

Distinct DNA Methylation Patterns of Subependymal Giant Cell Astrocytomas in Tuberous Sclerosis Complex

Anika Bongaarts

Amsterdam UMC - Locatie AMC: Amsterdam UMC Locatie AMC <https://orcid.org/0000-0003-1451-4240>

Caroline Mijnsbergen

Amsterdam UMC - Locatie AMC: Amsterdam UMC Locatie AMC

Jasper J. Anink

Amsterdam UMC - Locatie AMC: Amsterdam UMC Locatie AMC

Floor E. Jansen

UMC Utrecht: Universitair Medisch Centrum Utrecht

Wim G.M. Spliet

UMC Utrecht: Universitair Medisch Centrum Utrecht

Wilfred F.A. den Dunnen

UMCG: Universitair Medisch Centrum Groningen

Roland Coras

Erlangen University Hospital: Universitätsklinikum Erlangen

Ingmar Blümcke

Erlangen University Hospital: Universitätsklinikum Erlangen

Werner Paulus

Universitätsklinikum Münster: Universitätsklinikum Munster

Victoria E. Gruber

Medical University of Vienna: Medizinische Universität Wien

Theresa Scholl

Medical University of Vienna: Medizinische Universität Wien

Johannes A. Hainfellner

Medical University of Vienna: Medizinische Universität Wien

Martha Feucht

Medical University of Vienna: Medizinische Universität Wien

Katarzyna Kotulska

Children's Memorial Health Institute: Instytut Pomnik -Centrum Zdrowia Dziecka

Sergiusz Jozwiak

Children's Memorial Health Institute: Instytut Pomnik -Centrum Zdrowia Dziecka

Wieslawa Grajkowska

Children's Memorial Health Institute: Instytut Pomnik -Centrum Zdrowia Dziecka

Anna Maria Buccoliero

Children Hospital Meyer: Azienda Ospedaliero Universitaria Meyer

Chiara Caporalini

Children Hospital Meyer: Azienda Ospedaliero Universitaria Meyer

Flavio Giordano

Children Hospital Meyer: Azienda Ospedaliero Universitaria Meyer

Lorenzo Genitori

Children Hospital Meyer: Azienda Ospedaliero Universitaria Meyer

Figen Söylemezoğlu

Hacettepe University: Hacettepe Universitesi

José Pimentel

Hospital de Santa Maria

David T.W. Jones

DKFZ: Deutsches Krebsforschungszentrum

Brendon P. Scicluna

Amsterdam UMC - Locatie AMC: Amsterdam UMC Locatie AMC

Antoinette Y.N. Schouten-van Meeteren

Prinses Máxima Centrum voor Kinderoncologie: Prinses Maxima Centrum voor Kinderoncologie

Angelika Mühlebner

Amsterdam UMC - Locatie AMC: Amsterdam UMC Locatie AMC

James D. Mills (✉ j.d.mills@amsterdamumc.nl)

Amsterdam UMC - Locatie AMC: Amsterdam UMC Locatie AMC

Eleonora Aronica

Amsterdam UMC - Locatie AMC: Amsterdam UMC Locatie AMC

Research Article

Keywords: SEGA, TSC, methylation, RNA-sequencing, low-grade glioma

Posted Date: June 16th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-586356/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Tuberous sclerosis complex (TSC) is a monogenic disorder caused by mutations in either the *TSC1* or *TSC2* gene, two key regulators of the mechanistic target of the rapamycin complex pathway. Phenotypically, this leads to growth and formation of hamartomas in several organs, including the brain. Subependymal giant cell astrocytomas (SEGAs) are low-grade brain tumors commonly associated with TSC. Recently, gene expression studies provided evidence that the immune system, the MAPK pathway and extracellular matrix organization play an important role in SEGA development. However, the precise mechanisms behind the gene expression changes in SEGA are still largely unknown, providing a potential role for DNA methylation. We investigated the methylation profile of SEGAs using the Illumina Infinium HumanMethylation450 BeadChip (SEGAs n = 42, periventricular control n = 8). The SEGA methylation profile was enriched for the adaptive immune system, T cell activation, leukocyte mediated immunity, extracellular structure organization and the ERK1 & ERK2 cascade. More interestingly, we identified two subgroups in the SEGA methylation data and show that the differentially expressed genes between the two subgroups are related to the MAPK cascade and adaptive immune response. Overall, this study shows that the immune system, the MAPK pathway and extracellular matrix organization are also affected on DNA methylation level, suggesting that therapeutic intervention on DNA level could be useful for these specific pathways in SEGA. Moreover, we identified two subgroups in SEGA that seem to be driven by changes in the adaptive immune response and MAPK pathway and could potentially hold predictive information on target treatment response.

Background

Tuberous sclerosis complex (TSC) is a multisystem monogenic disorder caused by mutations in either *TSC1* or *TSC2* and is characterized by hamartoma development in several organs, including the brain, kidneys, lungs, heart, eyes, and skin (Curatolo, Bombardieri, and Jozwiak 2008). Patients with TSC often have neurological manifestations including neurodevelopmental disorders (such as autism) and severe epilepsy (Curatolo, Moavero, and de Vries 2015). The majority of patients with TSC have seizure onset before the age of two (Davis et al. 2017). The hallmark brain lesions in TSC include cortical/subcortical tubers, subependymal nodules (SENs) and subependymal giant cell astrocytomas (SEGAs) (Aronica, Becker, and Spreafico 2012, Aronica and Crino 2014). SEGAs are benign, slow growing tumors classified as WHO grade I making up for 1–2% of all paediatric brain tumors (Jozwiak, Mander, and Mlynarski 2015, Louis et al. 2016). Usually, SEGAs develop during the first 2 decades of life of patients with TSC, with a mean age at presentation below 18 years (Jozwiak, Mander, and Mlynarski 2015, Adriaensen et al. 2009). They are typically located near the foramen of Monro where extended growth of the tumor can result in blockage of the cerebral fluid circulation and subsequent obstructive hydrocephalus (Cuccia et al. 2003). SEGAs are thought to arise from subependymal nodules (SEN) along the ependymal lining of the lateral ventricles (Buccoliero et al. 2009, Fujiwara et al. 1989, Morimoto and Mogami 1986). Histologically, they are characterized by spindle cells, gemistocytic-like cells and giant cells and demonstrate an immature neuroglial phenotype.

Tumor suppressors hamartin (*TSC1*) and tuberin (*TSC2*) can form an intracellular complex with TBC1 Domain Family Member 7 (TBC1D7) that exerts GTPase-activating protein (GAP) activity towards the small GTPase Ras homolog enriched in brain 1 (RHEB1) (Dibble et al. 2012, Inoki et al. 2003). Inhibition of RHEB1 is important in regulating the mechanistic target of rapamycin complex (mTOR) pathway, which can affect cell growth and proliferation. Pathogenic loss of function mutations in *TSC1* or *TSC2* result in constitutive activation of the mTOR pathway and uncontrolled cell cycle progression (Chan et al. 2004). Besides the mTOR pathway, the immune system, the Mitogen-activated protein kinase (MAPK) pathway and extracellular matrix (ECM) organization have been suggested to play a role in SEGA development based on gene expression studies (Bongaarts et al. 2019, Martin et al. 2017, Tyburczy et al. 2010). However, the precise mechanisms behind these gene expression changes in SEGA are still largely unknown.

Gene expression can be controlled through regulation of the epigenome, via epigenetic mechanisms (Keshet, Yisraeli, and Cedar 1985). DNA methylation is one of most recognized epigenetic markers and is generally associated with silencing of gene expression, and its role in tumorigenesis has become a topic of interest (Klutstein et al. 2016). It is characterized by the addition of a methyl or hydroxymethyl by DNA methyltransferases (DNMTs) to cytosine residues in CG (CpG sites), CXG and CXX DNA sequences (where X corresponds to A, T, or C). Changes in DNA methylation have been well studied in cancer including central nervous system (CNS) tumors (Binder et al. 2019, Jones and Baylin 2007) and profound changes of methylation profiles have also been seen in neuro-psychiatric diseases such as autism spectrum disorders, epilepsy and TSC (Martin et al. 2017, Henshall and Kobow 2015, Gos 2013). Furthermore, DNA methylation profiling is highly robust and reproducible and has therefore been successfully used to distinguish subtypes in CNS tumors and focal cortical dysplasia (Laffaire et al. 2011, Kobow et al. 2019, Capper, Jones, et al. 2018, Stone et al. 2018). These DNA methylation-based classifications of CNS tumors have proven helpful for better diagnostics especially in cases with ambiguous histology or contradictory molecular profiles. Although, SEGAs have been included in previous methylation based studies, none of these studies have gone into depth on the molecular contribution of DNA methylation in SEGAs specifically (Martin et al. 2017, Capper, Jones, et al. 2018, Capper, Stichel, et al. 2018). Therefore, in this study we aimed to identify distinct methylation patterns and pathways that might contribute to SEGA pathogenesis by performing a comprehensive analysis of genomic DNA methylation in SEGAs.

Results

The methylation profile of SEGAs

To characterize the methylation profile of SEGAs, DNA was extracted from SEGA samples and control brain samples and analyzed using the 450k methylation array. In total 42 SEGA samples were included from 39 TSC patients and 3 patients with no other manifestations of TSC (surgical specimens) and 8 location-matched periventricular controls (autopsy specimens; see materials and methods and Table 1). After filtering for probes with a detection p-values of more than 0.01, located on the sex chromosomes, or on SNPs as well as cross-hybridization probes a total of 421,352 CpGs were analyzed with a principal

component analysis (PCA) indicating that the major source of variability was the diagnosis (SEGA or control; Fig. 1a), which was confirmed with a spearman's correlation matrix on the top 5% most variable CpGs (Fig. 1b). Furthermore, no specific clustering was seen based on the TSC mutation (Fig. 1b). To assess other potential confounders on the methylation profile a principal variance component analysis (PVCA) was performed, showing that the major contributor to the variance between the samples was the diagnosis (Supplementary Fig. 1a). The majority of the filtered CpGs were located outside of the TSS-associated regions of genes either on the body region (36.67%), the IGR (23.18%) or the 3'UTR (3.93%; Fig. 1c). Within the TSS-associated regions 13.55% was located on the TSS200 region of the gene, 17.71% on the TSS1500 region, 8.4% on exon 1 and 13.94% on the 5'UTR. After selecting for an adjusted p-value < 0.01 and a β -value difference of > 0.2 this distribution shifted towards the TSS-associated regions (20.93% TSS200, 42.55% TSS1500, 13.78% exon 1 and 18.85% 5'UTR), whereas 41.21% was located on the body region, 1.74% on the 3'UTR and no CpGs were located within the IGR (Fig. 1c).

Since most of the differentially methylated CpGs were located at the TSS-associated regions (TSS200, TSS1500, exon 1 and 5'UTR), we narrowed our data set to these CpGs and found 4616 CpGs hypomethylated and 2526 hypermethylated in SEGA compared to control (adjusted p-value 0.01, β -value difference of > 0.2, TSS-associated regions; Fig. 1d). The 7142 differentially methylated CpGs were located on 3875 genes. We identified 227 enriched GO terms (adjusted p-value < 0.05) for these genes (Fig. 1e; Table 2) including adaptive immune system, T cell activation, leukocyte mediated immunity, extracellular structure organization and the ERK1 & ERK2 cascade. The majority of the enriched pathways contained more hypomethylated genes than hypermethylated genes (Fig. 1f; Table 2).

