembles into a three-bladed propeller-shaped pattern that can flex in response to mechanical force (9). Given the overall homology to PIEZO1, it is likely that PIEZO2 forms a similar three-propeller structure that serves to transmit gentle touch pressure in the Merkel cell to the neural synapse (10).

Several signaling pathways have been described to influence Merkel cell development. Notch signaling antagonizes ATOH1 signaling by activating the HES family of the helix-loop-helix family of DNA binding proteins that competes with ATOH1 for specific binding to DNA. Knockout of HES1 or the essential Notch coactivator RBPJ results in increased numbers of Merkel cells (11). Polycomb repressive complex 1 (PRC1) and PRC2 repress the development of Merkel cells in neonatal skin, and loss of PRC1 catalytic activity also results in an increased number of Merkel cells (12, 13). Hedgehog and bone morphogenetic protein signaling contribute to Merkel cell development, particularly during formation of the touch dome and hair follicle (14–16).

MCC

MCC was first described in 1972 by Cyril Toker (17) as a trabecular carcinoma of the skin with carcinoid features. Later, he reported the presence of neurosecretory granules—membrane-bound granules containing dense cores—within the tumor cells. This feature is indistinguishable from tumor cells of neural crest origin and is also present in normal Merkel cells (18). The tumor name was changed to MCC to reflect the similarity in appearance of tumor cells to Merkel cells (19, 20).

MCC is an aggressive neuroendocrine carcinoma of the skin that frequently metastasizes to draining lymph nodes and distant organs, including liver, bone, pancreas, lung, and brain (21). MCC typically presents as a rapidly growing, erythematous lesion in the dermal layer of the skin. The most common presentation of MCC is in older, fair-skinned adults with a lifelong history of...
intense UV exposure from the sun. MCC occurs less frequently in non-sun-exposed skin as well as in children, young adults, and dark-skinned persons. Residence in a latitude closer to the equator is associated with increased incidence of MCC in North American men, but not women, possibly due to occupational sunlight exposure patterns (22). Risk for developing MCC is also increased in patients with severely immunocompromising conditions, including HIV/AIDS, or from medical treatment of autoimmune diseases, solid organ transplantation, and other types of cancers (23). The AEIOU mnemonic accounts for 90% of all MCC presentation: asymptomatic/lack of tenderness, expanding rapidly, immune suppression, older than 50 years, and UV-exposed/fair skin (24).

The most recent MCC staging system from the American Joint Committee on Cancer estimates a 5-year overall survival rate of 51% for local disease, 35% for nodal involvement, and 14% for metastatic disease (25, 26). Surgery and radiation therapy can be curative for local and nodal MCC, but systemic therapy is usually required for extensive, metastatic, and recurrent disease. Cytotoxic chemotherapy, based on cisplatin and etoposide regimens, has a high response rate but is limited by a short duration with a mean progression-free survival of just 94 days (27). A revolution in MCC care began recently when it was determined that checkpoint blockade therapy with antibodies to PD-1 or PD-L1 could induce frequent and durable responses (28–31). Predictions for overall survival may improve as experience with checkpoint blockade therapy increases.

**HISTOLOGY OF MCC**

MCC can vary from a pure neuroendocrine histology to a variant form with mixed histologic features. High-grade neuroendocrine MCC cells have a high nuclear to cytoplasmic ratio with scant cytoplasm, giving them the appearance of small blue cell tumors when stained by hematoxylin and eosin. The tumor nuclei have an open, pepper-and-salt-appearing chromatin pattern with frequent mitotic figures indicative of a high proliferative rate (Figure 1). Immunohistochemistry (IHC) staining of MCC for neuroendocrine markers is typically positive for chromogranin, synaptophysin, CD56, and neurofilament. MCC also stains specifically for CK20, which typically shows a paranuclear dot-like pattern. In contrast, CK20 staining in normal Merkel cells is more uniformly distributed throughout the cytoplasm. CK20 staining can distinguish MCC from other more common neuroendocrine tumors such as small cell lung carcinoma (SCLC) (32). SCLC stains positive for TTF-1 (thyroid-specific transcription factor 1, encoded by the NKX2-1 gene), while MCC is negative for this stain. INSM1 has recently emerged as a useful IHC marker for MCC and Merkel cells as well as for other neuroendocrine carcinomas (6).

There are several reports of combined MCC tumors with a mixed histology showing a neuroendocrine component and other tumor types such as squamous cell carcinoma. There have been suggestions that these tumors represent collision tumors from two separate origins. However, several reports demonstrate that certain mutations and histologic markers are shared between the two components, which suggests that the MCC tumor developed from a preexisting neoplasm (33, 34).

