Histopathology

Visual histological assessment of morphological features reflects the underlying molecular profile in invasive breast cancer: a morphomolecular study

Emad A Rakha,[1,*,†] Mansour Alsaleem,[1,*] Khloud A ElSharawy,[1] (iD) Michael S Toss,[1,†] (iD) Sara Raafat,[1] Raluca Mihai,[1] Fayyaz A Minhas,[2,†] Andrew R Green,[1] (iD) Nasir M Rajpoot,[2,†] Leslie W Dalton[3] & Nigel P Mongan[4,5]

[1]*School of Medicine, University of Nottingham, Nottingham City Hospital, Nottingham,* [2]*Department of Computer Science, University of Warwick, Coventry, UK,* [3]*Department of Histopathology, South Austin Hospital, Austin, TX,* [4]*Department of Pharmacology, Weill Cornell Medicine, New York, NY, USA, and* [5]*Faculty of Medicine and Health Sciences, School of Veterinary Medicine and Science, University of Nottingham, University of Nottingham Biodiscovery Institute, Nottingham, UK*

Rakha E A, Alsaleem M, ElSharawy K A, Toss M S, Raafat S, Mihai R, Minhas F A, Green A R, Rajpoot N M, Dalton L W & Mongan N P.

(2020) *Histopathology.* https://doi.org/10.1111/his.14199

Aims: Tumour genotype and phenotype are related and can predict outcome. In this study, we hypothesised that the visual assessment of breast cancer (BC) morphological features can provide valuable insight into underlying molecular profiles.

Methods and results: The Cancer Genome Atlas (TCGA) BC cohort was used (*n* = 743) and morphological features, including Nottingham grade and its components and nucleolar prominence, were assessed utilising whole-slide images (WSIs). Two independent scores were assigned, and discordant cases were utilised to represent cases with intermediate morphological features. Differentially expressed genes (DEGs) were identified for each feature, compared among concordant/discordant cases and tested for specific pathways. Concordant grading was observed in 467 of 743 (63%) of cases. Among concordant case groups, eight common DEGs (*UGT8, DDC, RGR, RLBP1, SPRR1B, CXorf49B, PSAPL1* and

SPRR2G) were associated with overall tumour grade and its components. These genes are related mainly to cellular proliferation, differentiation and metabolism. The number of DEGs in cases with discordant grading was larger than those identified in concordant cases. The largest number of DEGs was observed in discordant grade 1:3 cases (*n* = 1185). DEGs were identified for each discordant component. Some DEGs were uniquely associated with well-defined specific morphological features, whereas expression/co-expression of other genes was identified across multiple features and underlined intermediate morphological features.

Conclusion: Morphological features are probably related to distinct underlying molecular profiles that drive both morphology and behaviour. This study provides further evidence to support the use of image-based analysis of WSIs, including artificial intelligence algorithms, to predict tumour molecular profiles and outcome.

Keywords: breast, digital pathology, grade, molecular profiles, morphology

---

Address for correspondence: Professor E. A. Rakha, School of Medicine, University of Nottingham, Nottingham City Hospital, Nottingham NG7 2UH, UK. e-mail: mrzear1@exmail.nottingham.ac.uk
*Both these authors contributed equally to this study.
†The PathLAKE Consortium.

Introduction

The histological grade of breast cancer (BC) assessed using the Nottingham Grading System is one of the strongest prognostic factors in early stage BC.[1-4] It comprises the assessment of morphological features that represent the degree of similarity between the tumour and the normal breast parenchymal counterparts (i.e. degree of differentiation/de-differentiation) and the rate of tumour proliferation. It is well known that histological grade and tumour type reflect underlying molecular profiles, which are associated with distinct genomic features in BC.[5-7] Examples of such tumours include lobular breast carcinoma in which loss of *CDH1* gene function results in a discohesive growth pattern[8] and tall cell carcinoma with reverse polarity that has *isocitrate dehydrogenase 2* (*IDH2*) mutations.[9-11] Some BC types with unique histological features also show specific genomic alterations, including mucoepidermoid (*CRTC3–MAML2* fusion gene),[12] secretory (*ETV6–NTRK3* fusion gene)[13] and adenoid cystic (*MYB–NFIB* fusion gene)[14] carcinomas.

Gene expression profiling studies of BC have identified that molecular profile is strongly correlated with histological grade,[5] hormone receptor and human epidermal growth factor receptor 2 (HER2) status.[15] However, previous studies which attempted to link the morphology with the molecular profiles have focused on the association between the extreme morphological features (e.g. grades 1 and 3) as an overall measure of tumour differentiation and the expression of a specific set of genes. These gene sets were used to stratify tumours with borderline features (grade 2 tumours) into two subgroups. This binary approach was also applied to hormone receptor and HER2 status to stratify tumours into positive and negative for these receptors. In these studies, intermediate values for these bioindicators do not define intermediate tumour subsets.

Recently, there has been an increasing interest in leveraging the power of image analysis and artificial intelligence (AI) algorithms to identify the various morphological features of BC from digitalised haematoxylin and eosin (H&E) whole-slide images (WSIs) and to link these features to tumour behaviour, response to therapy or specific genomic profiles,[16] with varying degrees of accuracy.[17,18] Validation of such tools on large multicentric cohorts would allow the development of image-based tools to predict these variables in a cost-effective manner. Because the underlying molecular profiles represent the drivers of tumour behaviour and can predict response to therapy and determine tumour morphology and subtype,[19] assessment of these morphological features can be seen as a surrogate of the underlying molecular biology of the tumour. However, a detailed association between various morphological features and molecular profiles remains to be defined.

Using the Nottingham Grading System, pathologists assign BC grade using visual assessment of three morphological features including: (i) tubular differentiation: the spatial arrangement of the cells and whether they form tubules, as well as the proportion of tumour cells arranged in such well-formed tubular structures, (ii) nuclear pleomorphism: departure of cytonuclear features (such as size, shape and texture) from those of the normal ductal cell nuclei and (iii) mitotic count: the number of mitotic figures per 10 high-power fields (thresholds are adapted to account for field diameter as per the *World Health Organisation classification of tumours of the breast*[20]); this results in an overall three-tier grading scale. Final tumour grades are used to predict outcome and guide therapy.[3] Due to its inherent subjective nature, there is some degree of discordance between independent pathological assessment of histological grade features, which can result in some tumours being scored differently by individual trained pathologists.[3,21,22] However, it is possible that tumours that are most challenging to be assigned to a specific grade by all observers reflect intrinsically different biology and molecular make-up driving their borderline morphological features. Therefore, characterising the distinct molecular features of discordantly graded tumours may provide further insights into morphomolecular correlations. In addition, other specific morphological features with prognostic significance, such as nucleolar prominence,[23] need to be investigated to assess not only their correlation with genomic profiles, but also their relationship with other various morphological features of BC linked to differentiation, behaviour and tumour outcome.

In a previous study, we assessed the impact of BC grade discordance on patients' outcome.[24] In this study, we hypothesised that subjectivity in grade assignment of BC is related to the presence of borderline morphological features which are a reflection of their underlying genomic and molecular features. Akin to morphological features, the molecular profiles of BC represent a spectrum, with some tumours having a distinct molecular make-up and hence clearly defined morphology, while other tumours are in the borderline zones of these molecular profiles. These tumours are those which show less distinct morphological features and overlap between scoring grades. Deciphering the molecular profiles and genomic features of tumours with borderline morphological features could, at least in part, explain the discrepancies

in the level of concordance seen even among expert well-trained pathologists. This will provide further evidence to explore the use of computer vision and AI for assessing morphological features and to further develop algorithms to extrapolate many variables correlated with morphology.