Methylation of the mTOR pathway in SEGAs

In order to compare the *TSC1* mutated SEGA samples with the *TSC2* mutated SEGA samples, three differential analyses were carried out: *TSC1* mutated SEGAs compared to control (TSC1-control) and *TSC2* mutated SEGAs compared to control (TSC2-control) and *TSC1* mutated SEGAs compared to *TSC2* mutated SEGAs (TSC1-TSC2). We identified 4008 hypomethylated and 2111 hypermethylated CpGs in TSC1-control (Fig. 2a), 4721 hypomethylated and 2345 hypermethylated CpGs in TSC2-control (Fig. 2b) and no CpGs differentially methylated in TSC1-TSC2 (adjusted p-value 0.01, β -value difference of > 0.2, TSS-associated regions). Furthermore, the majority of the differentially methylated CpGs in TSC1-control and TSC2-control were overlapping (5293 CpGs), whereas 1773 CpGs were only found differentially methylated in TSC1-control and 826 CpGs in TSC2-control (Fig. 2c).

We further evaluated the methylation of mTOR pathway related genes by extracting CpGs that were located on genes from the Reactome based mTOR pathway or mTORC1 signaling pathway. A total of 459 CpGs were located on 35 mTOR pathway related genes of which 18 CpGs located on 6 genes (2/17 CpGs on *AKT2*, 1/17 on *EEF2K*, 4/21 on *EIF4E*, 4/33 on *EIF4G1*, 1/19 on *MTOR* and 6/43 on *PRKAG2*) were differentially methylated (Fig. 2d). The majority of these CpGs were hypomethylated (15/18), whereas 3 CpGs located on *PRKAG2* were hypermethylated.

Expression of inflammation, mTOR activation, glial and neuronal markers in SEGAs

SEGAs are considered mixed glio-neuronal tumors, with mTOR activity and presence of inflammation markers. Therefore, we wanted to evaluate the commonalities and differences in the expression of CD3, HLA-DP/DQ/DR, GFAP, MAP2 and pS6 in 42 SEGAs and 8 location-matched controls. In periventricular control tissue CD3, MAP2 and pS6 were not detected, whereas a moderate expression of HLA-DP/DQ/DR and high expression of GFAP was seen (Fig. 3a). In SEGA we found several positive CD3 cells and observed an overall increase in positive area for CD3 in SEGA compared to control tissue (Fig. 3b; $p < 0.0001$). HLA-DP/DQ/DR, GFAP, MAP2 and pS6 were expressed in a heterogeneous manner in SEGAs (Fig. 3a). The percentage of positive area of HLA-DP/DQ/DR ($p < 0.0001$), MAP2 ($p = 0.0114$) and pS6 ($p < 0.0001$) were increased in SEGA compared to control, whereas the positive area for GFAP was decreased in SEGA ($p = 0.0016$). Spearman's rank correlation revealed a weak positive correlation between the expression of CD3 and HLA-DP/DQ/DR ($r = 0.347$; $p = 0.026$), pS6 and HLA-DP/DQ/DR ($r = 0.368$; $p = 0.016$), and GFAP and HLA-DP/DQ/DR ($r = 0.325$; $p = 0.036$) in SEGA. Spearman's rank correlation with clinical data revealed a weak positive correlation between age at surgery and CD3 ($r = 0.3197$; $p = 0.0416$) and a negative correlation between tumor size and CD3 ($r = -0.4331$; $p = 0.0118$), HLA-DP/DQ/DR ($r = -0.4370$; $p = 0.0098$), MAP2 ($r = -0.4746$; $p = 0.0046$) and pS6 ($r = -0.4884$; $p = 0.0034$).

Two distinct methylation groups in SEGAs

To evaluate potential subgroups within the SEGA samples the top 5% most variable CpGs were analysed with hierarchical clustering, consensus clustering and silhouette. Hierarchical clustering indicated 2 major groups with one group subdividing into two smaller groups, which could not be explained by age at SEGA surgery, gender or *TSC* mutation (Fig. 4a). This was confirmed by both consensus clustering (Fig. 4c-d) and silhouette (Fig. 4e-g), which indicated that $k = 3$ was most robust. To assess other potential confounders on the methylation profile another PVCA was performed, showing that the major variance between the SEGA samples matched with the identified subgroups $k = 3$, followed by the subgroups $k = 2$ (Supplementary Fig. 1b). Other clinical data including *TSC1/TSC2* mutation status, gender, localization of the SEGA, size of the tumor, epilepsy, age at seizure onset, seizure frequency, drug management at time of surgery (including treatment with mTORC1 inhibitors), and presence of other TSC related malformations as well as their various interactions contributed minimally to the overall variance seen amongst the samples.

We first wanted to further investigate the two largest groups identified and performed differential testing between group 1 compared to control (SEGA1-control), group 2 compared to control (SEGA2-control) and group 1 compared to group 2 (SEGA1-SEGA2). We found 4377 hypomethylated and 1411 hypermethylated CpGs in SEGA1-control (Fig. 5a), 4883 hypomethylated and 3132 hypermethylated CpGs in SEGA2-control (Fig. 5b) and 321 hypomethylated and 70 hypermethylated CpGs in SEGA1-SEGA2 (Fig. 5c; adjusted p-value 0.01, β -value difference of > 0.2 , TSS-associated regions). In order to identify differentially methylated genes that were unique for each group, genes corresponding to the differentially

methylated CpGs were extracted. Genes that were overlapping between SEGA1-control and SEGA1-SEGA2 and did not overlap with SEGA2-control were considered unique for SEGA1, whereas genes that overlapped between SEGA2-control and SEGA1-SEGA2 but did not overlap with SEGA1-control were considered unique for SEGA2 (70 SEGA1 unique genes and 58 SEGA2 unique genes; Fig. 5d). GO analysis revealed 15 GO terms enriched for these 128 unique genes, which were related mainly to the MAPK cascade and adaptive immune response (P-adjusted < 0.05; Fig. 5e). We further evaluated the RNA expression of these 128 genes and performed correlations between the normalized count matrix of genes that were expressed (106/128 genes) and the β -values of the corresponding CpGs. We identified 11 genes (*MFAP5*, *SLC15A2*, *GPT2*, *CLCC1*, *CYP4F12*, *KLHL33*, *C12orf29*, *SELENBP1*, *MTERFD1*, *ICAM2*, *CHD7*) that inversely correlated with their corresponding CpG (Table 3). Based on the RNA expression of the expressed genes no clear clustering was found between SEGA1 and SEGA2, although the number of cases in each group was relatively small (Fig. 5f).

Next, we investigated if the expression of CD3, HLA-DP/DQ/DR, GFAP, MAP2 or pS6 could explain the subgroups found in the methylation data. We found a higher positive area of CD3 in SEGA2 compared to SEGA1 ($p = 0.0068$; Fig. 6a boxplot). No difference was found between SEGA1 and SEGA2 for the other markers (Fig. 6b-6e; boxplots) or between the subgroups of SEGA2, SEGA2a and SEGA2b (Fig. 6a-e; boxplots). Using ROC analysis we identified CD3 positive area as the best predictor for dividing SEGA1 from SEGA2 (Fig. 6a; AUC = 0.825) compared to HLA-DP/DQ/DR (Fig. 6b; AUC = 0.561), GFAP (Fig. 6c; AUC = 0.450), MAP2 (Fig. 6d; AUC = 0.481) and pS6 (Fig. 6e; AUC = 0.539). Furthermore, using Random forest we found that none of the clinical data (*TSC1/TSC2* mutation status, gender, localization of the SEGA, age at seizure onset, seizure frequency, drug management at time of surgery, size of the tumor, tumor recurrence/regrowth and presence of other TSC related malformations) could properly separate between SEGA1 and SEGA2 (Fig. 6f) or between SEGA1, SEGA2a and SEGA2b (Fig. 6g). The tumor size contributed most to separating the two groups but showed no significant difference between SEGA1 and SEGA2 (Fig. 6h).

Discussion

In this study we performed DNA methylation profiling of SEGAs from TSC patients and showed that the differential methylation profile between SEGAs and control tissue is enriched for GO terms including the adaptive immune system, T cell activation, leukocyte mediated immunity, extracellular structure organization and the ERK1 & ERK2 cascade. Histological markers for T cells, microglia reactivity, mTOR activation and neurons were higher expressed, whereas the glial marker GFAP was lower expressed in SEGA compared to periventricular control tissue. Furthermore, we identified two robust subgroups in the DNA methylation of SEGA, with a distinct methylation profile of genes related to the adaptive immune response and the MAPK pathway. Moreover, we found differences in positivity of the T cell marker CD3 between the two largest subgroups.

Previous studies on SEGA methylation have shown that SEGAs are a unique entity among CNS tumors (Martin et al. 2017, Capper, Jones, et al. 2018, Capper, Stichel, et al. 2018). However, the molecular

mechanisms targeted by methylation changes in SEGA have not been well studied yet. In the present study, we identified substantial methylation changes in SEGAs compared to periventricular control tissue that appeared to be independent of the *TSC1/TSC2* mutation or other clinical information available. GO analysis showed an enrichment of the adaptive immune system, T cell activation, leukocyte mediated immunity, extracellular structure organization and the ERK1 & ERK2 cascade. Previous gene expression studies on SEGA found differential expression of similar pathways, indicating that these pathways are already affected on DNA level and might therefore be important drivers in SEGA pathogenesis (Bongaarts et al. 2019, Martin et al. 2017, Tyburczy et al. 2010). Differential expression of genes related to the immune system and the ECM organization has also been seen in cortical tubers (Martin et al. 2017, Mills et al. 2017). Furthermore, several studies have documented dysregulation of inflammation and ECM organization related pathways in cortical tubers, suggesting that these processes might be conserved across TSC pathology (Boer et al. 2010, Boer et al. 2008, Prabowo et al. 2013, Broekaart et al. 2019). Therefore, it would be of interest to investigate if DNA methylation changes related to these processes are also present in cortical tubers and other TSC lesions.

The role of mTOR pathway activation due to loss of function mutations in *TSC1/TSC2* in TSC is well established. Furthermore, loss of heterozygosity (LOH) of *TSC1* or *TSC2* has been reported in approximately 80% of SEGAs and has also been found in other TSC hamartomas (Martin et al. 2017, Bongaarts et al. 2017, Chan et al. 2004). *In vitro* experiments show that mTOR inhibitors can reduce cell size and cell proliferation of SEGA cells. Currently, mTOR inhibitors are amongst the treatment options for SEGA associated with TSC (Franz et al. 2015). Therefore, we aimed to investigate methylation changes on genes related to the mTOR signaling pathway. GO term analysis on the differentially methylated CpGs between SEGA and control tissue did not reveal the mTOR pathway as a principal target. Moreover, by directly investigating mTOR pathway related genes, we found only a small number of differentially methylated CpGs on 6/35 mTOR related genes, indicating that DNA methylation changes most likely do not contribute to the mTOR activation in SEGA. It has been suggested that *TSC1/TSC2* epigenetic silencing might contribute to tumor formation in TSC and could explain cases where the second hit mutation in *TSC1/TSC2* is not found (Jiang et al. 2005, Lesma et al. 2009). In accordance with previous study, we did not find evidence of epigenetic silencing of the promoter of *TSC1* or *TSC2* in SEGA (Martin et al. 2017).

