Clifford Steven C (Orcid ID: 0000-0002-8573-8009)

Timing is everything: A connection between medulloblastoma prognosis and fetal cerebellar development.

Daniel Williamson[1], Edward C Schwalbe[1,2], Simon Bailey[1], Steven C Clifford[1].

[1] Wolfson Childhood Cancer Research Centre, Translational and Clinical Research Institute, Newcastle University Centre for Cancer, Newcastle University, Newcastle upon Tyne, UK.

[2] Department of Applied Sciences, Northumbria University, Newcastle upon Tyne, UK.

Corresponding Author: e-mail: daniel.williamson@newcastle.ac.uk

Abstract

The childhood brain tumour medulloblastoma is typically classified into multiple discrete molecular subgroups with characteristic DNA methylation and expression patterns. Several of these subgroups are used as, or proposed to be, an effective basis for treatment stratification. Here, we highlight the close connection between the findings described in a recent series of studies which, together, strongly imply a continuous association between survival outcome, the transcriptional profile of a Group3/Group4 (i.e., non-WNT/non-SHH) medulloblastoma and the specific point during early fetal cerebellar development at which initial pathogenic disruption took place. This has important implications for future efforts to model the disease by incorporating driving molecular features into their specific developmental context.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/nan.12903

This further suggests that instead of relying upon discrete DNA methylation subgroups, using expression biomarkers as the basis of a continuous risk predictor may produce a more effective risk stratification of patients with Group3/Group4 medulloblastoma.

Main Text

Molecular-biological prognostication in medulloblastoma has been driven, over the last decade, by the discovery of increasing numbers of molecular subgroups and their associated mutations. Initially, medulloblastoma was divided by transcriptional profiling into SHH, WNT, Group3 (MB_{Grp3}) and Group4 (MB_{Grp4})[1]; each has now been further subdivided into subgroups by DNA methylation patterns[2-4]. For example, the combined Group3/Group4 medulloblastoma ($MB_{Grp3/Grp4}$) was divided into eight further subgroups (I-VIII)[5]. Whilst of great biological interest, this atomization of an already rare disease presents a practical problem to clinical trialists aiming to preserve statistical power. Here, we highlight the close connections between the findings described in a series of papers published in close succession by ourselves and others[6-9], which together support the prognostic potential of gene expression across $MB_{Grp3/Grp4}$ medulloblastoma as a whole and its relationship to normal cerebellar development.

Our recent study[9] analysed transcriptional profiles – in $MB_{Grp3/Grp4}$ combined - to describe a single transcriptional continuum; a pattern of continuous expression changes that links all MB_{Grp3} and MB_{Grp4} patients. We devised a "G3/G4" score to represent this expression pattern and used this score to place each patient at a unique position along the continuum between two extremes (i.e., the archetypal MB_{Grp3} and MB_{Grp4} transcriptional states). Position on the continuum was reflective of an individual's clinicopathology and significantly related to 5-year survival.

Korshunov *et al.*[7] also recently performed a transcriptomic analysis of MB_{Grp3}, describing differences in gene expression between patients who died versus those who survived 5 years post-diagnosis. They identified six differentially expressed genes which were associated with high-risk disease - *MYC, KIRREL2, ITPRIPL1, DCAF4, NPW* and *CDT1* – highlighting *KIRREL2* expression in particular, as prognostic, independent of other clinico-molecular features, most notably $MB_{Grp3/Grp4}$ DNA methylation

subgroups (I-VIII)[5]. High *KIRREL2* expression was present in all MB_{Grp3} subgroups and, in combination with other risk factors, stratified disease risk with good accuracy.

Whilst Korshunov *et al.*[7] divided their patients into high and low expressors, the survival association we described was linear and continuous; the higher an individual's G3/G4 score (i.e., the more "MB_{Grp3}-like"), the worse the prognosis. Notably, within our $MB_{Grp3/Grp4}$ cohort, the expression of *KIRREL2* is log-linearly correlated to the G3/G4 score and a major contributor (22nd out of 56546 transcripts) to the $MB_{Grp3/Grp4}$ continuum signature (Figure 1A); indeed, each of the top 6 genes described are significantly correlated with the G3/G4 score and with survival (Figure 1B-D). This strongly suggests that the results described by Korshunov *et al.*[7] are substantially part of the phenomenon described by the G3/G4 continuum.