In this study, we have re-assessed the BC cohort included in The Cancer Genome Atlas (TCGA) database for several morphological features in order to investigate the relationship between BC grade concordance/discordance, grade components and nucleolar prominence and the underlying molecular features.

Materials and methods

STUDY COHORT

A large cohort of BC cases ($n = 743$) from the TCGA data set[8] (cBioPortal.org) having both RNA sequencing (RNA-Seq) data and available digital H&E-stained WSIs scanned at ×40 magnification was used in this study. These data provided access to mRNA expression from RNASeqV2, along with identified clinico-pathological factors and outcome. This study focused on BC grade and nucleolar prominence as the main histological features assessed for studying the correlation with genomic profiles, whereas the histological subtype of BC cases included in this cohort (summarised in Supporting information, Table S1) was not considered in the analysis.

ORIGINAL GRADING

Original grading of WSIs was carried out by Heng *et al.*[19] (herein after referred to as the 'original grade'), where cases were randomly assigned to breast pathologists, and the WSIs were graded by referring to an electronic scoring sheet adapted from the College of American Pathologists' (CAP) protocol for BC grading.[20] Grading pathologists held conference calls to discuss the grading criteria; they then circulated images for scoring and images with high consensus diagnoses were used as examples for standardising grading. The nucleoli scores were assessed in this study and were assigned a score from 1 to 3 based on their prominence, as previously published.[23] Briefly, score 1 was given if nucleoli were inconspicuous and difficult to see at ×20 magnification. If the nucleoli were prominent and easily seen at ×10, or dysmorphic/multiple nucleoli were present, score 3 was assigned. Nucleolar score 2 was assigned to tumours with nucleoli not scored 1 or 3 (Supporting information, Figure S1A–C).

RE-SCORING

In this study, all cases were re-scored to reduce the impact of subjectivity in the assessment of various morphological features. Concordant cases are considered to represent cases with distinct morphological features, whereas discordant cases are likely to represent cases with intermediate features that are difficult to assign to one category. Re-grading of these WSIs (herein after referred to as 're-score') was carried out by an experienced breast pathologist (L.W.D.) who previously validated the use of WSIs grade assignment as a predictor of patient outcome.[24] He assigned a second tumour grade by using CAP grading criteria in BC.[20] This was compared to the results obtained during the original grading, and tumours were grouped based on the resulting concordant (grades 1:1, 2:2 and 3:3) and discordant (grades 1:2, 1:3 and 2:3) grade assignments, which represent cases with distinct morphological features and cases with border-line morphological features, respectively. Examples of discordant grades 1:2 and 2:3 are shown in Supporting information, Figure S1D,E.

Additionally, during re-scoring, other specific morphological features were recorded individually, such as the mitotic index and nucleolar prominence. For the standardised assessment of mitotic index using WSIs of the study cohort, mitotic figures were counted per mm^2. Nucleolar prominence was not initially scored as a separate feature in the original grade, even though it is implied in one of the grade components (nuclear pleomorphism). For the purposes of selecting concordant cases when evaluating this feature during re-score, it was assessed by two observers (L.W.D. and K.E.S.), as previously detailed.[23]

ASSESSING CONCORDANCE AND IDENTIFYING DIFFERENTIALLY EXPRESSED GENES (DEGS)

To expand our understanding of the molecular mechanisms underlying grade concordance and discordance that drive both morphology and outcome, differentially expressed genes (DEGs) were identified. To do this, a composite score was calculated based on the scores obtained for a grade, its components and nucleolar prominence, by the two observers as described earlier,[19,23] to define the concordance/discordance. To avoid bias, only concordant cases were used in the analysis to define the DEGs associated with specific grade, grade components (tubular differentiation, nuclear pleomorphism and mitotic count) and with nucleolar prominence. Gene expression data for> 20 000 genes generated by the TCGA BRCA

Table 1. Concordance rates - The tables show cross comparison (these are also called "confusion matrices" in the machine learning literature) of TCGA whole slide images assessed for grade (A), grade component scores (B–D) and nucleolar prominence score (E) between the original grade and re score

(A)

Grade (Original grade)	Grade (Re-score)			Total percentage	Kappa value
	Grade 1	Grade 2	Grade 3		
Grade 1	89	57	12	21.3%	0.43
Grade 2	72	179	54	41%	
Grade 3	6	75	199	37.7%	
Total Percentage	22.5%	41.9%	35.7%	100%	

Percent exact agreement: 63%, Percent adjacent level: 35%, Percent high:low: 2%

(B)

Tubular differentiation (Original grade)	Tubular differentiation (Re-score)			Total percentage	Kappa value
	Score 1	Score 2	Score 3		
Score 1	10	53	14	10.4%	0.44
Score 2	0	80	68	19.9%	
Score 3	3	41	474	69.7%	
Total Percentage	1.7%	23.4%	74.8%	100 %	

Percent exact agreement: 76%, Percent adjacent level: 22%, Percent high:low: 2%

(C)

Nuclear Pleomorphism (Original grade)	Pleomorphism (Re-score)			Total Percentage	Kappa value
	Score 1	Score 2	Score 3		
Score 1	23	30	5	7.8%	0.41
Score 2	68	168	99	45.1%	
Score 3	4	49	297	47.1%	
Total Percentage	12.8%	33.2%	54%	100 %	

Percent exact agreement: 66%, Percent adjacent level: 33%, Percent hi:low: 1%

(D)

Mitotic count (Original grade)	Mitotic count (Re-score)			Total Percentage	Kappa value
	Score 1	Score 2	Score 3		
Score 1	293	37	17	46.7%	0.35
Score 2	97	41	36	23.4%	
Score 3	41	72	109	29.9%	
Total Percentage	58%	20.2%	21.8%	100%	

Percent exact agreement: 60%, Percent adjacent level: 32%, Percent high:low: 8%

Table 1. *(Continued)*

(E)

Nucleolar prominence score* (1st observer)	Nucleolar Prominence score (2nd observer)			Total Percentage	Kappa value
	Score 1	Score 2	Score 3		
Score 1	180	47	3	40.5%	0.4
Score 2	55	96	9	28.2%	
Score 3	23	87	68	31.3%	
Total Percentage	45.4%	40.5%	14.1%	100%	

Percent exact agreement: 61%, Percent adjacent level: 35%, Percent hi:low: 4%

* The nuclear prominence is scored twice in this study by 2 scorers (1st scorer and 2nd scorer) as no original nucleolar score was available online.