Initial studies suggested an astrocytic nature of SEGAs, whereas more recent studies demonstrate a mixed glio-neuronal phenotype, with mTOR activity and presence of inflammation markers (Bongaarts et al. 2017, Boer et al. 2008, Chan et al. 2004, Buccoliero et al. 2009, Buccoliero et al. 2016). We evaluated the expression of CD3, HLA-DP/DQ/DR, GFAP, MAP2 and pS6 using whole slide scanning in 42 SEGAs and 8 location and age matched controls. In accordance with previous literature, we confirmed the presence of inflammation markers and mTOR pathway activation in SEGA compared to control tissue (Bongaarts et al. 2017, Boer et al. 2008, Chan et al. 2004, Buccoliero et al. 2009, Buccoliero et al. 2016). Previous research showed that the mTOR activation is mainly present in giant cells and not in spindle cells of SEGA, which could explain the variability between the SEGA samples seen in this study (Buccoliero et al. 2016). The mTOR pathway can also regulate inflammatory responses and a previous

study has shown that HLA-DR positive microglial cells were localized around giant cells in SEGA (Boer et al. 2008, Lim et al. 2003). In accordance, we found a positive correlation between the expression HLA-DP/DQ/DR and pS6. Furthermore, we found expression of both GFAP and MAP2 in SEGAs confirming a glio-neuronal nature of SEGAs. However, the percentage of positive GFAP area was lower in SEGA compared to control tissue. This lower expression could be explained by the diffuse staining of GFAP, which has been reported in prior immunohistochemical studies, indicating that some SEGA cells might lose their glial phenotype (Buccoliero et al. 2009, Buccoliero et al. 2016). It could be that GFAP negative SEGA cells have a more neuronal expression, however no negative correlation was found between the expression of GFAP and MAP2 in SEGA.

Interestingly, we identified two subgroups (SEGA1 and SEGA2) in the SEGA methylation data with one group subdividing further into two smaller groups. These two groups are distinct in the methylation of genes related to the adaptive immune response and the MAPK cascade. However, no correlation was found between the methylation and RNA expression of these specific genes. Due to the complexity of RNA expression regulation, the effect of methylation changes might not be directly reflected in the RNA expression data. Higher expression of the T cell marker CD3 was found in SEGA2 compared to SEGA1 confirming differences in the adaptive immune response between the two groups. It could be possible that these methylation subgroups reflect differences in inflammatory cell content. However, it must be noted that the differences are small and only detectible with quantification. Furthermore, CD3 was not found differentially methylated between SEGA1 and SEGA2, suggesting an indirect effect of methylation on CD3 expression. The precise role of T-cells in SEGAs is still unknown, it could indicate a more responsive immune response to tumor cells but we also know that, neuroinflammation can increase the expression and activity of MMPs, which are increased in SEGAs and can play a role in tumorigenesis (Bongaarts et al. 2020). The MAPK pathway has been shown to be an important pathway for SEGA growth and has been suggested as a novel target for therapy for patients with SEGA (Bongaarts et al. 2019, Tyburczy et al. 2010, Mi et al. 2009). Therefore, the differences in methylation of this pathway between the two groups found in this study is interesting and could potentially reflect how these tumors would respond to MAPK inhibitors. Since MAPK inhibitors are not routinely used for SEGA treatment this could not be evaluated within this study. Furthermore, none of the clinical data available could explain the subgroups found, including the TSC mutation. We did find any confounding effects, however we cannot exclude confounding of other parameters such as Body Mass Index (BMI). It is very well possible that these two groups do reflect other important clinical features that were not evaluated in this study, such as tumor progression, or any differences in the clinical phenotypes either related or unrelated to the TSC mutation. Moreover, since the majority of patients included were not treated with mTOR inhibitors, it cannot be excluded that the biological make up of these subgroups reflect potential response to mTOR inhibitors. Therefore, DNA methylation analysis on SEGAs in retrospective studies are highly needed in order to unravel the clinical relevance of these subgroups.

Conclusions

Overall, this study shows that the DNA methylation profile of SEGAs is enriched for the immune system, the MAPK pathway and the ECM organization, strengthening the importance of these pathways in SEGA development and suggests that therapeutic intervention on DNA level could be useful. Moreover, we identified two subgroups in SEGA that seem to be driven by changes in the adaptive immune response and MAPK pathway. Although the clinical relevance of these subgroups remains uncertain they could potentially reflect tumor progression or response to treatment other than anti-epileptic drugs and deserve further investigation.

Materials & Methods

SEGA tumor specimens

A total of 55 SEGA specimens were obtained from the following sites: the Amsterdam University Medical Center (location AMC), the University Medical Center Utrecht, University Medical Center Groningen, Medical University of Vienna, Children's Memorial Health Institute in Warsaw, Meyer Children's Hospital in Florence, University Hospital Erlangen, University Hospital Münster, Hacettepe University in Ankara, the University Hospital of Santa Maria (CHLUN) in Lisbon, and the North Bristol NHS Trust as part of the UK Brain Archive Information Network (BRAIN UK: Ref.: 16/002). Fifty-one of the SEGA samples were obtained from patients that met the clinical diagnostic criteria for TSC. Histological diagnosis was confirmed following the current WHO classification guidelines by two independent neuropathologists (Louis et al. 2016). *TSC1/TSC2* mutation analysis was performed in blood or tumor sample DNA (for samples with a sufficient amount of DNA) as part of routine clinical care or was detected using massively parallel sequencing as described previously (Table 1) (Bongaarts et al. 2017). The following clinical data was collected from medical records: *TSC1/TSC2* mutation status, gender, localization of the SEGA, size of the tumor, age at seizure onset, seizure frequency, drug management at time of surgery (including treatment with mTORC1 inhibitors), tumor recurrence/regrowth and presence of other TSC related malformations. The ethnicity of all patients was Caucasian. Periventricular brain tissue was used as control and was obtained from autopsy controls without a history of TSC, epilepsy, brain tumor or other neurological manifestations. Specimens were obtained and used in accordance with the Declaration of Helsinki and this study was approved by the Medical Ethics Committees of each institution.

DNA extraction & 450k methylation analysis

DNA was extracted from FFPE SEGA tumor samples (n = 42) and location matched controls (n = 8). Representative tumor regions were identified on hematoxylin & eosin sections for cases in which hemorrhages were prominent and were macrodissected from unstained sections (10- μ m-thick) to reach a tumor cell content of 70% or more. For controls a punch of 2 mm diameter and up to 3 mm depth was taken from the periventricular zone. DNA was extracted using BiOstic FFPE Tissue DNA Isolation kit (MO BIO) according to the manufacturer's instructions. 200–300 ng of DNA per sample was analysed using Illumina Infinium HumanMethylation450 BeadChip (450k) arrays according to the manufacturer's instructions at the Genomics and Proteomics Core Facility of the DKFZ (Heidelberg, Germany).

RNA isolation and RNA sequencing

Frozen tissue of 12 SEGAs was homogenized with Qiazol Lysis Reagent (Qiagen Benelux, Venlo, The Netherlands) and total RNA was isolated using the miRNeasy Mini kit (Qiagen Benelux, Venlo, the Netherlands) according to manufacturer's instructions. RNA concentration was determined using a Qubit® 2.0 Fluorometer (Life Technologies, Carlsbad, CA, USA) and the RNA integrity was assessed using a Bioanalyser 2100 (Agilent). Library preparation and sequencing were completed at GenomeScan (Leiden, the Netherlands). The Illumina (San Diego, California, USA) NEBNext Ultra Directional RNA Library preparation kit was used to prepare sequencing libraries in accordance to manufacturers guidelines. Clustering and DNA sequencing was performed using the Illumina cBot and HiSeq 4000 according to manufacturer's protocols. Each library was subjected to paired-end sequencing, producing reads of 150 nucleotides in length with a read-depth of 36 million reads.

Data was processed as previously described (Bongaarts et al. 2019). Briefly, sequence reads were trimmed and filtered using FastQC v0.11.5 (Babraham Institute, Babraham, Cambridgeshire, UK) and Trimmomatic v0.36 (Bolger, Lohse, and Usadel 2014). Paired-end reads were aligned to the human reference genome (GRCh38) with TopHat2 v2.0.13 and default settings (Kim et al. 2013). The number of reads that mapped to each gene, based on Gencode v25, was determined using featureCounts from the SubRead package (Liao, Smyth, and Shi 2014). The count matrix was normalized using the R package DESeq2 (Love, Huber, and Anders 2014).

Bioinformatic analysis

Raw IDAT files from the 450k were passed to the minfi package in R and quality control was performed. Samples that failed quality control were excluded from the analysis. Normalization included a Noob background correction and dye-correction based on the control probes using the function preprocessFunnorm from the R package minfi, which removes any between-array variation (Fortin et al. 2014). Probes with detection p-values of more than 0.01, located on the sex chromosomes, or in SNPs were removed as well as cross-hybridization probes. After these steps, beta (β)-values ranging from 0.0 to 1.0 from 421,352 probes were used for further analysis.

Using the ConsensusClusterPlus package (Wilkerson and Hayes 2010), consensus clustering was performed with h-clust average linkage to detect robust clusters, where the metric was 1 minus the Spearman's correlation coefficient. The procedure was run over 1000 iterations and with a sub-sampling ratio of 0.99. Additionally, we applied silhouette to identify robust clusters. Principal component analysis (PCA) was performed considering all CpG probes. Hierarchical clustering was performed on the top 5% most variable CpG probes using h-clust with average linkage. Principal variance component analysis (PVCA), Receiver Operating Characteristic (ROC) analysis and Random Forest were used to assess the contribution of histological or clinical data to the clusters identified. To determine the CpGs that were differentially methylated between groups (e.g. SEGA compared to control) a non-parametric Mann-Whitney U was used at each CpG probe. The distribution of CpGs on the gene (TSS200, TSS1500, 5'UTR and Exon 1, Intergenic region (IGR), 3'UTR or gene body) each CpG was evaluated by calculating the

percentage of CpGs per gene region. CpGs that were located on multiple genes or transcript variants were counted as one for each corresponding gene region. CpG probes located at the TSS-associated regions (TSS200, TSS1500, 5'UTR and Exon 1) with a benjamini-Hochberg adjusted p-value < 0.01 and a β -value difference of > 0.2 were considered differentially methylated. Gene ontology (GO) analysis was performed using the R package clusterProfiler (Yu et al. 2012). In order to identify mTOR pathway methylation changes the mTOR pathway or mTORC1 signaling pathway gene list from the Reactome database was extracted (Croft et al. 2011). CpGs mapping to these genes with a benjamini-Hochberg adjusted p-value < 0.01 and a β -value difference of > 0.2 were considered differentially methylated. Clustering of subgroups was re-evaluated in an additional independent SEGA cohort from Heidelberg (50 additional cases) using hierarchical clustering and t-distributed stochastic neighbor embedding (TSNE) plot.