Our G3/G4 expression continuum - of which *KIRREL2* expression is a key constituent – was also projected onto an scRNA-seq expression atlas of early human fetal cerebellar development[10], showing that the continuum is mirrored by a specific developmental trajectory - beginning with early rhombic lip (RL) precursors (most MB_{Grp3}-like) and ending with differentiated unipolar brush cells (UBC) (most MB_{Grp4}-like) - linking individual tumours to specific developmental states by their position on the continuum. *KIRREL2* is a gene involved in cerebellar development, usually regarded as an early GABAergic cell fate determinant. This is perhaps counter-intuitive for an RL/UBC (primarily glutamatergic) derived tumour, although recent descriptions of a multipotent posterior transitory zone between the ventricular zone and RL border may provide some explanation[11].

Smith *et al.*[8] and Hendrikse *et al.*[6] also recently published similar findings; i.e., a single developmental RL/UBC lineage as the origin of the $MB_{Grp3/Grp4}$ subgroups based on analysis of the same scRNA-seq atlas[12]. Both studies aligned similar G3/G4 transcriptional patterns to two spatially distinct ventricular and subventricular compartments within early human RL development. Hendrikse *et al.*[6] further explored how the development of cells might be interrupted by disruption of the CBFA complex. They proposed the existence of PeRLs (Persistent Rhombic Lip), a postulated pre-malignant dysplasia within the cerebellar nodulus, resulting from disrupted fetal development/differentiation.

They conflated PeRLs with previous reports of nodulus hyperplasia/genetic alterations dating back to the 1960s[13] and 1990s[14], surmising that further genetic hits could result in medulloblastoma.

Taken together, this series of recent studies forms a throughline suggesting that timing of initial disruption of early fetal development lineages fixes aspects of their tumour transcriptional profile at a given point and that this tracks closely to an individual's chance of surviving following treatment (Figure 1E).

If a more "MB_{Grp3}-like" position upon the continuum, exemplified by high levels of *KIRREL2* expression, denotes an "earlier" developmental disruption, why should those tumours be more aggressive and difficult to treat? It could be by virtue of their age at diagnosis – position on the continuum tracks closely to the average age of diagnosis [9] – although when controlling for age a significantly worse prognosis is still observed. Is there something intrinsic to the retained developmental biology which creates a more aggressive or treatment-resistant cell? We may speculate about the expression of *MYC* or more broadly about an undifferentiated/proliferative phenotype, but it is not yet obvious specifically why they should be more resistant to therapy.

If a disruption occurs during fetal development, there is a substantial latent period, but why should pre-malignancies at one end of the continuum/developmental lineage lie dormant for 3-4 years on average and at the other end, for several years more? It does not appear that medulloblastomas occurring in older patients (e.g., subtype VIII) require time to accrue more mutations or that the mutations required are - by virtue of the mode of mutation - inherently less likely to occur [4]. One interpretation is that "earlier" disruption in a more progenitor-like/undifferentiated cell allows more rounds of division before post-mitotic stalling, producing greater or longer-lasting dysplasia. Consequently, a larger pool of pre-malignant cells is available to suffer a second hit and develop into cancer, therefore occurring statistically sooner. For the present, this remains speculation, although we note that methods to estimate the number of symmetric/asymmetric cell divisions prior to tumour initiation exist. For instance, by monitoring SNVs accumulating naturally through cell division coupled with multi-regional tissue sampling and/or single-cell sequencing [15] or by cell tracing experiments within mouse models [16].

Regardless, the answer surely lies in observing and/or modelling these putative PeRLs whose pre-determined character may dictate an individual's tumour biology and outcome.