Table 2. Differential gene expression analysis showing the numbers of genes expressed in association with the overall grade, grade components and nucleolar prominence (Analysis was restricted to the concordant cases)

Differentiation parameter	Total differentially expressed	Upregulated genes	Downregulated genes
Tubular differentiation	595	348	247
Nuclear pleomorphism	728	492	236
Mitotic count	620	543	77
Nucleolar prominence	352	250	102
Overall grade	519	361	158
Grade 1 versus (Grade 2 and Grade 3)	649	249	400*
Grade 2 versus (Grade 1 and Grade 3)	533	388	145
Grade 3 versus (Grade 1 and Grade 2)	589	301	288

*When these 400 down-regulated genes were investigated for enriched pathways, they were significantly enriched within the Integrin signalling pathway and the top master regulators were *COL2A1, COL11A2, COL9A3, DMBT1 and COL9A1*.

study were obtained and patients were stratified based on the composite score generated for each analysed component. DEGs were identified using the RobiNA implementation of the Edge R statistical tool.[25,26] Genes were filtered based on fold change ($> \pm2$) combined with *P*-value (<0.05). Common DEGs were identified between groups using the Venny version 2.0 tool (https://bioinfogp.cnb.csic.es/tools/venny/). These DEGs were further analysed comparatively between different subgroups of cases. The web-based gene set enrichment analysis tool (WebGestalt) was used to calculate significantly enriched pathways and gene ontologies (GO) based on the identified DEGs.[27-29]

As the TCGA data have a limited number of events (disease recurrence or related mortality), outcome analysis was computed using the Kaplan–Meier (KM) Plotter data set ($n = 1764$) as a validation for the prognostic value of the observed gene signature at the mRNA level.[30] The KM Plotter data set has 1764 cases with recurrence data, while the overall survival data was available for 626 patients only. The survival of patients was stratified by the collective mean

mRNA expression of the identified common DEGs. The best-performing threshold against outcome was used to categorise the cases into high and low risk as generated by the KM Plotter (determined by the public domain database) regardless of the tumour grade. No mutational analysis was performed in this study.

STATISTICAL ANALYSIS

BM SPSS version 24.0 (SPSS, Inc., IBM Corp, Chicago, IL, USA) software was used for statistical analysis. The concordance rate for assigning tumour grade and for estimating different grade components between observers was assessed using the kappa test.

## Results

CONCORDANCE ASSESSMENT

Concordant grading was observed in 63% of cases (grades 1:1 in 12%, 2:2 in 24% and 3:3 in 27%), whereas grade discordance was observed in the remaining 36% cases (grades 1:2 in 17%, 2:3 in 17%

and 1:3 in 2%). In terms of morphological features, concordance rates were largely similar: 76% for tubular differentiation, 66% for nuclear pleomorphism, 60% for mitotic count and 61% for nucleolar prominence (Table 1). The kappa value for concordance between the two grade assignment sessions (original grade and re-grade) was 0.43, while for the assessment of each morphological feature kappa values were as follows: 0.44 for tubular differentiation, 0.41 for nuclear pleomorphism, 0.35 for mitotic count and 0.4 for nucleolar prominence. These values mainly show a moderate level of agreement.

TRANSCRIPTOMIC ANALYSIS OF CONCORDANCE AND GRADE COMPONENTS

The data analysis for the individual morphological features assessed in concordant cases (tubular differentiation, nuclear pleomorphism, mitotic count and nucleolar prominence) identified the following DEG sets (Table 2):
1. DEGs (595) associated with tubular differentiation, where 348 genes were significantly up-regulated with prominent tubular differentiation.
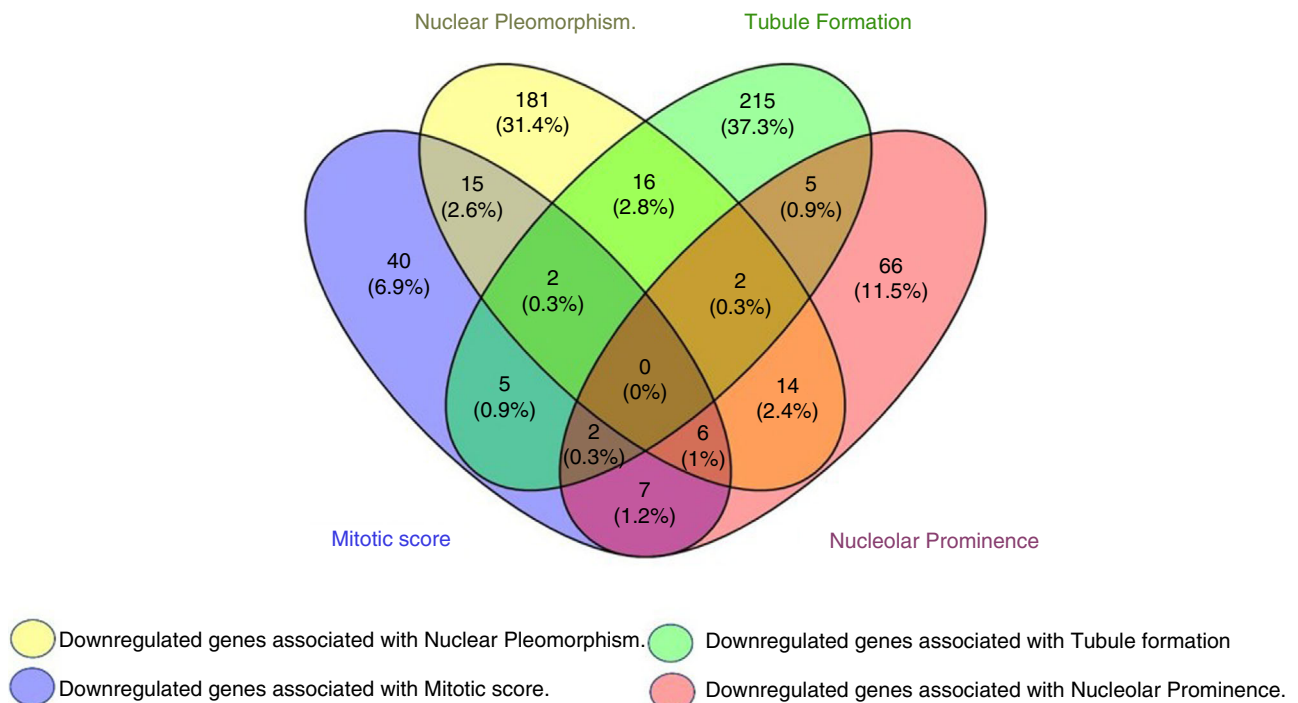
Figure 1. The overlap between the down-regulated differentially expressed genes (DEGs) associated with each of grade component in relation to different grade component (investigated using the Venny tool). The overlap show sets of genes included exclusively in each feature alone, and in combination, for 40 DEGs included exclusively in mitoses, 15 common DEGs in mitoses and pleomorphism, two common DEGs in mitoses, pleomorphism and tubule formation, no common DEGs in mitoses, pleomorphism, tubule formation and nucleolar prominence.

2. DEGs (728) associated with nuclear pleomorphism, where 492 genes were significantly up-regulated with high degree of nuclear pleomorphism.
3. DEGs (620) associated with mitotic count, where 543 genes were significantly up-regulated with high mitotic count.
4. DEG (352) associated with nucleolar prominence, where 250 genes were significantly up-regulated with higher nucleolar prominence score.

Sixteen genes were commonly up-regulated across the four features, whereas no common genes were down-regulated across the four morphological components assessed (Figures 1 and 2). The DEGs associated with the tumour grade identified 361 genes which were significantly up-regulated (Table 2). By overlapping the DEGs associated with all four morphological features (*n* = 16) with the DEGs associated with the tumour grade (*n* = 361), we identified eight core common genes significantly up-regulated. The Web-Gestalt over-representation analysis tool (ORA) was used to perform GO biological process analysis for the eight significantly up-regulated common genes; this indicated that some up-regulated genes (*PSAPL1*,

SPRR1B and *SPRR2G*) were involved in the epithelial cell differentiation pathways, whereas other genes were involved in other related pathways, such as the sphingolipid metabolic process (*PSAPL1* and *UGT8*) and the alkaloid metabolic process (*DDC*). Figures 3 and 4 and Tables 3 and 4 show details of these up-regulated genes.