Immunohistochemistry

Immunohistochemical staining was performed on 42 SEGAs and 8 controls with a Ventana semiautomated staining machine (Benchmark ULTRA; Ventana, Illkirch, France) and the Ventana DAB staining system according to the manufacturer's protocol. The following antibodies have been used: glial fibrillary acidic protein (GFAP; polyclonal rabbit, DAKO, Glostrup, Denmark; 1:4000; monoclonal mouse; DAKO; 1:50), microtubule-associated protein (MAP2; mouse clone HM2; Sigma 1:100), human leukocyte antigen class II (HLA-DP, DQ, DR; mouse clone CR3/43; DAKO; 1:100), cluster of differentiation 3 (CD3; mouse monoclonal, clone F7.2.38; DAKO; 1:200; T-lymphocytes), phospho-S6 ribosomal protein (pS6 Ser235/236; rabbit polyclonal, Cell Signaling Technology, Beverly, MA, USA; 1:50).

Evaluation of immunohistochemistry

Olympus dotSlide system (vs 2.5, Olympus, Tokyo, Japan) was used for whole slide scanning at a 20x magnification with a resolution of 0.32 μm /pixel. The scans were converted into TIFF files. The DAB positive area was separated from background using an adjusted protocol for segmentation and was corrected for total area of the tissue. For samples with blood contamination, regions of interest (ROI) were selected in order to select representative tumor regions. The overall percentage of positivity was assessed for each case and used for statistical analysis.

Statistical analysis

Statistical analysis was performed with GraphPad Prism software (Graphpad software Inc., La Jolla, CA) using the nonparametric Mann-Whitney U-test or, for multiple groups, the non-parametric Kruskal-Wallis test followed by Mann-Whitney U-test. Correlations were assessed with R using the Spearman's rank correlation test. An adjusted P-value < 0.05 was considered statistically significant except for the differentially methylation analysis where an adjusted P-value < 0.01 was considered statistically significant.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on request.

List Of Abbreviations

TSC - tuberous sclerosis complex; mTOR - the mechanistic target of the rapamycin complex; SEGA - subependymal giant cell astrocytomas; SEN - subependymal nodules; TBC1D7 - TBC1 Domain Family Member 7; GAP - GTPase-activating protein; RHEB1 - ras homolog enriched in brain 1; ECM - extracellular matrix; MAPK - mitogen-activated protein kinase; DNMTs - DNA methyltransferases; CNS - central nervous system; 450k - Illumina Infinium HumanMethylation450 BeadChip; PCA - principal component analysis; PVCA - principal variance component analysis; ROC - receiver operating characteristic; IGR - intergenic region; GO - gene ontology; GFAP - glial fibrillary acidic protein; MAP2 - microtubule-associated protein; HLA-DP, DQ, DR - human leukocyte antigen class II; CD3 - cluster of differentiation 3; pS6 - phospho-S6 ribosomal protein; ROI - regions of interest; LOH - loss of heterozygosity

Declarations

Acknowledgements

The authors thank all supporters of the TSC brain bank (Laboratory of Molecular and Cellular Neurobiology, International Institute of Molecular and Cell Biology, Warsaw, Poland: J. Jaworski, A Tempes; The Service d' Anatomie Pathologique, CHI de Creteil and Inserm U676, Hospital Robert Debre, Paris, France: H. Adle-Biassette; Department of Pediatrics, Division of Neuropathology and Neurochemistry, Department of Neurology, Department of Neurosurgery, Medical University Vienna, Austria: M. Feucht, T. Scholl, J. Hainfellner, T. Czech; Department of Neurology and Pathology and Molecular Medicine, Charles University, 2nd Faculty of Medicine, Motol University Hospital, Prague, Czech Republic: P. Krsek, J. Zamecnik; Department of Neuropathology, John Radcliffe Hospital, Oxford, UK: C. Kennard; Department of Anatomic Pathology Sciences, Università Sapienza, Rome, Italy: M. Antonelli, F. Giangaspero; Institute of Neuropathology, Westfälische Wilhelms – Universität Münster, Münster, Germany: W. Paulus, M. Hasselblatt; Department of Neuropathology, University Hospital Erlangen, Erlangen, Germany: R. Coras, I. Blümcke; Bethel Epilepsy Centre, Bielefeld, Germany: T. Polster, C.G. Bien; Laboratory of Neuropathology, Department of Neurology, Hospital de Santa Maria (CHLN), Lisbon, Portugal: J. Pimentel; Department of Human Pathology and Oncology, University of Florence and Division of Neurosurgery, 'Anna Meyer' Pediatric Hospital, Florence, Italy: A. M. Buccoliero, F. Giordano; Department of Pathology, Faculty of Medicine, Hacettepe University, Ankara, Turkey: F. Söylemezoğlu). In this regard we would like to acknowledge all personnel involved in sending us the material. Furthermore, the authors would like to thank Dr. Mark Nellist (Department of Clinical Genetics, Erasmus Medical Centre, Rotterdam, The Netherlands) and Dr. David J. Kwiatkowski, MD, PhD (Division of Experimental Medicine and Medical Oncology, Brigham and Women's Hospital, Boston, USA) for performing *TSC1/TSC2* mutation analysis.

Authorship

EA, JM and AM conceived the study and participated in its design and coordination. FJ, WS, WD, RC, IB, WP, VG, TS, JH, MF, KK, SJ, WG, AMB, CC, FG, LG, FS, JP, AM and EA contributed to the collection and selection of tissue samples or clinical data. DJ performed the methylation 450K analysis. AB, JM, DJ and BS performed the bioinformatics. DNA/RNA isolation, immunohistochemistry and evaluation of immunohistochemistry was performed by AB, supported by AM, CM and JA. Analysis of the data was performed by AB, JM and AM. AB, JM, DJ, BS, AS, AM and EA wrote the paper. All authors read and approved the final manuscript.

Funding

This work was supported by KIKA (Stichting Kinderen Kankervrij; AB, AM, AS, BS, EA); Stichting AMC Foundation (EA); Stichting TSC Fonds (EA); the Austrian Science Fund (FWF, no. J3499; AM); the European Union 7th framework program: acronym EPISTOP (grant agreement no. 602391; FJ, VG, TS, MF, SJ, AM, JM, EA) and acronym DESIRE (grant agreement no. 602531; IB); the European Union's Horizon 2020 WIDESPREAD-05-2020–Twinning, (EpiEpiNet; EA, JM), grant agreement no. 952455; the Polish Ministerial funds for science (years 2013-2018) for the implementation of international co-financed project (KK, SJ) and internal research project of the Children's Memorial Health Institute no.S132/2013 (KK, SJ).

Conflict of Interest

The authors have no duality or conflicts of interest to declare.

Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on request.

Ethics approval

Specimens were obtained and used in accordance with the Declaration of Helsinki and this study was approved by the Medical Ethics Committees of each institution.

References

1. Adriaansen ME, Schaefer-Prokop CM, Stijnen T, Duyndam DA, Zonnenberg BA, Prokop M (2009) "Prevalence of subependymal giant cell tumors in patients with tuberous sclerosis and a review of the literature". *Eur J Neurol* 16(6):691–696. doi:10.1111/j.1468-1331.2009.02567.x

2. Aronica E, Becker AJ, Spreafico R (2012) Malformations of cortical development. *Brain Pathol* 22(3):380–401. doi:10.1111/j.1750-3639.2012.00581.x
3. Aronica E, Crino PB (2014) "Epilepsy related to developmental tumors and malformations of cortical development." *Neurotherapeutics* 11 (2):251 – 68. doi: 10.1007/s13311-013-0251-0
4. Binder H, Willscher E, Loeffler-Wirth H, Hopp L, Jones DTW, Pfister SM, Kreuz M, Gramatzki D, Fortenbacher E, Hentschel B, Tatagiba M, Herrlinger U, Vatter H, Matschke J, Westphal M, Krex D, Schackert G, Tonn JC, Schlegel U, Steiger HJ, Wick W, Weber RG, Weller M, Loeffler M (2019) "DNA methylation, transcriptome and genetic copy number signatures of diffuse cerebral WHO grade II/III gliomas resolve cancer heterogeneity and development". *Acta Neuropathol Commun* 7(1):59. doi:10.1186/s40478-019-0704-8
5. Boer K, Crino PB, Gorter JA, Nellist M, Jansen FE, Spliet WG, van Rijen PC, Wittink FR, Breit TM, Troost D, Wadman WJ, Aronica E (2010) Gene expression analysis of tuberous sclerosis complex cortical tubers reveals increased expression of adhesion and inflammatory factors. *Brain Pathol* 20(4):704–719. doi:10.1111/j.1750-3639.2009.00341.x
6. Boer K, Jansen F, Nellist M, Redeker S, van den Ouweland AM, Spliet WG, van Nieuwenhuizen O, Troost D, Crino PB, Aronica E (2008) Inflammatory processes in cortical tubers and subependymal giant cell tumors of tuberous sclerosis complex. *Epilepsy Res* 78(1):7–21. doi:10.1016/j.epilepsyres.2007.10.002
7. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120. doi:10.1093/bioinformatics/btu170
8. Bongaarts A, de Jong JM, Broekaart DWM, van Scheppingen J, Anink JJ, Mijnsbergen C, Jansen FE, Spliet WGM, den Dunnen WFA, Gruber VE, Scholl T, Hainfellner JA, Feucht M, Borkowska J, Kotulska K, Jozwiak S, Grajkowska W, Buccoliero AM, Caporalini C, Giordano F, Genitori L, Scicluna BP, Schouten-van AYN, Meeteren EA, van Vliet A, Muhlebner JD, Mills, Aronica E (2020) Dysregulation of the MMP/TIMP Proteolytic System in Subependymal Giant Cell Astrocytomas in Patients With Tuberous Sclerosis Complex: Modulation of MMP by MicroRNA-320d In Vitro. *J Neuropathol Exp Neurol* 79(7):777–790. doi:10.1093/jnen/nlaa040
9. Bongaarts A, Giannikou K, Reinten RJ, Anink JJ, Mills JD, Jansen FE, Spliet GMW, den Dunnen WFA, Coras R, Blumcke I, Paulus W, Scholl T, Feucht M, Kotulska K, Jozwiak S, Buccoliero AM, Caporalini C, Giordano F, Genitori L, Soylemezoglu F, Pimentel J, Nellist M, Schouten-van Meeteren AYN, Nag A, Muhlebner A, Kwiatkowski DJ, Aronica E (2017) Subependymal giant cell astrocytomas in Tuberous Sclerosis Complex have consistent TSC1/TSC2 biallelic inactivation, and no BRAF mutations. *Oncotarget* 8(56):95516–95529. doi:10.18632/oncotarget.20764
10. Bongaarts A, van Scheppingen J, Korotkov A, Mijnsbergen C, Anink JJ, Jansen FE, Spliet WGM, den Dunnen WFA, Gruber VE, Scholl T, Samuelli S, Hainfellner JA, Feucht M, Kotulska K, Jozwiak S, Grajkowska W, Buccoliero AM, Caporalini C, Giordano F, Genitori L, Coras R, Blumcke I, Kresk P, Zamecnik J, Meijer L, Scicluna BP, A. Y. N. Schouten-van Meeteren, A. Muhlebner, J. D. Mills, and E. Aronica (2019) "The coding and non-coding transcriptional landscape of subependymal giant cell astrocytomas." *Brain*