Critically, we note that two independent studies now confirm significant prognostic information, not readily apparent from DNA methylation subtyping, is encapsulated in a continuous manner within transcriptional patterns. This suggests a different approach to risk prognostication among $MB_{Grp3/Grp4}$ – as an alternative or adjunct to sub-categorization by subgroup - whereby a continuous base level of risk can be assigned to an individual patient based on their transcriptional profile, after which further independent risk modifiers (e.g., *MYC* amplification, presence of metastases, $MB_{Grp3/Grp4}$ Subgroup VII) may then be applied (Figure 1F).

The opportunities afforded by transcriptional analysis are perhaps best illustrated by those individuals who, by DNA methylation profiling belong to $MB_{Grp3/Grp4}$ subgroups II and III. These subgroups contain exclusively MB_{Grp3} individuals and, taken as a whole, demonstrate the classic poor prognosis historically associated with MB_{Grp3} [5]. Nevertheless, according to our transcriptional study[9], only ~40% of those patients could be considered to have a 5-year survival <50% by virtue of their biology alone, i.e., with no other risk modifiers taken into account. Similarly, Korshunov *et al.* [7] reported that the addition of *KIRREL2* expression to their biomarker scheme altered the risk category of up to 50% of MB_{Grp3} patients, with ~25% of such patients assigned to a low-risk category with 5-year OS 95%. Put simply, expression profiling can be deployed to reassign substantial proportions of patients more accurately within each $MB_{Grp3/Grp4}$ subtype into "better" or "worse" treatment stratification groups. The job of upcoming clinical trials/biological studies will be to determine how best to deploy such information in the future.

We note that a surrogate G3/G4 continuum score can be derived from DNA methylation profiles - widely used for diagnosis - if no expression profile exists[9]. Conversely, a focussed diagnostic transcriptomic assay (qRT-PCR, NanoString, *etc*) could technically be used to achieve prognostication on its own, where no DNA methylation profile exists; however, we stress that we do not believe this to be desirable. We would expect progressive efforts in clinical trials and standard-of-care to continue, for the time being, to rely, in part, upon information more obtainable by DNA methylation analysis and

therefore, we imagine both expression and methylation assays operating side-by-side for the foreseeable future.

In short, there is now good evidence, backed up by a developmental/biological rationale, that expression is an effective prognostic determinant - particularly for MB_{Grp3}. Further prospective clinical trials data will be required to definitively settle whether continuous or subgroup-based prognostication, or a combination of both approaches, is most efficacious.

Acknowledgements

This work was supported by Cancer Research UK (C8464/A13457 and C8464/A23391), the Tom Graham Trust/CCLG, LoveOliver, Star for Harris and the INSTINCT network, co-funded by The Brain Tumour Charity, Great Ormond Street Children's Charity, and Children with Cancer UK (grant no. 16/193).

References

1 Taylor MD, Northcott PA, Korshunov A, Remke M, Cho Y-J, Clifford SC, Eberhart CG, Parsons DW, Rutkowski S, Gajjar A, Ellison DW, Lichter P, Gilbertson RJ, Pomeroy SL, Kool M, Pfister SM. Molecular subgroups of medulloblastoma: the current consensus. Acta Neuropathol 2012; 123: 465-72