ASSOCIATION OF GENE SIGNATURE WITH PATIENTS' OUTCOME

Within the TCGA cohort; 20% of cases showed high expression of the eight-gene signature (146 of 743). To evaluate the clinical value of the eight morphology-associated DEGs, we tested them against BC patient outcome using the KM Plotter database using default settings with 1764 cases having recurrence data and 626 with overall survival data.[30] High expression of the eight-gene signature was associated with shorter overall and recurrence-free survival ($P = 0.003$ and $P = 0.00016$, respectively; Figure 5), which confirmed the link between morphology, underlying molecular profiles and outcome.
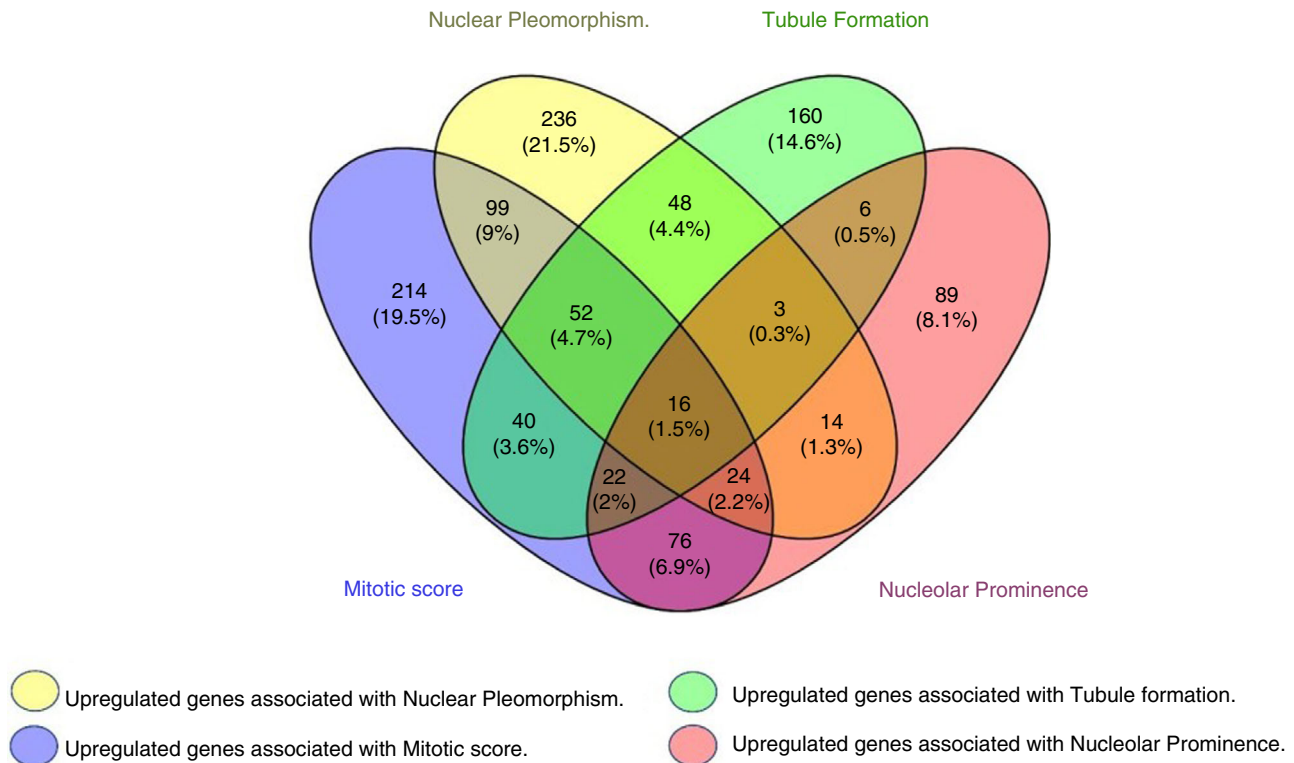
Figure 2. The overlap between the up-regulated differentially expressed genes (DEGs) associated with each morphological feature (investigated using the Venny tool). The overlap, for instance, shows 236 DEGs included exclusively in nuclear pleomorphism, 214 DEGs included exclusively in mitotic count and 160 DEGs included exclusively in tubular differentiation.
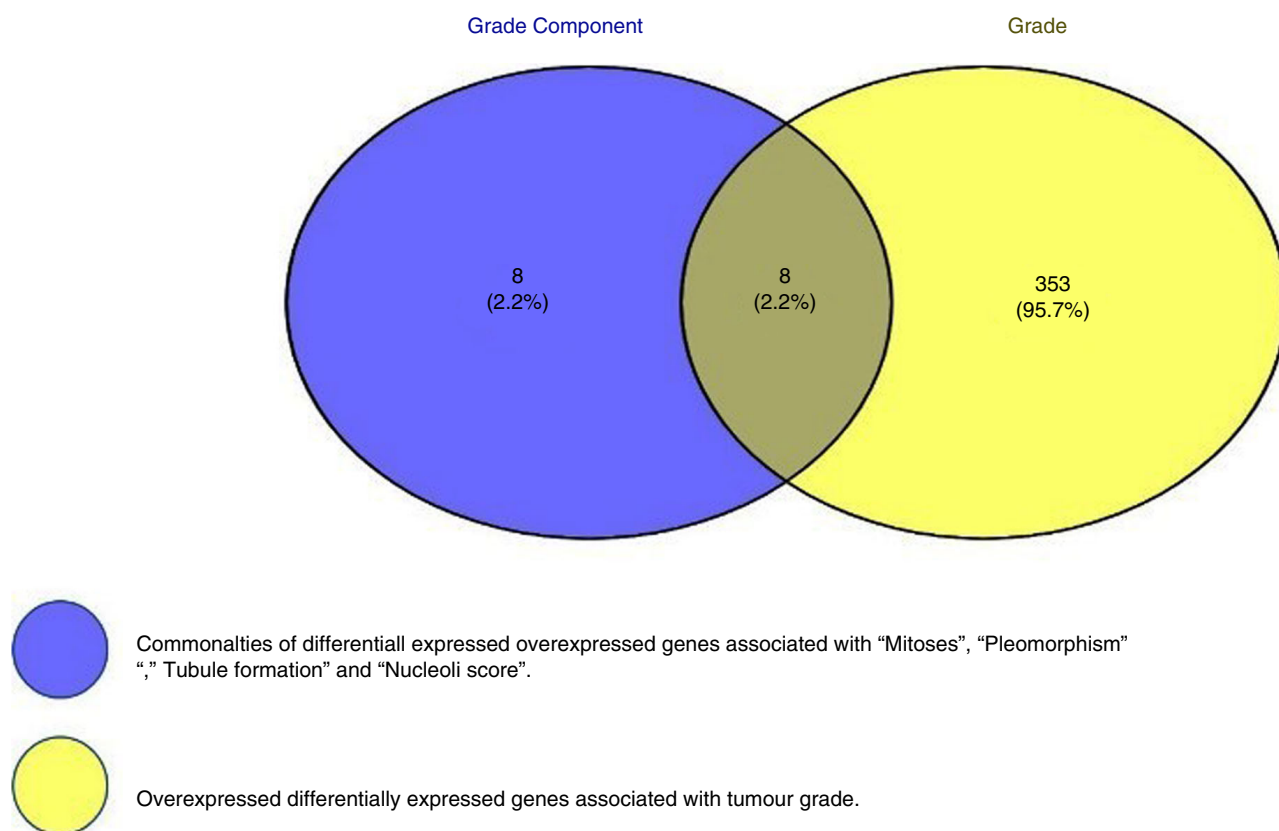
Figure 3. The overlap between the up-regulated differentially expressed genes (DEGs) associated with grade component in relation to overall tumour grade up-regulated DEGs were investigated using the Venny tool. The overlapping has shown eight common up-regulated DEGs in 'grade component' and 'grade'.