11. Broekaart DWM, van Scheppingen J, Anink JJ, Wierts L, van Het Hof B, Jansen FE, Spliet WG, van Rijen PC, Kamphuis WW, de Vries HE, Aronica E, van Vliet EA (2019) "Increased matrix metalloproteinases expression in tuberous sclerosis complex: modulation by microRNA 146a and 147b in vitro". *Neuropathol Appl Neurobiol*. doi:10.1111/nan.12572
12. Buccoliero AM, Caporalini C, Giordano F, Mussa F, Scagnet M, Moscardi S, Baroni G, Genitori L, Taddei GL (2016) Subependymal giant cell astrocytoma: a lesion with activated mTOR pathway and constant expression of glutamine synthetase. *Clin Neuropathol* 35(5):295–301. doi:10.5414/NP300936
13. Buccoliero AM, Franchi A, Castiglione F, Gheri CF, Mussa F, Giordano F, Genitori L, Taddei GL (2009) Subependymal giant cell astrocytoma (SEGA): Is it an astrocytoma? Morphological, immunohistochemical and ultrastructural study. *Neuropathology* 29(1):25–30. doi:10.1111/j.1440-1789.2008.00934.x
14. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, Koelsche C, Sahm F, Chavez L, Reuss DE, Kratz A, Wefers AK, Huang K, Pajtler KW, Schweizer L, Stichel D, Olar A, Engel NW, Lindenberg K, Harter PN, Braczynski AK, Plate KH, Dohmen H, Garvalov BK, Coras R, Holsken A, Hewer E, Bewerunge-Hudler M, Schick M, Fischer R, Beschorner R, Schittenhelm J, Staszewski O, Wani K, Varlet P, Pages M, Temming P, Lohmann D, Selt F, Witt H, Milde T, Witt O, Aronica E, Giangaspero F, Rushing E, Scheurlen W, Geisenberger C, Rodriguez FJ, Becker A, Preusser M, Haberler C, Bjerkvig R, Cryan J, Farrell M, Deckert M, Hench J, Frank S, Serrano J, Kannan K, Tzirigos A, Bruck W, Hofer S, Brehmer S, Seiz-Rosenhagen M, Hanggi D, Hans V, Rozsnoki S, Hansford JR, Kohlhof P, Kristensen BW, Lechner M, Lopes B, Mawrin C, Ketter R, Kulozik A, Khatib Z, Heppner F, Koch A, Jouvet A, Keohane C, Muhleisen H, Mueller W, Pohl U, Prinz M, Benner A, Zapatka M, Gottardo NG, Driever PH, Kramm CM, Muller HL, Rutkowski S, von Hoff K, Fruhwald MC, Gnekow A, Fleischhack G, Tippelt S, Calaminus G, Monoranu CM, Perry A, Jones C, Jacques TS, Radlwimmer B, Gessi M, Pietsch T, Schramm J, Schackert G, Westphal M, Reifenberger G, Wesseling P, Weller M, Collins VP, Blumcke I, Bendszus M, Debus J, Huang A, Jabado N, Northcott PA, Paulus W, Gajjar A, Robinson GW, Taylor MD, Jaunmuktane Z, Ryzhova M, Platten M, Unterberg A, Wick W, Karajannis MA, Mittelbronn M, Acker T, Hartmann C, Aldape K, Schuller U, Buslei R, Lichter P, Kool M, Herold-Mende C, D. W. Ellison, M. Hasselblatt, M. Snuderl, S. Brandner, A. Korshunov, A. von Deimling, and S. M. Pfister. 2018. "DNA methylation-based classification of central nervous system tumours." *Nature* 555 (7697):469–474. doi: 10.1038/nature26000
15. Capper D, Stichel D, Sahm F, Jones DTW, Schrimpf D, Sill M, Schmid S, Hovestadt V, Reuss DE, Koelsche C, Reinhardt A, Wefers AK, Huang K, Sievers P, Ebrahimi A, Scholer A, Teichmann D, Koch A, Hanggi D, Unterberg A, Platten M, Wick W, Witt O, Milde T, Korshunov A, Pfister SM, and A. von Deimling. 2018. "Practical implementation of DNA methylation and copy-number-based CNS tumor diagnostics: the Heidelberg experience" *Acta Neuropathol* 136 (2):181–210. doi:10.1007/s00401-018-1879-y
16. Chan JA, Zhang H, Roberts PS, Jozwiak S, Wieslawa G, Lewin-Kowalik J, Kotulska K, Kwiatkowski DJ (2004) Pathogenesis of tuberous sclerosis subependymal giant cell astrocytomas: biallelic

- inactivation of TSC1 or TSC2 leads to mTOR activation. *J Neuropathol Exp Neurol* 63(12):1236–1242
17. Croft D, O'Kelly G, Wu G, Haw R, Gillespie M, Matthews L, Caudy M, Garapati P, Gopinath G, Jassal B, Jupe S, Kalatskaya I, Mahajan S, May B, Ndegwa N, Schmidt E, Shamovsky V, Yung C, Birney E, Hermjakob H, PD'Eustachio, Stein L (2011) Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res* 39(Database issue):D691–D697. doi:10.1093/nar/gkq1018
 18. Cuccia V, Zuccaro G, Sosa F, Monges J, Lubienicky F, Taratuto AL (2003) Subependymal giant cell astrocytoma in children with tuberous sclerosis. *Childs Nerv Syst* 19(4):232–243. doi:10.1007/s00381-002-0700-2
 19. Curatolo P, Bombardieri R, Jozwiak S (2008) "Tuberous sclerosis". *Lancet* 372(9639):657–668. doi:10.1016/S0140-6736(08)61279-9
 20. Curatolo P, Moavero R, de Vries PJ (2015) Neurological and neuropsychiatric aspects of tuberous sclerosis complex. *Lancet Neurol* 14(7):733–745. doi:10.1016/S1474-4422(15)00069-1
 21. Davis PE, Filip-Dhima R, Sideridis G, Peters JM, Au KS, Northrup H, Bebin EM, Wu JY, Krueger D, Sahin M, and Network Tuberous Sclerosis Complex Autism Center of Excellence Research. 2017. "Presentation and Diagnosis of Tuberous Sclerosis Complex in Infants." *Pediatrics* 140 (6). doi:10.1542/peds.2016-4040
 22. Dibble CC, Elis W, Menon S, Qin W, Klekota J, Asara JM, Finan PM, Kwiatkowski DJ, Murphy LO, Manning BD (2012) "TBC1D7 is a third subunit of the TSC1-TSC2 complex upstream of mTORC1". *Mol Cell* 47(4):535–546. doi:10.1016/j.molcel.2012.06.009
 23. Fortin JP, Labbe A, Lemire M, Zanke BW, Hudson TJ, Fertig EJ, Greenwood CM, Hansen KD (2014) Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol* 15(12):503. doi:10.1186/s13059-014-0503-2
 24. Franz DN, Agricola K, Mays M, Tudor C, Care MM, Holland-Bouley K, Berkowitz N, Miao S, Peyrard S, Krueger DA (2015) Everolimus for subependymal giant cell astrocytoma: 5-year final analysis. *Ann Neurol* 78(6):929–938. doi:10.1002/ana.24523
 25. Fujiwara S, Takaki T, Hikita T, Nishio S (1989) "Subependymal giant-cell astrocytoma associated with tuberous sclerosis. Do subependymal nodules grow?". *Childs Nerv Syst* 5(1):43–44. doi:10.1007/BF00706748
 26. Gos M (2013) "Epigenetic mechanisms of gene expression regulation in neurological diseases". *Acta Neurobiol Exp (Wars)* 73(1):19–37
 27. Henshall DC, Kobow K (2015) "Epigenetics and Epilepsy" *Cold Spring Harb Perspect Med* 5 (12). doi:10.1101/cshperspect.a022731
 28. Inoki K, Li Y, Xu T, Guan KL (2003) Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling. *Genes Dev* 17(15):1829–1834. doi:10.1101/gad.1110003
 29. Jiang WG, Sampson J, Martin TA, Lee-Jones L, Watkins G, Douglas-Jones A, Mokbel K, Mansel RE (2005) "Tuberlin and hamartin are aberrantly expressed and linked to clinical outcome in human

- breast cancer: the role of promoter methylation of TSC genes". *Eur J Cancer* 41(11):1628–1636. doi:10.1016/j.ejca.2005.03.023
30. Jones PA, Baylin SB (2007) "The epigenomics of cancer" *Cell* 128(4):683–692. doi:10.1016/j.cell.2007.01.029
31. Jozwiak S, Mander M, Mlynarski W (2015) Natural History and Current Treatment Options for Subependymal Giant Cell Astrocytoma in Tuberous Sclerosis Complex. *Semin Pediatr Neurol* 22(4):274–281. doi:10.1016/j.spen.2015.10.003
32. Keshet I, Yisraeli J, Cedar H (1985) Effect of regional DNA methylation on gene expression. *Proc Natl Acad Sci U S A* 82(9):2560–2564. doi:10.1073/pnas.82.9.2560
33. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14(4):R36. doi:10.1186/gb-2013-14-4-r36
34. Klutstein M, Nejman D, Greenfield R, Cedar H (2016) "DNA Methylation in Cancer and Aging". *Cancer Res* 76(12):3446–3450. doi:10.1158/0008-5472.CAN-15-3278
35. Kobow K, Ziemann M, Kaipananickal H, Khurana I, Muhlebner A, Feucht M, Hainfellner JA, Czech T, Aronica E, Pieper T, Holthausen H, Kudernatsch M, Hamer H, Kasper BS, Rossler K, Conti V, Guerrini R, Coras R, Blumcke I, El-Osta A, Kaspi A (2019) "Genomic DNA methylation distinguishes subtypes of human focal cortical dysplasia" *Epilepsia* 60(6):1091–1103. doi:10.1111/epi.14934
36. Laffaire J, Everhard S, Idbaih A, Criniere E, Marie Y, de Reynies A, Schiappa R, Mokhtari K, Hoang-Xuan K, Sanson M, Delattre JY, Thillet J, Ducray F (2011) Methylation profiling identifies 2 groups of gliomas according to their tumorigenesis. *Neuro Oncol* 13(1):84–98. doi:10.1093/neuonc/noq110
37. Lesma E, Sirchia SM, Ancona S, Carelli S, Bosari S, Ghelma F, Montanari E, Di Giulio AM, Gorio A (2009) "The methylation of the TSC2 promoter underlies the abnormal growth of TSC2 angiomyolipoma-derived smooth muscle cells". *Am J Pathol* 174(6):2150–2159. doi:10.2353/ajpath.2009.080799
38. Liao Y, Smyth GK, Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30(7):923–930. doi:10.1093/bioinformatics/btt656
39. Lim HK, Choi YA, Park W, Lee T, Ryu SH, Kim SY, Kim JR, Kim JH, Baek SH (2003) "Phosphatidic acid regulates systemic inflammatory responses by modulating the Akt-mammalian target of rapamycin-p70 S6 kinase 1 pathway". *J Biol Chem* 278(46):45117–45127. doi:10.1074/jbc.M303789200
40. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW (2016) The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol* 131(6):803–820. doi:10.1007/s00401-016-1545-1
41. Love MI, Huber W, Anders S (2014) "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2". *Genome Biol* 15(12):550. doi:10.1186/s13059-014-0550-8