**POLYOMAVIRUS-POSITIVE MCC COMPARED TO NONVIRAL MCC**

A pathogenic cause for MCC was first suspected when it was recognized that the incidence of MCC was increased more than tenfold in people living with HIV-1/AIDS compared to the general population (35). The risk for developing MCC is also increased in patients with medically induced immunosuppression for autoimmune conditions such as rheumatoid arthritis and solid organ transplantation (36). Recognizing the increased risk for developing MCC by immunosuppression, Huichen Feng and Masahiro Shuda in the laboratory of Yuan Chang and Patrick Moore...
MCPyV: Merkel cell polyomavirus

Merkel cell carcinoma, nonviral (MCCN): often contains a high tumor mutational burden with a strong UV mutational signature

Tumor mutational burden (TMB): somatic mutations per megabase; high TMB gives rise to neoantigens

MCCP: Merkel cell carcinoma, polyomavirus-associated

began a search for a pathogenic cause for MCC (37). They performed whole transcriptome sequencing and searched for pathogens by first subtracting all human genes from their analysis. In the remaining unmapped sequences, novel transcripts distantly related to polyomaviruses were detected. Complete sequencing of the viral genome in MCC tumors led to the determination that it corresponded to a new human polyomavirus called Merkel cell polyomavirus (MCPyV or MCV) (37). MCPyV viral DNA was shown to be integrated into the tumor genome by Southern blotting in 8 of 10 tested MCC tumors. Evidence that MCPyV DNA was likely causative or at least an early event in MCC tumorigenesis was implied by an identical restriction fragment length polymorphism pattern observed in a primary skin tumor and a metastatic lymph node from the same patient.

Additional sequencing studies of MCC tumors revealed significant differences in MCC tumors containing MCPyV and those without a virus. Nonviral MCC (MCCN) tumors have a predominant UV mutational signature with a very high tumor mutational burden (TMB), often greater than 20 somatic mutations per megabase (38–40). In contrast, polyomavirus-associated MCC (MCCP) tumors have a low TMB of 6 or less, without a UV mutational signature (41). Whole genome sequencing revealed evidence for a highly damaged genome in MCCN with many somatic single nucleotide variants, copy number alterations, and translocations. In contrast, MCCP tumors typically have near-normal diploid genomes with few somatic mutations (Figure 2) (42, 43).

Figure 1
Merkel cell carcinoma. (a) A predominantly dermal nodule is shown consisting of small round blue cells. (b) The cells have a high nuclear to cytoplasmic ratio and pale chromatin with nuclear molding. (c) Lymphovascular invasion (arrow) is a common feature. (d) Perinuclear dot-like reactivity is shown for CK20 (hematoxylin and eosin stain). Figure adapted from Reference 128.
Somatic junctions
Density of homozygous SNPs
Density of heterozygous SNPs
Loss of heterozygosity (LOH)
Called level
Chromosome position (Mb)
Density of somatic variants
Chromosome number
Lesser allele fraction (LAF)
Karyotype
Gene symbols for impacted genes
Somatic junctions
LOH
Called level
Chromosome position (Mb)
Density of somatic variants
Chromosome number
Lesser allele fraction (LAF)
Karyotype
Gene symbols for impacted genes
SPOCK3
STOX2
UTP3

(Caption appears on following page)
Commonly mutated genes in MCCN include loss-of-function mutations in the tumor suppressor genes RB1 and TP53 (Figure 3). Both TP53 and RB1 are usually wild-type in MCCP, but inactivating mutations have been reported (44, 45). Loss-of-function mutations in NOTCH1, KMT2C, and KMT2D are also frequently observed in MCCN (41, 43). PI3K signaling is likely activated in both MCCP and MCCN with activating mutations present in PIK3CA or loss of the negative regulators PTEN, TSCI, and TSC2 (39, 46).

In addition to loss of tumor suppressors, amplification of the MYC paralog MYCL (MYCL1 or L-MYC) is frequently observed in MCCN (43, 47). MYCL was first described in a subset of SCLC (48). Amplification of MYC, MYCN, and MYCL occurs frequently in SCLC in a mutually exclusive pattern, which indicates that these paralogs likely provide overlapping oncogenic functions (49). MYCL is not required for normal mouse development, although it is expressed in developing kidney, lung, and brain (50). More recently, MYCL has been shown to be required for the development of the Batf3-dependent subset of classical dendritic (cDC1) cells (51, 52).

**Figure 3** Molecular subtypes of MCC, showing an oncoprint of co-occurrence of TMB status, dominant mutational signature, viral status, and most commonly altered genes demonstrating mutual exclusivity of MCPyV, APOBEC integration, and UV damage. Each column corresponds to one unique MCC tumor. Tumors are sorted in descending order by TMB from the left. Abbreviations: Mb, megabase; MCC, Merkel cell carcinoma; MCPyV, Merkel cell polyomavirus; TMB, tumor mutational burden. Figure adapted from Reference 41.
**MERKEL CELL POLYOMAVIRUS**

MCPyV was the fifth human polyomavirus to be identified. There are now 13 known human polyomaviruses. MCPyV DNA can be readily detected on the skin of healthy individuals, although the specific skin cells that support viral replication in vivo are not known. Two additional human polyomaviruses, HPyV6 and HPyV7, were identified from the skin or hair follicles of healthy adults using a technique called rolling circle amplification, which takes advantage of the small circular nature of the polyomavirus double-stranded DNA (dsDNA) genome (53). In immunocompromised conditions, HPyV6 and HPyV7 can replicate and cause a hyperkeratotic skin condition characterized by pruritic and brown plaques with epidermal hyperplasia and virus-laden keratinocytes (54, 55). The condition is called HPyV6- and HPyV7-associated pruritic and dyskeratotic dermatoses.