BIOINFORMATICS ANALYSIS OUTPUT OF GRADE DISCORDANT CASES

DEGs of grade discordant cases revealed 877 genes that distinguished grades 1:2 from 1:1 tumours, of which 491 genes were significantly up-regulated, including genes associated with the integrin signalling pathway. Furthermore, there were 1558 genes that differentiated grades 1:3 from 1:1 tumours, involving 768 significantly up-regulated genes including those associated with the Beta1 adrenergic receptor signalling pathway. Similarly, there were 955 genes that differentiated tumour grades 1:2 from 2:2 tumours with 362 up-regulated genes and including genes significantly associated with inflammatory pathways mediated by the chemokine and cytokine signalling pathway.

There were 1460 genes that distinguished grades 2:3 from 2:2 tumours having 986 up-regulated genes and including genes significantly enriched in the plasminogen activating pathway. Finally, there were 2346 DEGs that differentiated between grades 2:3

and 3:3 tumours, of which 848 up-regulated genes were enriched in the heterotrimeric G-protein signalling pathway-Gq alpha and Go alpha-mediated pathway. These results are summarised in Table 5, Supporting information, Table S2 and illustrated in Figure 6, which is mapped to histological examples of concordant and discordance cases.

Discussion

Expert visual assessment of tumour morphological features remains the primary approach to diagnose and predict BC outcome. The correlation between morphology and genomic features of BC is also well documented, and there are several lines of evidence demonstrating that the distinct behaviour, aggressiveness and response to therapy of histological subtypes of BC are related to distinct molecular alterations.[31-37] In the era of digital pathology, computer-aided image analysis and image-based AI tools, there is the potential to more accurately relate cellular
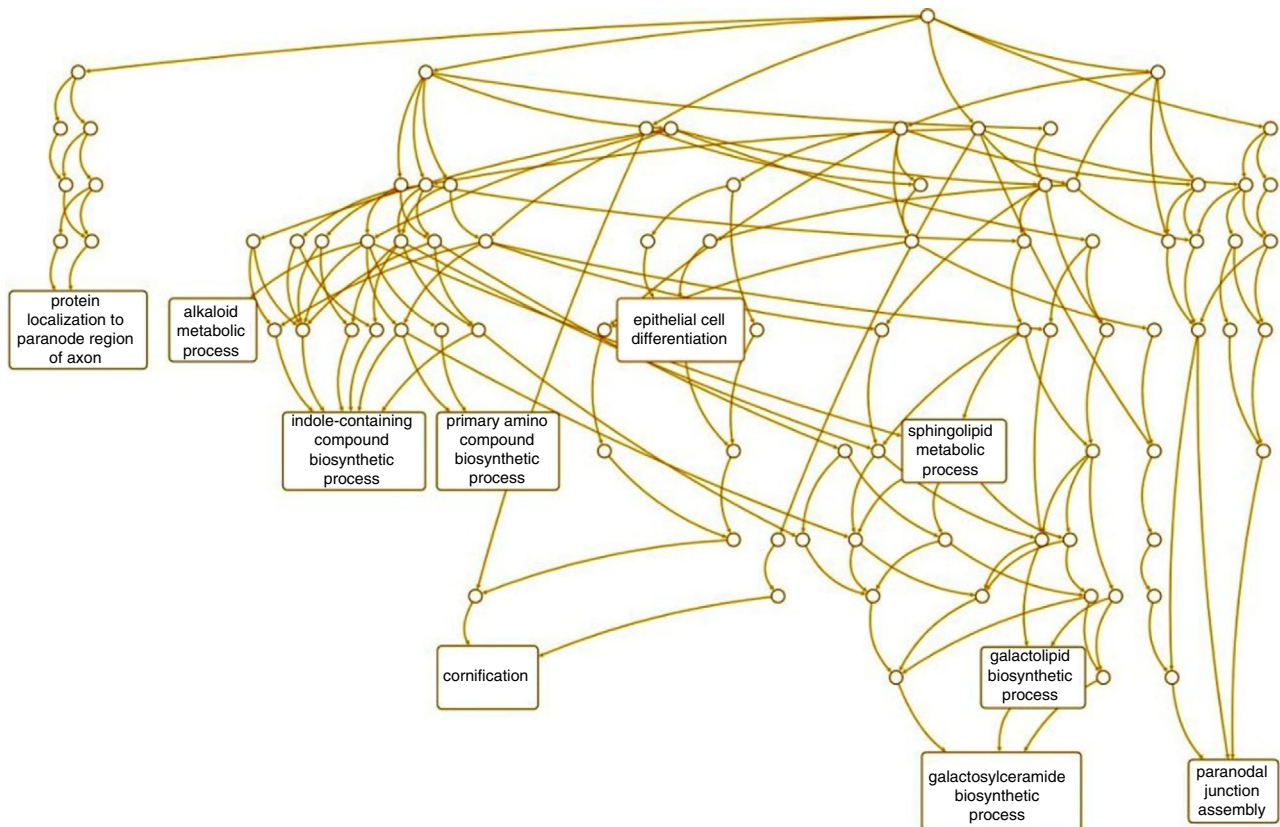
Figure 4. Pathway analysis using category 'biological process' of gene ontology for the eight genes with significant up-regulation across all specific morphological features investigated (grade components and nucleolar prominence) and overall tumour grade. Cellular differentiation, metabolism and proliferation are the main ontologies observed.

morphology and histology to the underlying molecular changes within tumours. Kather and colleagues have recently shown that it is possible to predict the microsatellite instability of gastrointestinal tumours by using AI-based morphological features.[17] In BC, Couture *et al.* also showed that an AI model can predict the oestrogen receptor (ER) status with more than 75% accuracy using the histological features alone.[18] Associations between morphological features and patients' outcome are strong in BC and, as such, tumour grade is commonly used to inform treatment decisions.[3] However, concordance among pathologists in assessing such morphological features is not perfect.[22] In our study, discordance between assessors was most common in the assessment of mitotic count and nucleolar prominence. The subjectivity in identifying the relevant regions within WSIs to assess these features and the challenges inherent in differentiating mitotic figures from other similar structures, such as apoptosis, were probably the main reason for the discordance, as previously noted.[23,38] In addition, the time interval between case assessments (sometimes

referred to as the 'wash-out period') and the large number of cases included in the study may have contributed discordant assessment in some cases, including the rare examples of extreme discordance, such 1:3 and 3:1 grade assignment.