2 Cavalli FMG, Remke M, Rampasek L, Peacock J, Shih DJH, Luu B, Garzia L, Torchia J, Nor C, Morrissy AS, Agnihotri S, Thompson YY, Kuzan-Fischer CM, Farooq H, Isaev K, Daniels C, Cho B-K, Kim S-K, Wang K-C, Lee J-Y, Grajkowska WA, Perek-Polnik M, Vasiljevic A, Faure-Conter C, Jouvet A, Giannini C, Nageswara Rao AA, Li KKW, Ng H-K, Eberhart CG, Pollack IF, Hamilton RL, Gillespie GY, Olson JM, Leary S, Weiss WA, Lach B, Chambless LB, Thompson RC, Cooper MK, Vibhakar R, Hauser P, van Veelen M-LC, Kros JM, French PJ, Ra Y-S, Kumabe T, López-Aguilar E, Zitterbart K, Sterba J, Finocchiaro G, Massimino M, Van Meir EG, Osuka S, Shofuda T, Klekner A, Zollo M, Leonard JR, Rubin JB, Jabado N, Albrecht S, Mora J, Van Meter TE, Jung S, Moore AS, Hallahan AR, Chan JA, Tirapelli DPC, Carlotti CG, Fouladi M, Pimentel J, Faria CC, Saad AG, Massimi L, Liau LM, Wheeler H, Nakamura H, Elbabaa SK, Perezpeña-Diazconti M, Chico Ponce de León F, Robinson S, Zapotocky M, Lassaletta A, Huang A, Hawkins CE, Tabori U, Bouffet E, Bartels U, Dirks PB, Rutka JT, Bader GD, Reimand J, Goldenberg A, Ramaswamy V, Taylor MD. Intertumoral Heterogeneity within Medulloblastoma Subgroups. Cancer Cell 2017; 31: 737-54.e6

3 Schwalbe EC, Lindsey JC, Nakjang S, Crosier S, Smith AJ, Hicks D, Rafiee G, Hill RM, Iliasova A, Stone T, Pizer B, Michalski A, Joshi A, Wharton SB, Jacques TS, Bailey S, Williamson D, Clifford SC. Novel molecular subgroups for clinical classification and outcome prediction in childhood medulloblastoma: a cohort study. Lancet Oncol 2017:

4 Northcott PA, Buchhalter I, Morrissy AS, Hovestadt V, Weischenfeldt J, Ehrenberger T, Gröbner S, Segura-Wang M, Zichner T, Rudneva VA, Warnatz H-J, Sidiropoulos N, Phillips AH, Schumacher S, Kleinheinz K, Waszak SM, Erkek S, Jones DTW, Worst BC, Kool M, Zapatka M, Jäger N, Chavez L, Hutter B, Bieg M, Paramasivam N, Heinold M, Gu Z, Ishaque N, Jäger-Schmidt C, Imbusch CD, Jugold A, Hübschmann D, Risch T, Amstislavskiy V, Gonzalez FGR, Weber UD, Wolf S, Robinson GW, Zhou X, Wu G, Finkelstein D, Liu Y, Cavalli FMG, Luu B, Ramaswamy V, Wu X, Koster J, Ryzhova M, Cho Y-J, Pomeroy SL, Herold-Mende C, Schuhmann M, Ebinger M, Liau LM, Mora J, McLendon RE, Jabado N, Kumabe T, Chuah E, Ma Y, Moore RA, Mungall AJ, Mungall KL, Thiessen N, Tse K, Wong T, Jones SJM, Witt O, Milde T, von Deimling A, Capper D, Korshunov A, Yaspo M-L, Kriwacki R, Gajjar A, Zhang J, Beroukhim R, Fraenkel E, Korbel JO, Brors B, Schlesner M, Eils R, Marra MA, Pfister SM, Taylor MD, Lichter P. The whole-genome landscape of medulloblastoma subtypes. Nature 2017; 547: 311-7

5 Sharma T, Schwalbe EC, Williamson D, Sill M, Hovestadt V, Mynarek M, Rutkowski S, Robinson GW, Gajjar A, Cavalli F, Ramaswamy V, Taylor MD, Lindsey JC, Hill RM, Jäger N, Korshunov A, Hicks D, Bailey S, Kool M, Chavez L, Northcott PA, Pfister SM, Clifford SC. Second-generation molecular subgrouping of medulloblastoma: an international meta-analysis of Group 3 and Group 4 subtypes. Acta Neuropathol 2019; 34: 396-18