42. Martin KR, Zhou W, Bowman MJ, Shih J, Au KS, Dittenhafer-Reed KE, Sisson KA, Koeman J, Weisenberger DJ, Cottingham SL, DeRoos ST, Devinsky O, Winn ME, Cherniack AD, Shen H, Northrup H, Krueger DA, MacKeigan JP (2017) The genomic landscape of tuberous sclerosis complex. *Nat Commun* 8:15816. doi:10.1038/ncomms15816
43. Mi R, Ma J, Zhang D, Li L, Zhang H (2009) Efficacy of combined inhibition of mTOR and ERK/MAPK pathways in treating a tuberous sclerosis complex cell model. *J Genet Genomics* 36(6):355–361. doi:10.1016/S1673-8527(08)60124-1
44. Mills JD, Iyer AM, van Scheppingen J, Bongaarts A, Anink JJ, Janssen B, Zimmer TS, Spliet WG, van Rijen PC, Jansen FE, Feucht M, Hainfellner JA, Krsek P, Zamecnik J, Kotulska K, Jozwiak S, Jansen A, Lagae L, Curatolo P, Kwiatkowski DJ, Pasterkamp RJ, Senthilkumar K, von Oerthel L, Hoekman MF, Gorter JA, Crino PB, Muhlebauer A, Scicluna BP, Aronica E (2017) Coding and small non-coding transcriptional landscape of tuberous sclerosis complex cortical tubers: implications for pathophysiology and treatment. *Sci Rep* 7(1):8089. doi:10.1038/s41598-017-06145-8
45. Morimoto K, Mogami H (1986) Sequential CT study of subependymal giant-cell astrocytoma associated with tuberous sclerosis. Case report. *J Neurosurg* 65(6):874–877. doi:10.3171/jns.1986.65.6.0874
46. Prabowo AS, Anink JJ, Lammens M, Nellist M, van den Ouweland AM, Adle-Biassette H, Sarnat HB, Flores-Sarnat L, Crino PB, Aronica E (2013) "Fetal brain lesions in tuberous sclerosis complex: TORC1 activation and inflammation". *Brain Pathol* 23(1):45–59. doi:10.1111/j.1750-3639.2012.00616.x
47. Stone TJ, Keeley A, Virasami A, Harkness W, Tisdall M, Izquierdo Delgado E, Gutteridge A, Brooks T, Kristiansen M, Chalker J, Wilkhu L, Mifsud W, Apps J, Thom M, Hubank M, Forshew T, Cross JH, Hargrave D, Ham J, Jacques TS (2018) Comprehensive molecular characterisation of epilepsy-associated glioneuronal tumours. *Acta Neuropathol* 135(1):115–129. doi:10.1007/s00401-017-1773-z
48. Tyburczy ME, Kotulska K, Pokarowski P, Mieczkowski J, Kucharska J, Grajkowska W, Roszkowski M, Jozwiak S, Kaminska B (2010) "Novel proteins regulated by mTOR in subependymal giant cell astrocytomas of patients with tuberous sclerosis complex and new therapeutic implications". *Am J Pathol* 176(4):1878–1890. doi:10.2353/ajpath.2010.090950
49. Wilkerson MD, Hayes DN (2010) ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26(12):1572–1573. doi:10.1093/bioinformatics/btq170
50. Yu G, Wang LG, Han Y, He QY (2012) "clusterProfiler: an R package for comparing biological themes among gene clusters." *OMICS* 16 (5):284-7. doi: 10.1089/omi.2011.0118

Figures

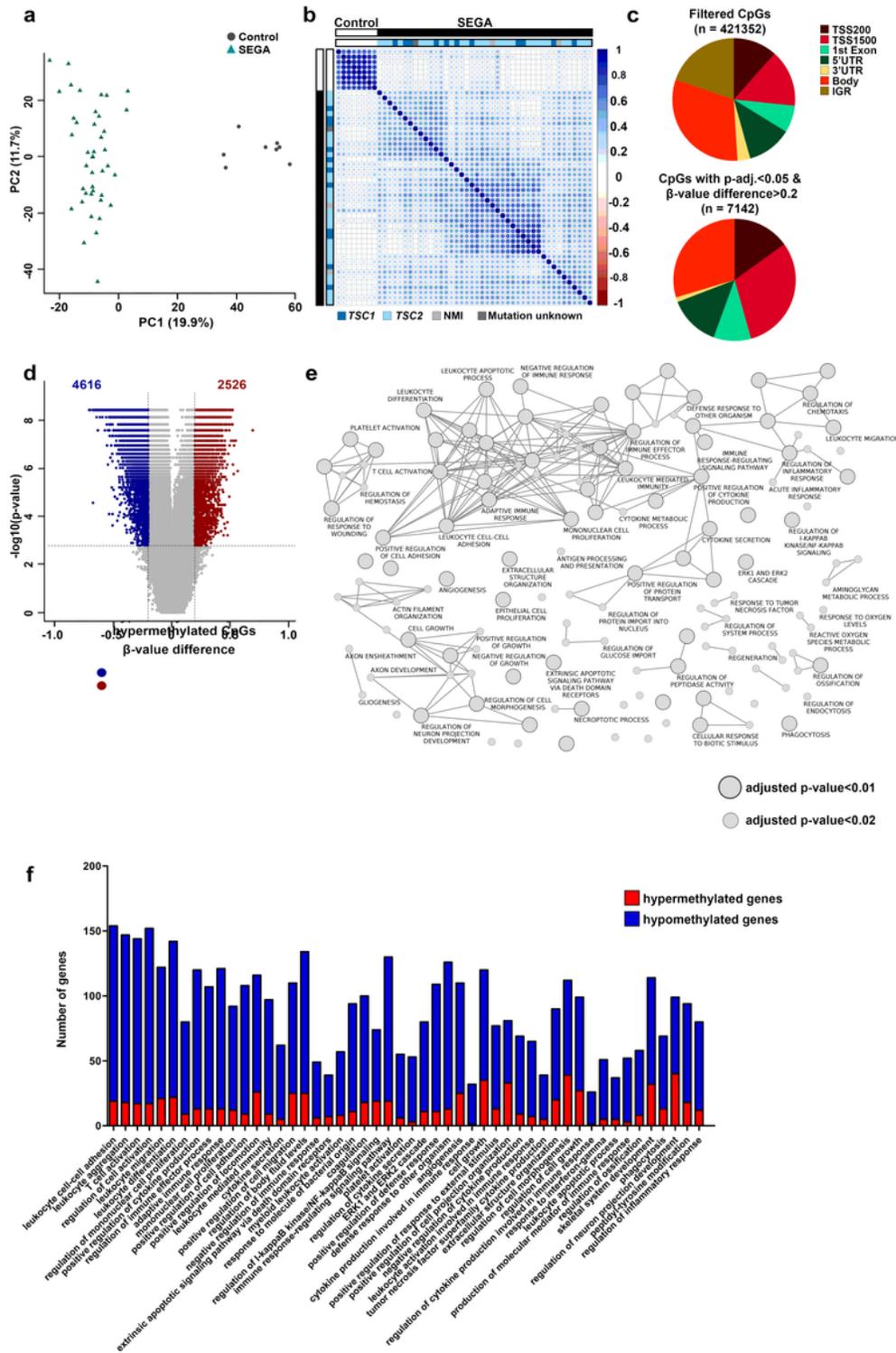


Figure 1

The methylation profile of SEGAs. a. A principal component analysis (PCA) of the methylation data in SEGAs (n=42) and periventricular control tissue (n=8) showing that the major source of variability in CpG methylation is the diagnosis. X-axis: the first principal component (PC); y-axis: the second PC. b. Spearman's rank correlation matrix of the methylation data showing separate clustering of SEGAs from periventricular control tissue. The scale bar indicates the strength of the correlation with 1 indicating a

strong positive correlation (dark blue) and -1 indicating a negative correlation (dark red) between samples. c. Pie charts showing the distribution of CpGs on the gene region (TSS200, TSS1500, 5'UTR and Exon 1, IGR, 3'UTR or gene body). The upper pie chart shows the distribution for 421,352 CpGs selected after filtering for probes with a detection p-values of more than 0.01, located on the sex chromosomes, or in SNPs were removed as well as cross-hybridization probes. The lower pie chart shows the gene distribution after selecting for an adjusted p-value<0.01 and a β -value difference of >0.2. d. Volcano plot showing the differentially methylated CpGs on the TSS-associated regions (adjusted p-value<0.01 and a β -value difference of >0.2) between SEGAs and control tissue. A total of 4616 CpGs were hypomethylated and 2526 were hypermethylated in SEGAs compared to control tissue. e. Schematic overview using Cytoscape of GO terms enriched in SEGAs compared to control tissue (adjusted p-value<0.02). Lines indicate genes in common between GO terms. f. Graphical representation of hypermethylated (red) and hypomethylated (blue) in the top 50 GO terms.

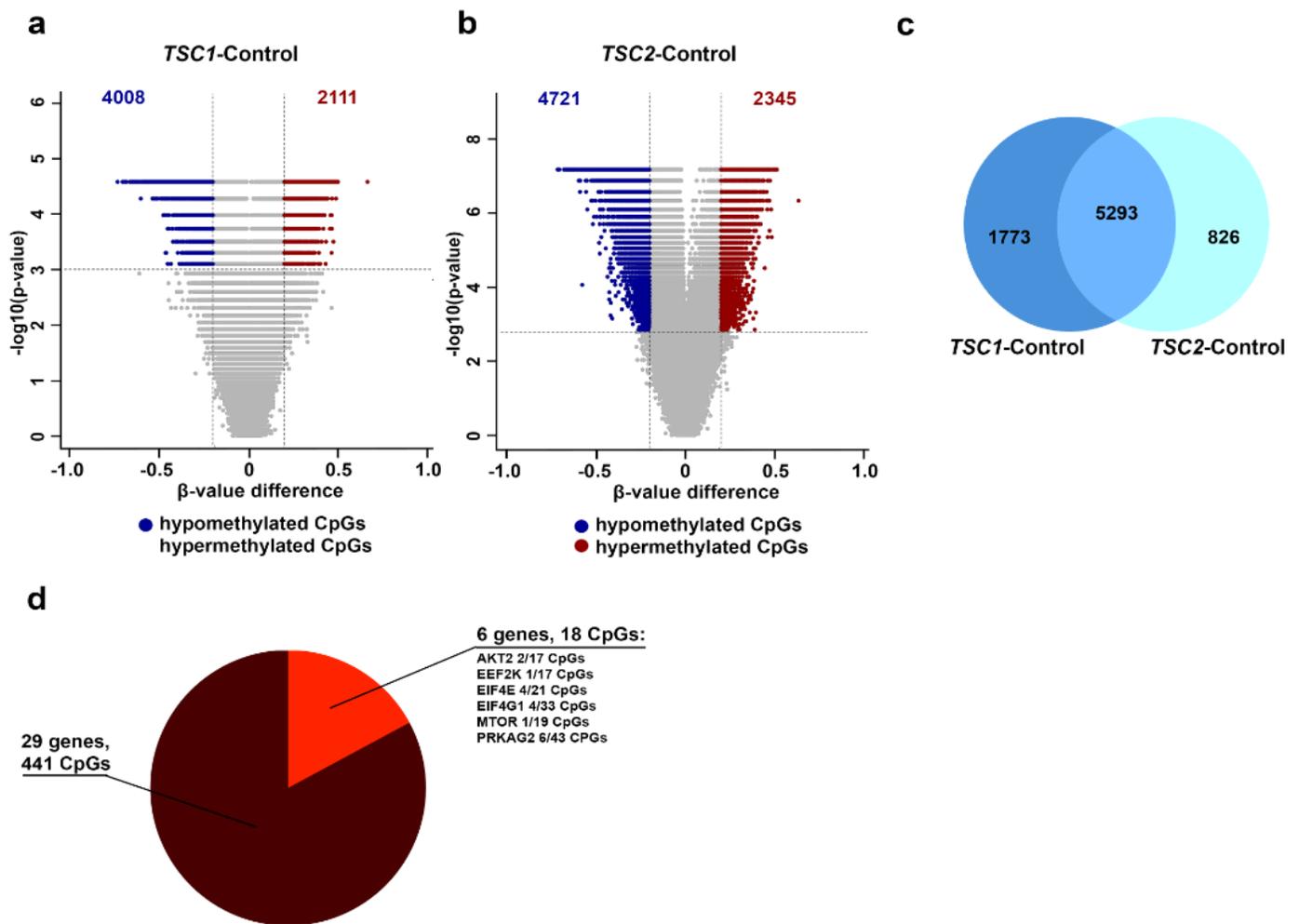


Figure 2

Methylation of the mTOR pathway in SEGAs. a. Volcano plot showing the differentially methylated CpGs on the TSS-associated regions (adjusted p-value<0.01 and a β -value difference of >0.2) between TSC1 mutated SEGAs and control tissue (TSC1-control). A total of 4008 CpGs were hypomethylated and 2111