As one of the main aims of this study was to relate specific morphological features with their underlying molecular characteristics, only cases with concordant scoring were considered in the analysis for such correlations. Discordant groups were used to assess the molecular features of cases with intermediate morphological features. Unlike our previous study of BC grade concordance that utilised WSIs of cases assessed at ×20 magnification,[24] the slides used in the current study were assessed at ×40 magnification. However, the proportion of cases with grade concordance/discordance was similar in both studies, implying that the scanner magnification has a limited contribution to the grade concordance.

In a previous study, we demonstrated that cases with grade discordance were associated with distinct outcomes and suggested that BC grade discordance is

Table 3. The Biological process category by the Gene Ontology Consortium shows the common biological processes that are significantly associated with grade, grade components and nucleolar prominence

| Gene Set	Description	Size/Overlap*	Expect[†]	Ratio/P-value[‡]	Gene symbol
GO:0070268	Cornification	111/ 2	0.047	42.89/0.001	*SPRR1B, SPRR2G*
GO:0030855	Epithelial cell differentiation	747/3	0.314	9.56/0.003	*PSAPL1, SPRR1B, SPRR2G*
GO:0042435	Indole-containing compound biosynthetic process	5/1	0.002	476.11/0.002	*DDC*
GO:1901162	Primary amino compound biosynthetic process	5/1	0.002	476.11/0.002	*DDC*
GO:0009820	Alkaloid metabolic process	7/1	0.004	340.08/0.003	*DDC*
GO:0006665	Sphingolipid metabolic process	153/2	0.064	31.12/0.002	*PSAPL1, UGT8*
GO:0002175	Protein localization to paranode region of axon	5/1	0.002	476.11/0.002	*UGT8*
GO:0006682	Galactosylceramide biosynthetic process	6/1	0.002	396.76/0.002	*UGT8*
GO:0019375	Galactolipid biosynthetic process	6/1	0.002	396.76/0.002	*UGT8*
GO:0030913	Paranodal junction assembly	7/1	0.004	340.08/0.003	*UGT8*

*The category size is calculated based on the number of overlapping genes between the annotated genes in the category and the reference gene list for the "ORA" method. The category overlap is the overlapping between the genes in the input and those in the database.
[†]The category expect is the number of categories expected from set cover" indicates the number of the expected reduced sets of the weighted set cover algorithm for redundancy reduction in the report.
[‡]The category ratio is the enrichment ratio (for ORA) as it indicates the gene ontology sets with FDR < 0.05. The category *P* value is indicative of the weighted set cover and maximum coverage called size-constrained weighted set cover where weights are assigned to gene sets with smaller enrichment *P* values.

Table 4. Differential gene expression analysis of the overall grade, grade components and nucleolar prominence detailing the 8 most common upregulated genes

	Gene description	Gene ID	Function
1	UDP glycosyltransferase 8	*UGT8*	Catalyses the transfer of galactose to ceramide, a key enzymatic step in the biosynthesis of galactocerebrosides.
2	Dopa decarboxylase	*DDC*	Catalyses the decarboxylation of L-3,4- dihydroxyphenylalanine (DOPA) to dopamine, L-5-hydroxytryptophan to serotonin and L-tryptophan to tryptamine.
3	Retinal G protein coupled receptor	*RGR*	Receptor for all-trans- and 11-cis-retinal. Binds preferentially to the former and may catalyse the isomerization of the chromophore by a retinochrome-like mechanism; Opsin receptors.
4	Retinaldehyde binding protein 1	*RLBP1*	Soluble retinoid carrier essential for the proper function of both rod and cone photoreceptors.
5	Small proline rich protein 1B	*SPRR1B*	Function as both amine donor and acceptor in transglutaminase-mediated cross-linkage; Belongs to the cornify (SPRR) family.
6	Chromosome X open reading frame 49B	*CXorf49B*	Uncharacterised protein CXorf49.
7	Prosaposin like 1 (gene: pseudogene)	*PSAPL1*	May activate the lysosomal degradation of sphingolipids.
8	Small proline rich protein 2G	*SPRR2G*	A keratinocyte protein that first appears in the cell cytosol, but ultimately becomes cross-linked to membrane proteins by transglutaminase. All that results in the formation of an insoluble envelope beneath the plasma membrane.
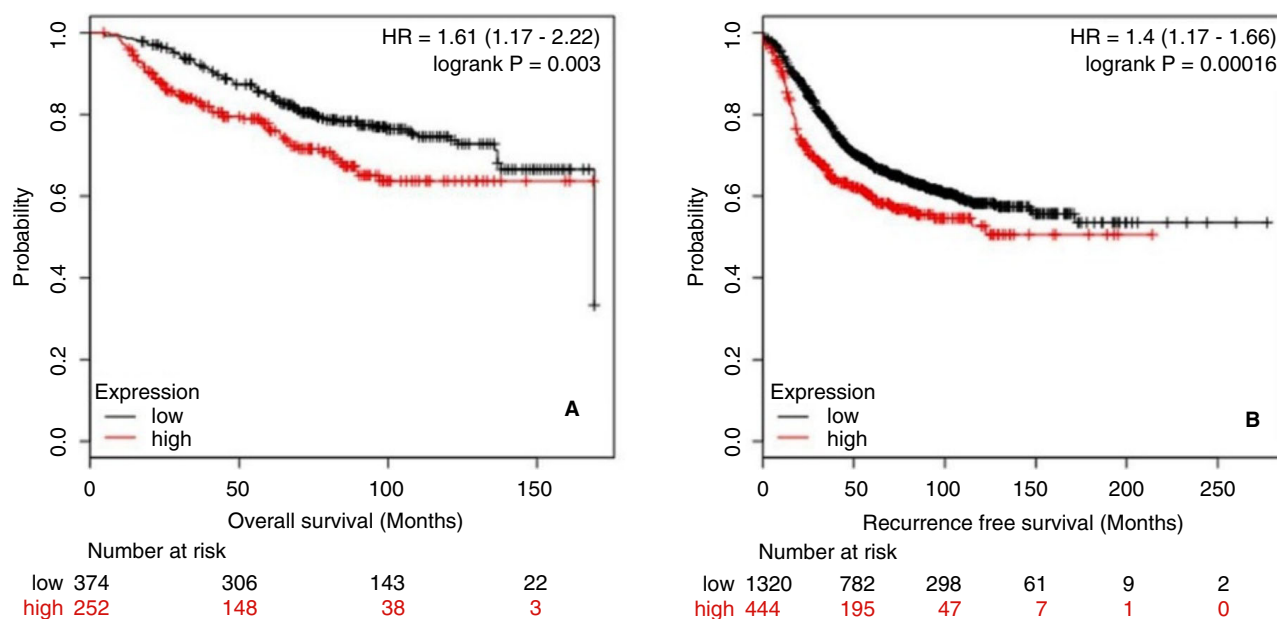
Figure 5. Kaplan–Meier survival plot showing association of the identified gene signature (patients were stratified by the cumulative mean mRNA expression of the identified eight-gene signature) with overall survival (**A**) and recurrence-free survival (**B**). **A,** The survival was significantly worse in 252 patients harbouring high expression of the identified eight-gene signature in comparison to 374 patient harbouring lower expression; **B,** 444 patients harbouring high expression of the identified eight-gene signature had worse recurrence-free survival in comparison to 1320 patients harbouring lower expression of the gene signature.

probably a reflection of biologically, and hence morphologically, distinct tumours.[24] In this study, we aimed to assess the impact of the underlying molecular profiles on tumour morphology, the ability of pathologists to assess them and to evaluate whether cases with discordant grading have distinct molecular profiles.