Rolling circle amplification identified a fourth human polyomavirus associated with the skin and was detected in a patient with a rare skin disease named trichodysplasia spinulosa (56). In this condition, trichodysplasia spinulosa–associated polyomavirus (TSPyV) replicates in the inner root sheath of the hair follicle, destroying the normal hair structure and leading to alopecia and folliculitis. Given that these four human polyomaviruses can be isolated from the skin and TSPyV can replicate in hair follicles, while HPyV6 and HPyV7 replicate in keratinocytes, it is likely that MCPyV can also replicate in one or both of these cell types. Of note, an image was reported of scalp folliculitis that presented in a double-lung transplant recipient that was immunostained with antibodies to MCPyV large T antigen (LT), suggesting that MCPyV could replicate in hair follicles and cause destruction in a manner similar to TSPyV (57). It should be noted that MCC typically presents on hairy skin, which is consistent with the idea that MCPyV may replicate in hair follicles. Although there have been reports of MCC presenting in mucosal tissues, most cases of mucosal MCC appear to be nonviral. However, at least some MCC tumors of the nasopharynx may contain MCPyV (58, 59).

The MCPyV dsDNA circular genome is approximately 5.4 kb and can be divided into three regions: The early viral gene region (EVGR) encodes genes that are expressed prior to the onset of viral DNA replication; the late viral gene region (LVGR) encodes genes expressed after viral DNA replication commences; and the regulatory region, called the noncoding control region (NCCR), contains the viral origin of replication and the promoters and enhancers that drive expression of the early and late viral genes.

The EVGR encodes LT, a spliced form of LT called 57kT, and small T antigen (ST). In addition, ALTO is encoded in an alternative open reading frame from LT (60). The LVGR encodes the viral coat proteins VP1 and VP2 as well as a microRNA that regulates T antigen levels. The polyomavirus virion is composed of 72 pentamers of VP1, with each pentamer lined by one molecule of VP2 on the inner surface (61). When expressed in bacteria or yeast, VP1 will spontaneously form pentamers and assemble into virus-like particles that can serve as a useful capture antigen to detect antibodies in serum indicative of prior infection (62, 63). Based on the VP1 serology assay, infection with MCPyV occurs as early as several months of age and increases in frequency until adulthood, when 70–90% of all adults show evidence for persistent infection (64–66).

While antibodies to MCPyV VP1 are widespread in the general population, antibodies to LT and ST are present in less than 1% of healthy individuals. In contrast, antibodies to MCPyV ST and LT can be detected in at least half of patients with MCCP (67). When present, antibody titers to the T antigens can decrease upon successful treatment of MCCP and can be used as a biomarker to follow disease status (67). Of note, MCC patients often have higher titers of antibodies to VP1 than do normal healthy individuals, although the significance of this observation is unclear (68).
**VIRAL INTEGRATION**

The MCPyV viral genome becomes highly mutated when it becomes clonally integrated in MCCP (37). In MCCP tumors, the NCCR and at least part of the EVGR region are retained, including an intact ST and a truncated form of LT (Figure 4a) (43). It can be assumed that the NCCR is conserved to promote expression of LT and ST, although additional functions of the NCCR may also contribute to oncogenesis. The N termini of LT and ST are shared and encode for a DnaJ or J domain. The J domain is usually wild-type or contains very few mutations in MCCP. The unique region of ST is also near wild-type in most MCCP tumors. In contrast, the second exon of LT contains many mutations, including point substitutions, deletions, and frameshifts that...
It should be noted that viral integration is not a normal phase of the MCPyV life cycle. The integrated viral DNA can no longer produce any viable virus. The integration of viral DNA can be viewed as a random genetic accident. However, when the combination of an intact NCCR, ST, and a truncated form of LT integration occurs in the appropriate cell type that permits expression of the T antigens, an MCC tumor emerges.

Integrated MCPyV DNA in the tumor genome can exist as multiple copies with the same or near-identical mutations in each copy of the viral genome. This observation implies that the viral DNA was already mutated prior to its insertion into the tumor DNA. In addition to amplification of the viral genome, some MCCP tumors show that the surrounding cellular DNA was coamplified with the virus (42, 43, 69). These observations have led to a model where the mutated viral DNA integrates into the tumor genome, generating a circular DNA form that is subsequently amplified by rolling circle amplification before resolving the amplified insertion back into the host genome (Figure 4b). Additional studies are needed to provide a more accurate molecular description of the amplification and insertion process. Clarification of this viral mechanism could also provide insight into the process of gene amplification observed for cellular oncogenes.

### MCPyV LARGE T ANTIGEN

The wild-type full-length LT encodes a protein of 817 residues. An alternatively spliced form of LT, 57kT, has an in-frame deletion of the central region that deletes most of the DNA origin binding and helicase domains but retains the C-terminal 100 residues of full-length LT (70, 71). The function of 57kT is not known, although the C-terminal 100 residues have growth-suppressing activities (71, 72).