were hypermethylated in TSC1-control. b. Volcano plot showing the differentially methylated CpGs on the TSS-associated regions (adjusted p-value<0.01 and a β -value difference of >0.2) between TSC2 mutated SEGAs and control tissue(TSC2-control). A total of 4008 CpGs were hypomethylated and 2111 were hypermethylated in TSC2-control. c. Venn diagram showing the overlap of CpGs between TSC1-control and TSC2-control. 5293 CpGs were overlapping, whereas 1773 CpGs were only found differentially methylated in TSC1-control and 826 CpGs in TSC2-control. d. Pie chart showing the distribution of differentially methylated and not differentially methylated mTOR pathway related genes (based on Reactome). A total of 18 CpGs located on 6 mTOR pathway related genes were differentially methylated, whereas 441 CpGs located on 29 mTOR pathway related genes were not differentially methylated.

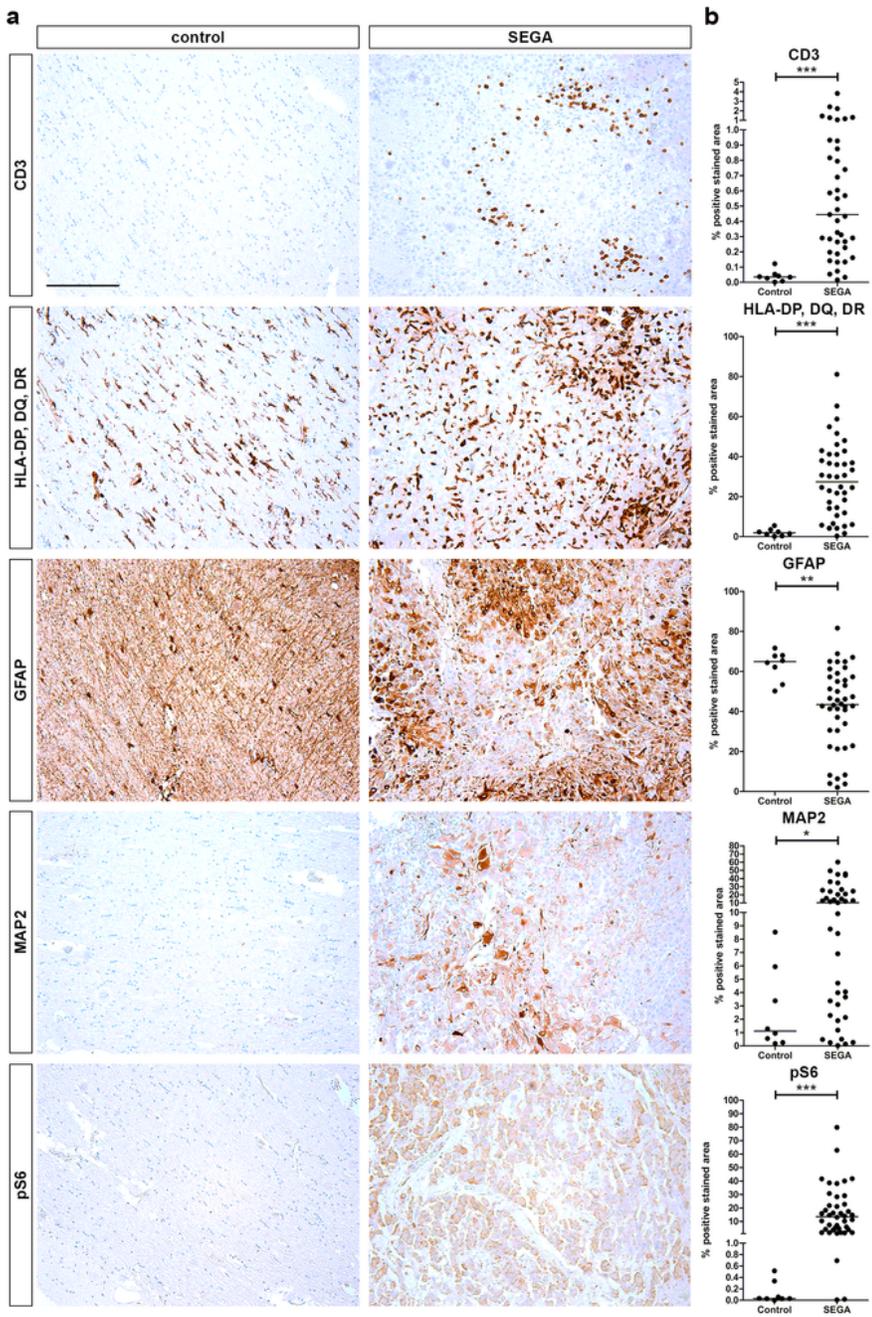


Figure 3

Expression of inflammatory, mTOR activation, glial and neuronal markers SEGAs. a. Immunohistochemistry showing the expression of CD3, HLA-DP/DQ/DR, pS6, GFAP and MAP2 in SEGA and periventricular control. Scale bar = 200 um. b. Quantification of immunohistochemistry indicated in % of positive area of CD3, HLA-DP/DQ/DR, pS6, GFAP and MAP2 in SEGA and periventricular control. An

higher expression of CD3, HLA-DP/DQ/DR, pS6 and MAP2 and a lower expression of GFAP was found in SEGA compared to control. *p-value<0.05, **p-value<0.01, ***p-value<0.001, Mann-Whitney U test.

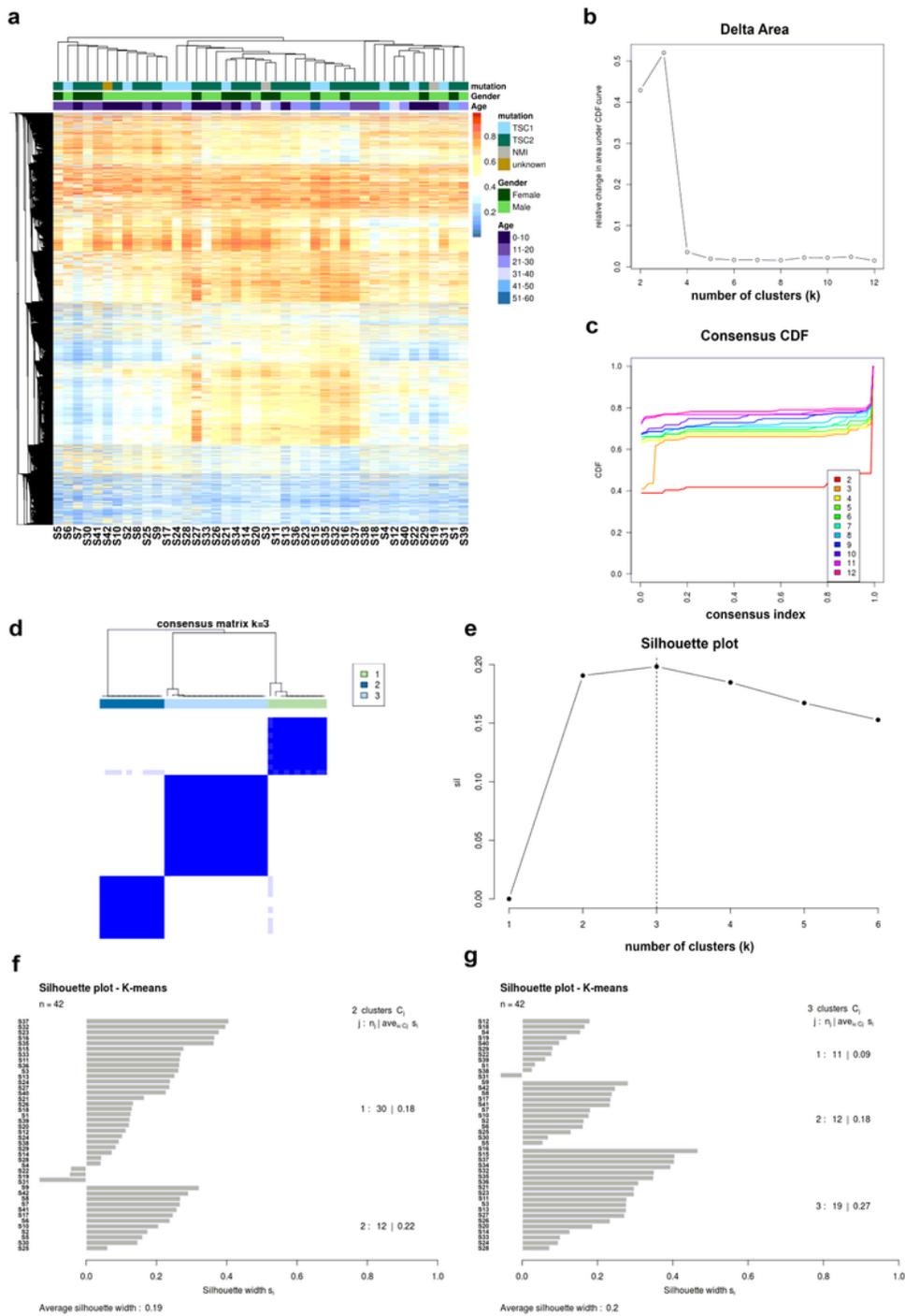


Figure 4

Two robust groups found in the methylation data of SEGAs. a. Heatmap using hierarchical clustering showing two/three subgroups in SEGA. b-d. Consensus clustering showing that k=3 is most robust indicated by delta area plot (b) consensus Cumulative Distribution Function (CDF) plot (c) and consensus

matrix for k=3 (d). e. Silhouette plot showing the average silhouette with for each k, indicating that k=3 is most robust. f. Barchart of the silhouette with for k=2 showing which samples cluster together. g. Barchart of the silhouette with for k=3 showing which samples cluster together.