When the concordant/discordant cases are taken into account, BC grade could be regarded as a five-category risk scale[24] [three concordant (1:1, 2:2, 3:3) and two discordant (1:2 and 2:3) grade categories] that are associated with specific genomic/molecular profiles. Grade 1 concordant cases represent the very well-differentiated, lowest-risk group of cancers at one end of the differentiation continuum, whereas grade 3 concordant cases are the least-differentiated tumours at the other end of the spectrum. In this study, we used the TCGA BC cohort, which is a publicly available resource that has WSIs for a large cohort of BC together with their linked molecular, genetic and de-identified clinical data, to assess the molecular differences between cases that were concordantly/discordantly graded between experienced breast pathologists. As expected, the number of DEGs in concordant cases is much smaller than the number of DEGs seen with discordant cases. This probably reflects that concordant cases possess

unambiguously similar morphological features, whereas discordant cases harbour subtle distinct morphological features. It also supports the hypothesis that discordant cases reflect intrinsically, although subtly, different tumour morphological features and, by inference, different tumour molecular characteristics, and that discordance is not simply the result of subjectivity of eyeballing assessment of such morphological features.

Regarding the genes that were differentially expressed in association with various morphological features used as grade components, a significant overlap between genes associated with nuclear pleomorphism, nucleolar prominence and cellular proliferation was identified. This could reflect the underlying biology of BC in terms of tumour differentiation and proliferation and provide insight regarding how this is reflected in the morphological features recorded with the conventional grading of BC. The results of this study indicated that grade 1 tumours were enriched with the integrin signalling pathway, which is important for the functional differentiation of the epithelia. Grade 2 tumours were enriched with the heterotrimeric G-protein signalling pathway-Gi alpha and Gs alpha-mediated pathways, which regulate a range of endothelial cell functions including differentiation, proliferation and migration, with the top
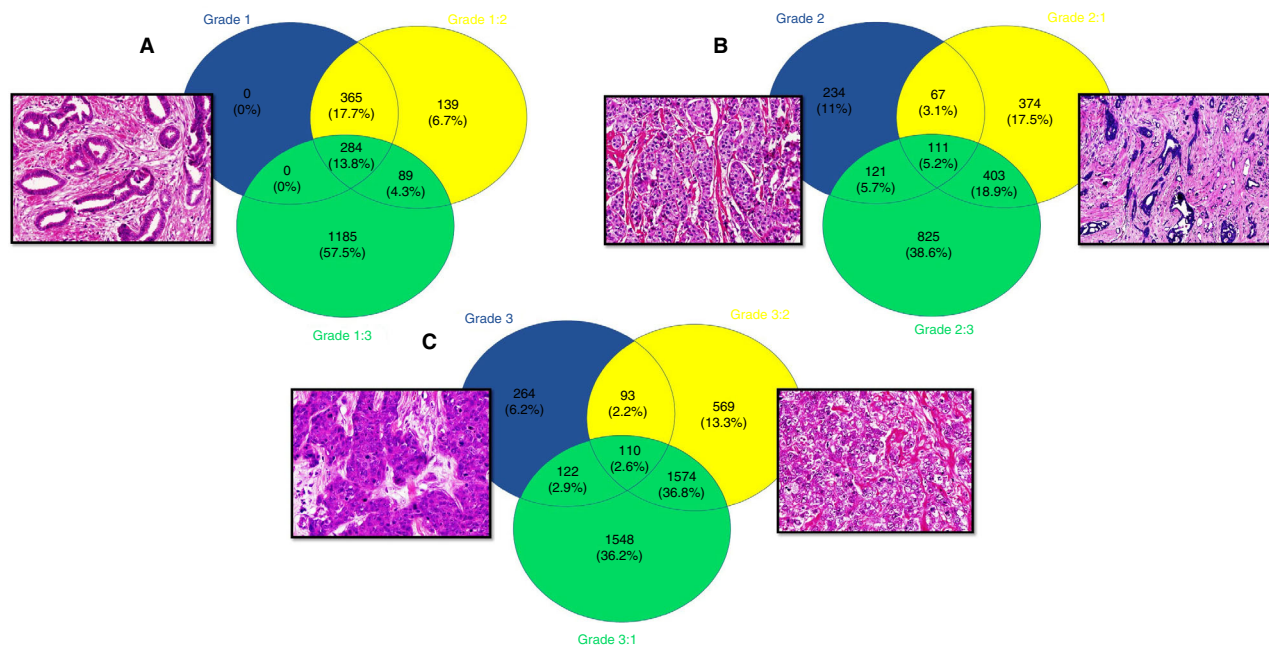
Table 5. Number of differentially expressed genes and the top regulated pathway associated with grade concordance/discordance as defined by both differential gene expression and pathway analyses.[28]

Differentiation parameter	Total DEG	Top regulated pathway
Unique genes underlying concordance grades		
Grade 1 versus (Grade 2 and Grade 3)	649	Integrin signalling pathway
Grade 2 versus (Grade 1 and Grade 3)	533	Heterotrimeric G-protein signalling pathway-Gi alpha and Gs alpha mediated pathway
Grade 3 versus (Grade 1 and Grade 2)	589	Wnt signalling pathway
Unique genes underlying discordance grades		
Grade 1:2 versus Grade 1	877	Integrin signalling pathway
Grade 1:3 versus Grade 1	1558	Beta1 adrenergic receptor signalling pathway
Grade 2:1 versus Grade 2	955	inflammatory pathways mediated by chemokine and cytokine signalling pathway
Grade 2:3 versus Grade 2	1460	plasminogen activating pathway
Grade 2:3 versus Grade 3	2346	Heterotrimeric G-protein signalling pathway-Gi alpha and Gs alpha mediated pathway
Common genes underlying concordance and discordance grades		
Grade 1 versus Grade 1:2	365	Nicotine degradation pathway
Grade 1 versus Grade 1:3	0	–
Grade 1:2 versus Grade 1:3	89	Adenine and hypoxanthine salvage pathway
Grade 1 versus Grade 1:2 versus Grade 1:3	284	Adrenaline and noradrenaline
Grade 2 versus Grade 2:1	67	Vitamin D metabolism pathway
Grade 2 versus Grade 2:3	121	Arginine biosynthesis pathway
Grade 2 versus Grade 2:1 versus Grade 2:3	111	Plasminogen activating cascade pathway
Grade 2:1 versus Grade 2:3	403	Heterotrimeric G-protein signalling pathway-rod outer segment phototransduction pathway
Grade 3 versus Grade 3:2	93	Heterotrimeric G-protein signalling pathway-Gi alpha and Gs alpha mediated pathway
Grade 3 versus Grade 3:1	122	Muscarinic acetylcholine receptor 1 and 3 signalling pathway
Grade 3 versus Grade 3:2 versus Grade 3:1	110	Inflammatory pathways mediated by chemokine and cytokine signalling pathway
Grade 3:2 versus Grade 3:1	1574	Nicotine degradation pathway

master regulator genes being *CBP2*, *FGA*, *FGB*, *FGG* and *MMP1*. Finally, grade 3 tumours were enriched with the Wnt signalling pathway whose overactivation triggers the oncogenic transformation and proliferation of many cancers, including triple-negative breast cancers. Some pathways, such as the Wnt-, CCKR- and cadherin-signalling pathways, were involved in driving multiple morphological features of differentiation, whereas other pathways are more frequently activated in the discordant cases (those with borderline features), such as the plasminogen activating cascade, nicotine degradation and iono-trophic glutamate receptor pathways. Moreover, the main pathways detected with grade-discordant cases can be summarised under two main categories which are related to cellular proliferation and differentiation. Our study also identified a gene signature comprising eight up-regulated genes overlapped between the four