MCPyV LT binds specifically to the retinoblastoma tumor suppressor protein RB to inactivate RB and thereby activate E2F cell cycle–regulated genes (45, 73, 74). The LXCXE motif is responsible for binding to RB and is an essential component of the transforming activity of MCPyV LT. An interesting report demonstrated that an MCCP tumor and a derivative cell line, LoKe, contained a deletion in the RB1 gene and that continued LT expression was not required for proliferation of this cell line (45). This result indicates that LT binding and inactivation of RB are required to maintain proliferation of this MCCP tumor. Other LT activities may also contribute to MCCP oncogenesis but are not essential for the maintenance of the tumor cells.

Unlike LT from other polyomaviruses such as SV40, MCPyV LT does not bind or inactivate p53 (71, 72). Instead, MCPyV LT, through its association with RB, activates p53 (75). LT may also activate p53 indirectly through its association with USP7 (76). USP7 (also known as HUA SP) normally functions as a deubiquitinating enzyme. An important substrate of USP7 is MDM2, and deubiquitination of MDM2 by USP7 leads to increased levels of MDM2, which can in turn bind to p53 and decrease p53 levels. Truncated LT in MCCP may inhibit USP7 activity or affect its ability to deubiquitinate MDM2. It has been reported that USP7 increases the binding of full-length LT to the viral origin of replication and reduces viral DNA replication (76). In contrast, BRD4 binding to full-length LT has been reported to promote viral replication (77).

LT can bind specifically to VPS39 (also known as Vam6) (78, 79). LT binding to VPS39 sequesters it from involvement in lysosomal trafficking, although the significance of this activity in MCC or in MCPyV replication is not known (79). An LT point substitution, W209A, that disrupts binding to VPS39 has wild-type transforming potential (74). Phosphorylation of LT at residues S220 and S239 has been reported (80). Point substitution of these serine residues leads to increased levels of LT and may represent binding sites to FBW7, although this observation has not been confirmed (81).
Figure 5
Large T antigen (LT) appears in the nucleus and cytoplasm. Immunohistochemistry staining for Merkel cell polyomavirus LT with CM2B4 antibody is shown with (a) predominantly nuclear and (b) both nuclear and cytoplasmic staining. Figure adapted from Reference 75.

All MCC tumors reported to date express a truncated form of LT due to mutations to the integrated viral genome (70, 82). In MCCP tumors, the first exon of LT, encoding a DnaJ or J domain, is usually intact with very few point substitutions (Figure 4) (43). In contrast, almost immediately after the start of the second exon of LT, MCCP tumors contain point substitutions and deletions. The mutations spare the LXCXE (RB-binding) motif, but many other residues surrounding the LXCXE motif have been reported to be mutated. The LXCXE motif is surrounded by the MCPyV-unique regions MUR-1 and MUR-2, but they are unlikely to contribute to oncogenesis because deletions or point substitutions have been identified in several different MCC tumors and cell lines (43, 74). The nuclear localization sequence follows the MUR-2 domain, is retained in only about half of MCCP tumors, and is not required for cellular transformation (74). The truncated forms of LT that lose the nuclear localization sequence can be detected in the cytoplasm and nucleus and are apparently able to enter the nucleus because of their smaller size (Figure 5). The LT origin binding and helicase domains are necessary for normal polyomavirus replication but are typically absent in MCCP. Disruption of the origin binding and helicase activities eliminates the possibility that LT could bind to the integrated viral origin of replication within the NCCR and initiate the process of replicating viral DNA replication that leads to a DNA damage response (70, 83, 84). These LT truncating mutations will also reduce the number of potential viral antigens that could trigger an antitumor immune response. In addition, growth-suppressing activities in the C terminus of LT are lost with these deletions (71, 72).

ALTO is typically truncated by mutation or not expressed in MCC tumors. The C terminus of ALTO has a hydrophobic region required for binding to cellular membranes (60). Loss of a similar C-terminal hydrophobic region in mouse polyomavirus middle T antigen (MT) disables the transforming activity of MT (85). Given the similarity of ALTO to MT, it is unlikely that the truncated ALTO contributes to the transformed phenotype in MCCP.

MCPyV SMALL T ANTIGEN

The MCPyV ST is an essential contributor of MCPyV transforming activity (86, 87). While LT is heavily mutated in MCCP tumors, ST is typically wild-type and intact. Wild-type MCPyV ST contains 186 residues and shares the N-terminal 79 residues with LT. The remaining C-terminal...
The ST-MYCL-EP400 complex activates downstream target genes. (a) The ST-MYCL-EP400 complex activates MDM2, MDM4, and CK1α, which assemble into a ubiquitin ligase that inhibits p53. The ST-MYCL-EP400 complex has additional target genes not shown (75, 92). (b) Additional ST-MYCL-EP400 complex downstream target genes include components of the LSD1-RCOR2-INSM1 complex that form a transcriptional repressor complex that opposes the ATOH1 transcription factor and the ncBAF complex (75, 92, 109). Abbreviations: ncBAF, noncanonical BAF complex; ST, small T antigen. Panel a adapted from References 75 and 92; panel b adapted from Reference 109.