6 Hendrikse LD, Haldipur P, Saulnier O, Millman J, Sjoboen AH, Erickson AW, Ong W, Gordon V, Coudière-Morrison L, Mercier AL, Shokouhian M, Suarez RA, Ly M, Borlase S, Scott DS, Vladoiu MC, Farooq H, Sirbu O, Nakashima T, Nambu S, Funakoshi Y, Bahcheli A, Diaz-Mejia JJ, Golser J, Bach K, Phuong-Bao T, Skowron P, Wang EY, Kumar SA, Balin P, Visvanathan A, Lee JJY, Ayoub R, Chen X, Chen X, Mungall KL, Luu B, Bérubé P, Wang YC, Pfister SM, Kim S-K, Delattre O, Bourdeaut F, Doz F, Masliah-Planchon J, Grajkowska WA, Loukides J, Dirks P, Fèvre-Montange M, Jouvet A, French PJ, Kros JM, Zitterbart K, Bailey SD, Eberhart CG, Rao AAN, Giannini C, Olson JM, Garami M, Hauser P, Phillips JJ, Ra YS, de Torres C, Mora J, Li KKW, Ng H-K, Poon WS, Pollack IF, López-Aguilar E, Gillespie GY, Van Meter TE, Shofuda T, Vibhakar R, Thompson RC, Cooper MK, Rubin JB, Kumabe T, Jung S, Lach B, Lolascon A, Ferrucci V, De

Antonellis P, Zollo M, Cinalli G, Robinson S, Stearns DS, Van Meir EG, Porrati P, Finocchiaro G, Massimino M, Carlotti CG, Faria CC, Roussel MF, Boop F, Chan JA, Aldinger KA, Razavi F, Silvestri E, McLendon RE, Thompson EM, Ansari M, Garre ML, Chico F, Eguía P, Pérezpeña M, Morrissy AS, Cavalli FMG, Wu X, Daniels C, Rich JN, Jones SJM, Moore RA, Marra MA, Huang X, Reimand J, Sorensen PH, Wechsler-Reya RJ, Weiss WA, Pugh TJ, Garzia L, Kleinman CL, Stein LD, Jabado N, Malkin D, Ayrault O, Golden JA, Ellison DW, Doble B, Ramaswamy V, Werbowetski-Ogilvie TE, Suzuki H, Millen KJ, Taylor MD. Failure of human rhombic lip differentiation underlies medulloblastoma formation. Nature 2022; 609: 1021-8

7 Korshunov A, Okonechnikov K, Stichel D, Schrimpf D, Delaidelli A, Tonn S, Mynarek M, Sievers P, Sahm F, Jones DTW, von Deimling A, Pfister SM, Kool M. Gene expression profiling of Group 3 medulloblastomas defines a clinically tractable stratification based on KIRREL2 expression. Acta Neuropathol 2022; 144: 339-52

8 Smith KS, Bihannic L, Gudenas BL, Haldipur P, Tao R, Gao Q, Li Y, Aldinger KA, Iskusnykh IY, Chizhikov VV, Scoggins M, Zhang S, Edwards A, Deng M, Glass IA, Overman LM, Millman J, Sjoboen AH, Hadley J, Golser J, Mankad K, Sheppard H, Onar-Thomas A, Gajjar A, Robinson GW, Hovestadt V, Orr BA, Patay Z, Millen KJ, Northcott PA. Unified rhombic lip origins of group 3 and group 4 medulloblastoma. Nature 2022; 609: 1012-20

9 Williamson D, Schwalbe EC, Hicks D, Aldinger KA, Lindsey JC, Crosier S, Richardson S, Goddard J, Hill RM, Castle J, Grabovska Y, Hacking J, Pizer B, Wharton SB, Jacques TS, Joshi A, Bailey S, Clifford SC. Medulloblastoma group 3 and 4 tumors comprise a clinically and biologically significant expression continuum reflecting human cerebellar development. Cell reports 2022; 40: 111162

10 Aldinger KA, Thomson Z, Phelps IG, Haldipur P, Deng M, Timms AE, Hirano M, Santpere G, Roco C, Rosenberg AB, Lorente-Galdos B, Gulden FO, O'Day D, Overman LM, Lisgo SN, Alexandre P, Sestan N, Doherty D, Dobyns WB, Seelig G, Glass IA, Millen KJ. Spatial and cell type transcriptional landscape of human cerebellar development. Nat Neurosci 2021; 24: 1163-75