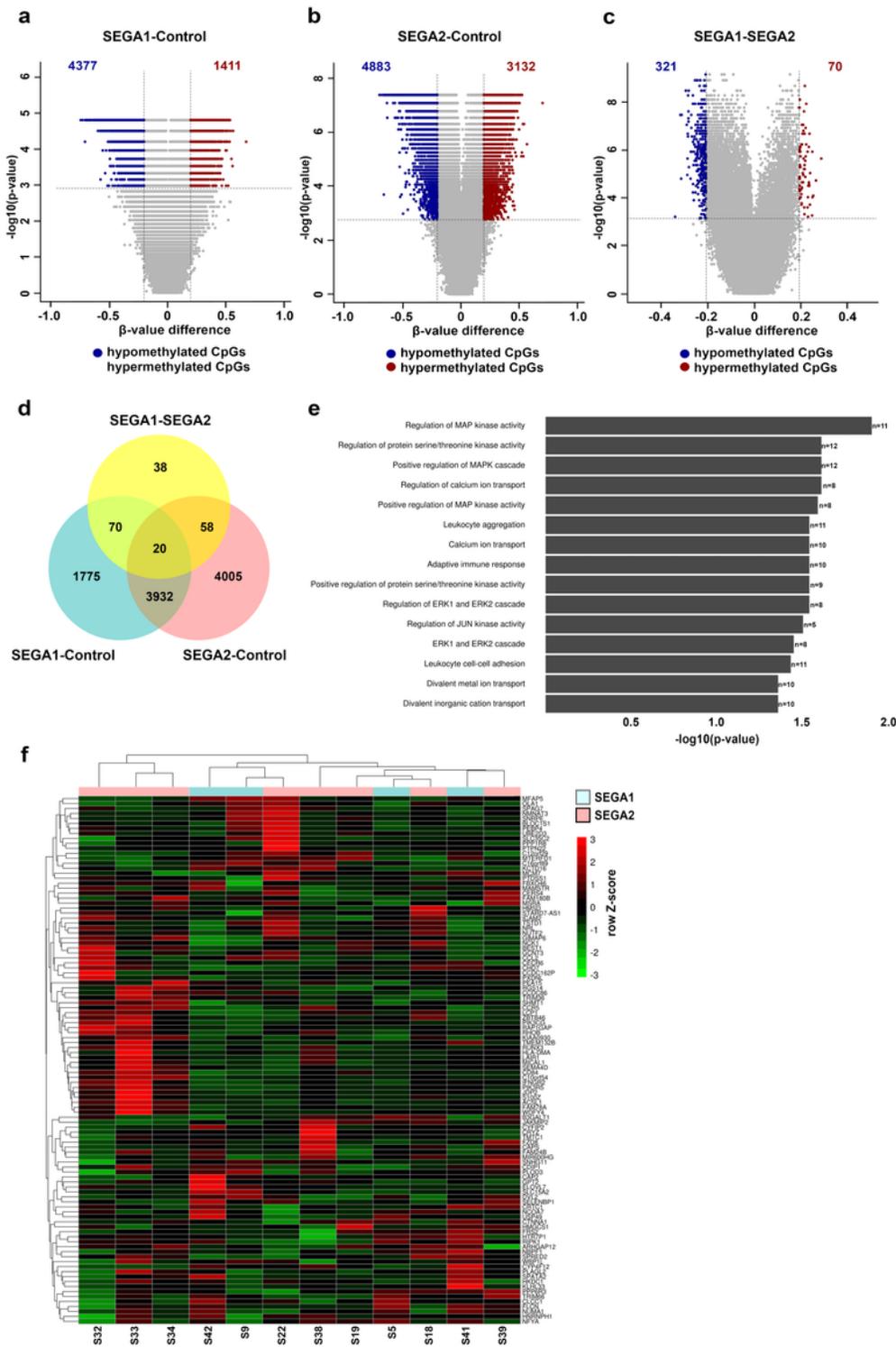


Figure 5

Unique gene expression between two SEGA subgroups. a. Volcano plot showing the differentially methylated CpGs on the TSS-associated regions (adjusted p-value<0.01 and a β -value difference of >0.2)

between one subgroup of SEGAs and control tissue (SEGA1-control). A total of 4377 CpGs were hypomethylated and 1411 were hypermethylated in SEGA1-control. b. Volcano plot showing the differentially methylated CpGs on the TSS-associated regions (adjusted p-value<0.01 and a β -value difference of >0.2) between the other subgroup of SEGAs and control tissue (SEGA2-control). A total of 4883 CpGs were hypomethylated and 3132 were hypermethylated in SEGA2-control. c. Volcano plot showing the differentially methylated CpGs on the TSS-associated regions (adjusted p-value<0.01 and a β -value difference of >0.2) between the two subgroups of SEGA (SEGA1-SEGA2). A total of 321 CpGs were hypomethylated and 70 were hypermethylated in SEGA1-SEGA2. d. Venn diagram showing the overlap of genes corresponding to the differentially methylated CpGs between SEGA1-control, SEGA2-control and SEGA1-SEGA2. Genes overlapping between SEGA1-control and SEGA1-SEGA2 but not with SEGA2-control were considered unique for SEGA1. Genes overlapping between SEGA2-control and SEGA1-SEGA2 but not with SEGA1-control were considered unique for SEGA2 (70 genes SEGA1 unique and 58 genes SEGA2 unique). e. Bar chart showing the GO terms enriched for the SEGA1 and SEGA 2 unique genes. Each GO term is listed on the y-axis, the $\log_{10}(1/\text{adjusted p-value})$ on the x-axis and the n is equal to the number of differentially methylated genes in each GO term. f. Heatmap with Z-score hierarchical clustering for the RNA expression data of 12 SEGAs. Each row indicates one of unique SEGA1 or SEGA2 genes based on the methylation data that was expressed on RNA level in SEGA. The color scale means the gene expression standard deviations from the mean, with green for low expression and red for the high expression levels.

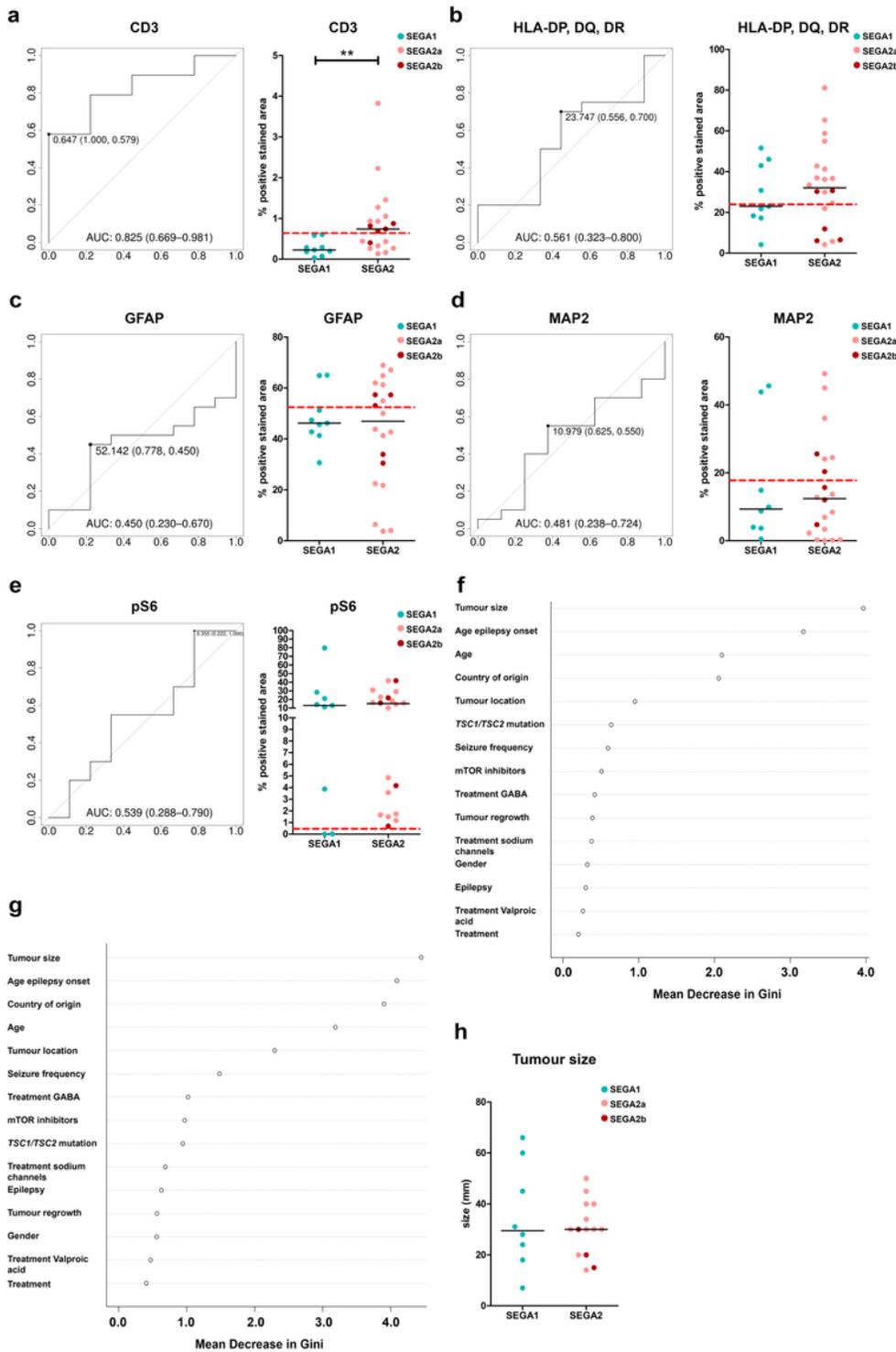


Figure 6

Detection of SEGA subgroups based on histological markers and clinical data. a-e. ROC curves and scatterplots for detection of SEGA1 or SEGA2 based on CD3 (a), HLA-DP/DQ/DR (b), GFAP (c), MAP2 and pS6 (e) positivity. The point in the ROC curve indicates the most optimal % of positivity to separate the SEGA1 and SEGA2 group, followed by the proportion of SEGA1 and SEGA2 that is correctly detected. Scatterplots show the spread between samples for the % of positive area for SEGA1 (blue) and SEGA2

(SEGA2a light red and SEGA2b dark red). The red line indicates the most optimal % of positivity to separate the SEGA1 and SEGA2 group based on the ROC curve. *p-value<0.05, **p-value<0.01, ***p-value<0.001, Mann-Whitney U test. f. Variable Importance plots obtained from Random Forest in R for detection of SEGA1 and SEGA2 based on the clinical data (Table 1). Each point represents the mean decrease Gini value, indicative of the importance of each variable. Variables are listed from most important to least important. Antiepileptic drugs were also categorized in 3 groups: GABA blockers (treatment GABA), valproic acid (treatment valproic acid) and sodium channel blockers (treatment sodium channels). g. Variable Importance plots obtained from Random Forest in R for detection of SEGA1, SEGA2a and SEGA2b based on the clinical data (Table 1). Each point represents the mean decrease Gini value, indicative of the importance of each variable. Variables are listed from most important to least important. Antiepileptic drugs were also categorized in 3 groups: GABA blockers (treatment GABA), valproic acid (treatment valproic acid) and sodium channel blockers (treatment sodium channels). h. Scatterplot showing no difference in tumor size (mm) in SEGA1 (blue) compared to SEGA2 (SEGA2a light red and SEGA2b dark red).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigure1.tif](#)
- [Supfigure2.tif](#)
- [Table1.xlsx](#)
- [Table2.xlsx](#)
- [Table3.xlsx](#)