The ST in all mammalian polyomaviruses binds to the protein phosphatase 2A (PP2A) complex. Mammalian PP2A consists of at least three subunits. The PP2A scaffold A subunit forms a horseshoe-like structure containing multiple HEAT domains that recruit the regulatory B and catalytic C subunits (88, 89). Polyomavirus ST typically displaces the PP2A B subunit and forms a trimeric complex with the A and C subunits (90). The PP2A scaffold subunit Aα (PPP2R1A) form is more abundant than the Aβ form (PPP2R1B). There are two forms of the PP2A catalytic C subunit (PPP2CA and PPP2CB). MCPyV ST can bind to both forms of the A and C subunits (91, 92). ST binding to PP2A contributes to the transforming activities of SV40 and mouse polyomavirus ST (93–95). However, it is not clear if MCPyV ST binding to PP2A contributes to its transforming potential (91).

MCPyV ST forms a complex with the MYC paralog MYCL (L-MYC) and its heterodimeric partner MAX (92). In addition, ST and MYCL/MAX recruit the EP400 (p400) chromatin remodeling complex (Figure 6a). The EP400 complex is made up of at least 15 unique proteins, including EP400 (p400), TTRAP, KAT5 (TIP60), ACTL6A (BAF53A), RUVBL1/RUVBL2 (TIP49/TIP48), MEAF6, MRGBP, YEATS4 (GAS41), MORF4L1/MORF4L2 (MRG15/MRGX), DMAP1, BRD8, VPS72 (YL1), EPC1/EPC2, MBTD1, and ING3. The EP400 complex has also been referred to as the p400, Tip60, BAF53, or TRRAP complex in
mammalian cells and the NuA4 complex in yeast (96–99). The EP400 complex has both histone acetylation and chromatin remodeling activities that participate in transcription and DNA damage responses.

While MYC, MYCN, and MYCL have each been reported to bind to the EP400 complex, or at least various components of the EP400 complex, MCPyV ST is notable for its apparent ability to increase the stability of the interaction of the complete 15-protein EP400 complex with MYCL (92). Although MYC, MYCN, and MYCL have been reported to bind several additional cellular factors, including Host cell factor 1 (HCF-1 or HCFC1) and WDR5, it is not known if MCPyV ST can also bind these other factors together with MYCL and MAX (99).

**MCPyV ST-MYCL-EP400 COMPLEX ACTIVATES KEY DOWNSTREAM TARGET GENES**

The ST-MYCL-EP400 complex binds together to the transcriptional start sites of several hundred genes and functions to activate their expression. These ST-MYCL-EP400 complex–activated downstream target genes contribute to MCPyV oncogenesis. Chromatin immunoprecipitation of ST, MAX, and EP400 followed by next-generation sequencing (ChIP-Seq) from MCCP cell lines identified similar and overlapping specific DNA binding sites that were predominantly located near the transcription start sites of several hundred genes (92). Notably, there was specific enrichment of binding of all three factors to the E-Box (CACGTG) or canonical MYC binding sites in gene promoters. Depletion of EP400, MYCL, or ST by RNA interference (RNAi) led to significantly decreased levels of genes whose promoters were bound by the ST-MYCL-EP400 complex. Identification of the ST-MYCL-EP400 target genes revealed a large number of known MYC target genes involved in ribosomal biogenesis, splicing, glycolysis, and other basic metabolic functions. These observations are consistent with a role for MYCL and the MAX heterodimer functioning similar to the MYC paralog (98, 99).

An additional set of ST-MYCL-EP400 complex target genes involves regulation of p53 activity. As indicated above, most MCCP tumors contain the wild-type TP53 gene. Furthermore, expression of MCPyV LT can activate p53 at least in part through inactivation of the RB tumor suppressor protein (75). An important question is whether MCPyV can also reduce p53 activity in MCCP tumors containing wild-type TP53. This was addressed by recognizing that the ST-MYCL-EP400 complex increased levels of MDM2, an E3 ubiquitin ligase, which specifically binds p53 and promotes its ubiquitination and subsequent degradation by the proteasome (Figure 6a). Perhaps similar to MYCL in MCCP, MDM2 had been previously recognized as a MYCN target gene functioning to inhibit p53 activity in neuroblastoma (100–102). In addition to MDM2, the ST-MYCL-EP400 complex increases levels of MDM4 and CK1α (CSNK1A1), which cooperate to activate the E3 ubiquitin ligase activity of MDM2 (75). It is likely that additional ST-MYCL-EP400 complex target genes cooperate with MDM2 to promote the degradation of p53 and its loss of function. Importantly, the ability of the ST-MYCL-EP400 complex to transactivate levels of MDM2 implies that inhibitors of MDM2 could prove to be effective at activating p53 in MCCP containing wild-type p53.

Expression of ST can increase levels of several proteins specifically involved in cell motility. Overexpression of MCPyV ST in HEK293 cells led to increased levels of proteins involved in microtubule destabilization, including CDC42, CFL1, CTNN, and RHOA, which lead to a motile and migratory phenotype. Of note, there is considerable overlap between the ST-MYCL-EP400 complex target genes and the proteomic analysis of ST expression in HEK293 cells, which include CDK2, CORO1C, CTNNA1, KPNA3, MAPT, MTPAP, RHOA, and PFN1 (92, 103,
104). It is possible that these ST-MYCL-EP400 target genes contribute to the highly metastatic potential of MCC.