11 Khouri-Farah N, Guo Q, Morgan K, Shin J, Li JYH. Integrated single-cell transcriptomic and epigenetic study of cell state transition and lineage commitment in embryonic mouse cerebellum. Sci Adv 2022; 8: eabl9156

12 Phoenix TN. The origins of medulloblastoma tumours in humans. Nature 2022; 609: 901-3

13 Rorke LB, Fogelson MH, Riggs HE. Cerebellar heterotopia in infancy. Dev Med Child Neurol 1968; 10: 644-50

14 Yachnis AT, Rorke LB, Trojanowski JQ. Cerebellar dysplasias in humans: development and possible relationship to glial and primitive neuroectodermal tumors of the cerebellar vermis. J Neuropathol Exp Neurol 1994; 53: 61-71

15 Werner B, Case J, Williams MJ, Chkhaidze K, Temko D, Fernandez-Mateos J, Cresswell GD, Nichol D, Cross W, Spiteri I, Huang W, Tomlinson IPM, Barnes CP, Graham TA, Sottoriva A. Measuring single cell divisions in human tissues from multi-region sequencing data. Nat Commun 2020; 11: 1035

16 Evano B, Khalilian S, Le Carrou G, Almouzni G, Tajbakhsh S. Dynamics of Asymmetric and Symmetric Divisions of Muscle Stem Cells In Vivo and on Artificial Niches. Cell Rep 2020; 30: 3195-206 e7

Figure 1 A: Scatterplot showing all genes ordered by rank and correlation (t-test statistic) with the G3/G4 score i.e., the transcriptional continuum. Colour represents a BH-adjusted p-value. All 6 genes highlighted by Korshunov *et al.* rank highly and are significantly positively correlated with *KIRREL2* ranking the highest. **B:** Scatterplot showing significant log-linear correlation ($p<0.001$) between *KIRREL2* expression (VST=Variance Stabilized Transform on a log scale) and G3/G4 score (continuum). Log-linear line of best fit is shown. **C:** Kaplan-Meier plot showing significant differences (Log-Rank test-for-trend) in $MB_{Grp3/Grp4}$ progression-free survival by *KIRREL2* expression (divided into 4 quartiles). **D:** Table of univariable Cox regression results showing relative risk (RR) of an event (progression/relapse or death) in $MB_{Grp3/Grp4}$ as a continuous function of G3/G4 score or expression of 6 genes highlighted by Korshunov *et al.* Note that relative risk for a continuous score is relative risk per unit of measurement and that the range of each variable is different i.e., G3/G4 score scaled between 0-1 hence the higher RR. **E.** mini-schema highlighting that transcription, subtype and risk of death are in large part fixed in utero prior to full transformation/tumorigenesis under the model proposed by Hendrikse *et al.* **F:** Conceptual schema highlighting two alternative approaches to assigning patients to treatment stratifications. "Subtype-based" stratification - largely the approach pursued at present - attempting to agglomerate DNA methylation subtypes into groups with similar average outcomes and subsequently to find relevant clinico-pathological characteristics that in combination place individuals into stratification windows (higher/intermediate/lower intensity) equate to a given average risk for the agglomerated groups. "Continuous Stratification" an alternative, perhaps more in keeping with the molecular biology, estimates a base level of 5-year risk as a continuous value derived from the transcriptional continuum (G3/G4 score), afterwards applying any independent risk factors as modifiers to the base risk (examples given, M+ = presence of metastases, whole chrom = whole chromosome aberration phenotype described by Goschzik et al, *MYC* amp = *MYC* Amplification). In this model, the treatment stratifications are fixed according to pre-defined risk windows. An individual's predicted 5-year risk rather than specifically their DNA methylation subtype is the key factor for treatment stratification.