**LSD1, RCOR2, AND INSM1 OPPOSE ATOH1**

While the ST-MYCL-EP400 complex functions as a transcriptional activator, it was recognized that when ST, MYCL, and EP400 were depleted by RNAi in MCCP cell lines, gene expression levels for a substantial number of genes increased. These genes were not directly repressed by the ST-MYCL-EP400 complex. Instead, it was demonstrated that the ST-MYCL-EP400 complex could transactivate several components of the lysine-specific demethylase 1 (LSD1) repressor complex (Figure 6b). LSD1 functions to remove activating H3K4me2 and H3K4me1 marks and thereby reduce transcriptional activity. LSD1 forms a complex with several proteins, including the CoREST factor RCOR2, HDAC1/2, and INSM1. INSM1 is a member of the SNAG domain protein family that includes Snail (SNAI1), Slug (SNAI2), Scratch (SCRT1, SCRT2), GFI1, GFI1B, OVOL1, and OVOL2 (105). Each of the SNAG domain–containing proteins contains a highly conserved SNAG motif at the N terminus that can become methylated and bind directly to LSD1 (106, 107).

Certain inhibitors target the LSD1 demethylating activity and lead to persistence of the methylated histone mark in treated cells. In addition, several groups have reported that some LSD1 inhibitors disrupt the interaction between LSD1 and SNAG domain–containing proteins (107, 108). This can lead to decreased DNA binding of the LSD1 complex, with loss of its repressive activity and increased levels of target genes. In MCC, LSD1 inhibitors disrupt binding of LSD1 to INSM1, which destabilizes the complex and leads to decreased DNA binding by LSD1 and RCOR2 (109). Decreased binding results in loss of repression by the LSD1 complex with corresponding increased levels of target genes.

Identification of LSD1-RCOR2 DNA binding sites by ChIP-Seq revealed an enrichment for ATOH1 binding sites in MCC cell lines. Further testing demonstrated that ATOH1 competed with the LSD1-RCOR2 complex for binding to promoters of ATOH1-dependent genes (Figure 6b). This result implies that the LSD1-RCOR2-INSM1 complex functions at least in part to repress ATOH1 transcriptional activity.

A genome-wide CRISPR-Cas9 screen was performed to identify genes that, when lost, enable MCC cell lines to survive in the presence of LSD1 inhibitors. As expected, loss of KMT2D or KMT2C, lysine methyl transferases for H3K4me2 and H3K4me1, allowed cells to tolerate LSD1 inhibitors (109). An unexpected result was that loss of several components of the noncanonical BAF (ncBAF) complex could also cause resistance to LSD1 inhibitors. The ncBAF complex functions to open chromatin and allow gene expression. In this context, the ncBAF complex may cooperate with ATOH1 to promote neuroendocrine differentiation gene expression (Figure 6b) (109). These recent results imply that the LSD1-RCOR2-INSM1 complex functions in part to oppose ATOH1 and ncBAF-dependent gene expression. In this manner, it appears that a major function of the LSD1 complex in MCC is to repress ATOH1–driven expression of neural specific genes. Whether ATOH1 functions to directly recruit the ncBAF complex is not known.

Two recent reports support a role for MCPyV T antigens in modulating ATOH1 activity. RNAi–mediated depletion of the T antigens in MCCP cell lines induced a neuron-like differentiation pattern with increased levels of neural-related genes and neurite outgrowths capable of supporting sodium–dependent action potentials (5). T antigen knockdown reduced levels of SOX2 and ATOH1. Conversely, T antigen overexpression in MCCN cell lines or fibroblasts led to increased levels of SOX2 and ATOH1. Furthermore, overexpression of ATOH1 in MCCN cell lines...
changed their growth pattern from adherent cells to suspension cells that reflect a neuroendocrine growth pattern similar to classical MCCP cell lines (110).

**MOLECULAR FEATURES COMMON TO MCCP AND MCCN**

Analysis of how MCPyV perturbs cellular function and contributes to oncogenesis in MCCP reveals that at least some of the same signaling pathways are perturbed by mutations in MCCN. These pathways can be grouped into three distinct gene expression patterns that affect the cell cycle, MYC activity, and ATOH1 signaling (Figure 7).

The first set of genes perturbed in MCC comprises the cell cycle regulatory genes. A key feature of the cell cycle genes is regulation by the tumor suppressor proteins RB and p53. RB and p53 act as checkpoints when they respond to external and internal cellular stresses and reduce cell cycle gene expression. RB can inhibit cell cycle progression by binding to and repressing the E2F family of transcription factors. E2F transcription factors bind and activate the promoters of genes required for entry into S phase. The principal mechanism for p53 control of the cell cycle is through direct activation of p21 (CDKN1A), which functions to inhibit the activity of cyclin-dependent kinases CDK1 and CDK2 during the G1/S and G2/M phases of the cell cycle. In the absence of both RB and p53, cells can enter into S phase unfettered by any extrinsic checkpoint that controls Cyclin D–CDK4 or Cyclin E–CDK2 activity. Furthermore, in the absence of p53 and p21, there is reduced inhibition of Cyclin E–CDK2 and Cyclin A–CDK1/2 activity. MCCN tumors have a high frequency of loss-of-function mutations in the RB1 and TP53 genes (38, 40). Conversely, in MCCP tumors, MCPyV LT binds and inactivates the RB protein while the ST-MYCL-EP400 complex transactivates MDM2 to promote the degradation of p53. Whether the absence of RB and p53, with the resulting deregulation of cell cycle–dependent gene expression, represents a significant vulnerability for targeted therapy in both forms of MCC is not known.
The second set of genes deregulated in MCC comprises the MYC-dependent genes. As described above, MCPyV ST recruits MYCL to the EP400 complex to activate MYC-dependent gene expression. The ST-MYCL-EP400 complex promotes the expression of a variety of genes involved in MYC signaling pathways. MYC signaling is also activated in MCCN by amplification of the MYCL gene (43, 47). Although it is not clear whether amplification of MYCL functions equivalently to the ST-MYCL-EP400 complex, it is notable that amplification of MYC or MYCN is not typically observed in MCC (41). Given this observation, it is likely that MYCL provides at least some unique oncogenic activity in MCC that cannot be readily substituted with MYC or MYCN amplification.

The third set of genes deregulated in MCC is controlled by the ATOH1 transcription factor. While ATOH1 is expressed in both MCCP and MCCN, it is clear that its transcriptional activity is at least partially attenuated, since MCC cells do not become fully differentiated into Merkel or neural cells. Indeed, decreased levels of MCPyV T antigens or overexpression of ATOH1 led to a terminally differentiated neural phenotype (5, 110). As shown recently, ATOH1 transcriptional activity is opposed by the LSD1-RCOR2-INSM1 complex in MCCP tumors (109). While ATOH1 is expressed in MCCN, the levels may be lower than those observed in MCCP (109, 111). Furthermore, loss of KMT2D and KMT2C, lysine methyl transferases whose activities oppose LSD1, may result in reduced ATOH1 activity. Of note, inactivating mutations in KMT2C and KMT2D are among the most frequently mutated genes in MCCN tumors (41).

It is likely that there are additional signaling pathways that are perturbed in MCC. PI3K signaling is often activated in both MCCP and MCCN. Heterozygous loss of chromosome 10 occurs in more than 30% of MCCN and MCCP tumors and leads to reduced levels of PTEN, thereby increasing AKT activity. Activation of PIK3CA by point mutations is observed in MCCP and MCCN. Other signaling pathways that may be perturbed in both forms of MCC include Notch, Hedgehog, and bone morphogenetic protein signaling pathways (38, 109, 112).

PRC2 activity is required for the proper development of Merkel cells and may play an important role in MCC. Low levels of EZH2 expression in MCC tumors, as determined by IHC, correlate with an improved prognosis compared to tumors that have moderate or strong EZH2 expression (113). Recently, it was shown that PRC2 coordinated transcriptional silencing of the major histocompatibility complex class I (MHC-I) antigen processing pathway in an MCC cell line (114). MCC tumors and cell lines typically have low levels of MHC-I expression, including HLA-A, HLA-B, and HLA-C, which could contribute to immune evasion (115, 116). At least some MCC cells can increase their levels of MHC-I in response to interferons as well as HDAC1 and EZH2 inhibitors (114, 115, 117).

CELL OF ORIGIN

The cell of origin for MCC remains a critical question in the field. A better understanding of what is the original cell type that leads to the development of MCC could lead to improved models that are more reflective of the disease and support preclinical therapeutic trials. Identification of the cell of origin could provide insight into risk factors and preventive strategies needed to minimize the risk of developing MCC.

The high degree of UV-associated DNA mutations in MCCN points to the likelihood that its cell of origin is a sun-exposed skin cell. The high TMB burden due to UV mutations observed in MCCN is also found in melanoma, invasive cutaneous squamous cell carcinoma, and basal cell carcinomas. It is highly likely that MCCN tumors derive from an epithelial cell in the keratinocyte lineage. One opposing argument against a keratinocyte origin for MCCN is the typical presentation of MCC tumors in the dermal layer. However, there are several reports of the in
situ appearance of MCC associated with a cutaneous squamous cell carcinoma consistent with the possibility that MCCN can originate in the epidermal layer (34,118,119).

Several mouse models support a keratinocyte lineage for MCCP (120, 121). Notably, coexpression of MCPyV ST and ATOH1 using keratinocyte specific promoters (KRT5) led to the development of neuroendocrine Merkel-like tumor cells (122). Further advancement of this mouse model could be used to explore additional phenotypes, including metastasis, invasion, and immune evasion, as well as provide an opportunity to test checkpoint blockade therapy in combination with additional agents.

Since MCCP tumors do not have a UV mutational signature, it is unlikely that they will originate from a sun-exposed keratinocyte. It is possible that keratinocytes deep within a hair follicle may avoid extensive UV-induced damage to the genome and permit transformation by MCPyV. Perhaps another requirement for the MCCP cell of origin is that it should be capable of expressing MYCL and ATOH1, since MYCL is required for MCPyV ST oncogenic activity and ATOH1 is required for the neuroendocrine phenotype. It could be imagined that the highly specialized cell types associated with hair follicles, including progenitor cells, would have the necessary malleability to induce expression of MYCL and ATOH1 in the presence of MCPyV.

Trichoblastomas are benign tumors that arise from hair follicle cells in the skin and are typically removed to rule out basal cell carcinoma. Trichoblastomas contain somatic mutations and often harbor a large number of normal-appearing Merkel cells (123). An interesting report noted the presence of an MCCP tumor within a trichoblastoma. Identical somatic mutations were observed in the trichoblastoma component as well as in the MCCP tumor, which suggests that the MCCP arose within a preexisting trichoblastoma (124).

Since all examples of MCCP contain mutations that truncate LT and inactivate its ability to support viral DNA replication, it is likely that the cell of origin for MCCP can support viral replication. If the MCCP cell of origin was unable to support viral replication, then there would be less selective pressure to eliminate the viral replication potential by eliminating the origin binding and helicase domains of LT. It is plausible that expression of wild-type MCPyV LT may promote replication and amplification of the mutated viral genome and associated cellular genome as the initial oncogenic event in MCCP. Viral replication may be an early event in the pathogenesis of MCCP that contributes to amplification of the viral genome (43,69).

It is plausible that the cell of origin for MCCP derives from a nonepithelial cell type. Candidate cell types proposed include B lymphocytes due to PAX5 expression detected in some MCC tumors (125, 126). Alternatively, dermal fibroblasts have been suggested as a potential cell of origin due to the facility of MCPyV replication in this cell type (127). Both B lymphocytes and dermal fibroblasts could partially explain the typical presentation of MCC tumors in the dermal layer of the skin.

**SUMMARY POINTS**

1. Our understanding of Merkel cell carcinoma (MCC) has been informed by studies of the development and physiology of Merkel cells. The ATOH1 transcription factor is essential for the development of Merkel cells.

2. There are two forms of MCC with similar presentation and prognosis but completely different genetic causes. One form is caused by integration of Merkel cell polyomavirus (MCPyV) with persistent expression of the viral T antigens. The nonviral form of MCC is caused by extensive UV-induced mutations.
3. The viral oncogenes are the major contributors to the pathogenesis of polyomavirus-associated MCC (MCCP). MCPyV LT inactivates RB, while MCPyV ST binds to MYCL and the EP400 transcription complex to activate downstream MYC target genes.

4. There are many downstream targets of the ST-MYCL-EP400 complex. These target genes include factors that inactivate p53 and others that contribute to cellular motility and invasion properties.

5. A key downstream target of the ST-MYCL-EP400 complex is the LSD1-RCOR2-INSM1 complex. The LSD1 complex functions as a lysine-specific demethylase that serves to repress genes. A relevant class of genes repressed by the LSD1 complex in MCC cells is driven by the ATOH1 transcription factor.

6. Mutations in oncogenes and tumor suppressor genes observed in nonviral MCC (MCCN) disrupt similar pathways that are targeted by the MCPyV T antigens in MCCP. These include the RB, p53, MYCL, and ATOH1 pathways.

FUTURE ISSUES

1. Downstream targets of the ST-MYCL-EP400 complex may represent opportunities for targeted therapy in MCCP tumors. Inhibitors of MDM2 and LSD1 may prove to be useful in the treatment of MCC. Additional downstream targets may also prove to be required for the survival of MCCP tumors.

2. The cell of origin for MCCP and MCCN remains incompletely described. It is assumed that MCCN requires a sun-exposed cell type to account for the UV mutational signature, while MCCP requires a non-sun-exposed cell type to account for the lack of a UV mutational signature. Additional requirements for the MCCP cell of origin include the requirement for expression of ATOH1 and MYCL.

3. The high response rate of MCCP and MCCN tumors to checkpoint blockade therapy with antibodies to PD-1 or PD-L1 remains unexplained. The high tumor mutational burden (TMB) in MCCN likely contributes to an increased number of neoantigens that represent targets for immune cell detection. The viral antigens in MCCP may similarly provoke a strong immune response to checkpoint blockade therapy. A better understanding of the specific mutations in MCCN and the activities of the MCPyV T antigens in MCCP may provide further insight into why these tumors respond so well and how to improve the response rate and durability of response.

4. Lessons learned by comparing the oncogenic activities in MCCN and MCCP can inform insights into other high-grade neuroendocrine carcinomas.

DISCLOSURE STATEMENT

J.A.D. received research funding from Constellation Pharmaceuticals. J.A.D. has served as a consultant to Merck & Co. and EMD Serono.
ACKNOWLEDGMENTS

This work was supported in part by US Public Health Service grants R35CA232128 and P01CA203655 and The Bridge Project, a partnership between the Koch Institute for Integrative Cancer Research at the Massachusetts Institute of Technology and the Dana-Farber/Harvard Cancer Center.

LITERATURE CITED


52. Anderson DA III, Murphy TL, Eisenman RN, Murphy KM. 2020. The MYCL and MXD1 transcription factors regulate the fitness of murine dendritic cells. PNAS 117:4885–93


68. Pastrana DV, Tolstov YL, Becker JC, Moore PS, Chang Y, Buck CB. 2009. Quantitation of human seroresponsiveness to Merkel cell polyomavirus. PLOS Pathog. 5:e1000578


80. Kwun HJ, Chang Y, Moore PS. 2017. Protein-mediated viral latency is a novel mechanism for Merkel cell polyomavirus persistence. PNAS 114:E4040–47


70. Characterization of truncated LT with loss of viral origin replication.