Figure 6. Overlapping differentially expressed genes (DEGs) between concordant and discordant cases. **A,** Overlapping DEGs between grades 1:1, 1:2 and 1:3 groups of cases. **B,** Overlapping DEGs between grades 2:2, 2:1 and 2:3 groups of cases. **C,** Overlapping DEGs between grades 3:3, 3:2 and 3:1 groups of cases. Examples of histological images of concordant and discordant cases are mapped adjacent to each category.

morphological features assessed. Up-regulation of one of this gene signature (UGTB; the endoplasmic reticulum-localised enzyme UDP-galactose:ceramide galactosyltransferase), was related to poor prognosis and was strongly associated with grade 3 BC.[39]

Consistent with the differences in tumour behaviour and outcome associated with distinct tumour grades, pathway analysis revealed that the DEGs identified are related to cellular proliferation and differentiation. Our findings also showed that the discordant cases in the grade categories 1:2 or 2:3 had more complex transcriptional profiles than concordant cases, and that the highest number of DEGs were seen in tumours showing the extreme ends of discordance (tumour with grades 1:3 or 3:1). This supports the hypothesis that the distinct morphological features assessed in BC grading are mirroring underlying molecular mechanisms. As the molecular make-up of tumours rather than the morphology *per se* is the driver of the behaviour, the associations between morphology and outcome should be explained by the underlying molecular profile which determines both, as previously described.[24] This is supported by our findings, where the cases that showed high expression of the DEGs associated with grade, and its various components, had a worse outcome than those with lower expression.

Interestingly, this study identified specific pathways and gene sets that are associated with specific morphological features. We also reported that some of these alterations may result in borderline morphological features that are difficult to be robustly assigned to a grade category when using visual assessment alone. Although the current study has limitations, it provides a proof of principle and further evidence to support the use of image analysis methods, including AI-based tools, in BC diagnosis, the prediction of tumour outcome and response to therapy. Validation of these results using independent cohorts and emerging computational approaches will provide more insights into the pathways associated with each histological feature.

FUTURE PERSPECTIVES

Developers of image-based AI tools should be aware of the challenges of grade assignment, and in particular note that existing grade categories elide subtle distinct morphological features that are associated with distinct molecular characteristics and outcomes. Our understanding of this diversity of molecular phenotypes resulted from relatively recently developed next-generation sequencing technologies, as their wider application to large BC patient cohorts has enabled a better understanding of the complex biology

underlying these histological features. Indeed, the further results obtained here highlighted how informative is tumour morphology. With advances in the application of image-based AI and machine-learning techniques to histopathology, it should be possible to infer underlying genetic and transcriptional phenotypes from morphological features and thereby more accurately guide therapies. Thus, AI-based histological analyses are likely to build upon this current study to enable superior patient stratification and more accurately predict molecular status and clinical outcomes, as such approaches will leverage the integration of large complex molecular data sets and pixel-level image analysis with the identification of morphological and architectural features not discernible during visual histological assessment.

Conclusion

This study shows that the underlying molecular alterations in BC are reflected in the morphological features of tumours and that grade-discordant cases have different genetic signatures when compared with the grade-concordant ones. It should also show that these underlying molecular differences have a role in disease outcome. Such observations could be used to understand the biology of BC of various grades, and could also provide a tool to better predict the behaviour of these tumours from their morphological features.

Acknowledgements

AUTHORS' CONTRIBUTIONS

Conception and design: EAR, LWD. Cohort Identification, Collection and assembly of data: MT, SR, AG, MAl. Data analysis and interpretation: EAR, MAl, KE, MT, SR, RM, FA, AG, NMR, LWD, EAR. Manuscript writing: EAR, MAl, KE, MT, SR, RM, FA, AG, NMR, LWD, EAR. All authors contributed to revise and approve the final version of the manuscript.

ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

This work obtained ethics approval by the North West – Greater Manchester Central Research Ethics Committee under the title; Nottingham Health Science Biobank (NHSB), reference number 15/NW/0685. All patients included were consented to participate in the study and to use their materials in research. All samples from Nottingham used in this study were pseudo-anonymised and stored in compliance with the UK Human Tissue Act. The study was performed in accordance with the Declaration of Helsinki.

COMPETING INTEREST

The authors declare no conflict of interest.

References

1. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991; **19**; 403–410.
2. Rakha EA, El-Sayed ME, Lee AH *et al.* Prognostic significance of nottingham histologic grade in invasive breast carcinoma. *J. Clin. Oncol.* 2008; **26**; 3153–3158.
3. Rakha EA, Reis-Filho JS, Baehner F *et al.* Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Res.* 2010; **12**; 207.
4. Rakha EA, El-Sayed ME, Menon S, Green AR, Lee AH, Ellis IO. Histologic grading is an independent prognostic factor in invasive lobular carcinoma of the breast. *Breast Cancer Res. Treat.* 2008; **111**; 121–127.
5. Sotiriou C, Wirapati P, Loi S *et al.* Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J. Natl Cancer Inst.* 2006; **98**; 262–272.
6. Yu K, Lee CH, Tan PH *et al.* A molecular signature of the Nottingham Prognostic Index in breast cancer. *Cancer Res.* 2004; **64**; 2962–2968.
7. Roylance R, Gorman P, Harris W *et al.* Comparative genomic hybridization of breast tumors stratified by histological grade reveals new insights into the biological progression of breast cancer. *Cancer Res.* 1999; **59**; 1433–1436.
8. Ciriello G, Gatza ML, Beck AH *et al.* Comprehensive molecular portraits of invasive lobular breast cancer. *Cell* 2015; **163**; 506–519.
9. Bhargava R, Florea AV, Pelmus M *et al.* Breast tumor resembling tall cell variant of papillary thyroid carcinomaa solid papillary neoplasm with characteristic immunohistochemical profile and few recurrent mutations. *Am. J. Clin. Oncol.* 2017; **147**; 399–410.
10. Chiang S, Weigelt B, Wen HC *et al.* IDH2 mutations define a unique subtype of breast cancer with altered nuclear polarity. *Cancer Res.* 2016; **76**; 7118–7129.
11. Lozada JR, Basili T, Pareja F *et al.* Solid papillary breast carcinomas resembling the tall cell variant of papillary thyroid

neoplasms (solid papillary carcinomas with reverse polarity) harbor recurrent mutations affecting IDH2 and PIK3CA: a validation cohort. *Histopathology* 2018; **73**; 339–344.

12. Nakayama T, Miyabe S, Okabe M *et al*. Clinicopathological significance of the CRTC3–MAMl2 fusion transcript in mucoepidermoid carcinoma. *Mod. Pathol.* 2009; **22**; 1575–1581.

13. Makretsov N, He M, Hayes M *et al*. A fluorescence *in situ* hybridization study of ETV6–NTRK3 fusion gene in secretory breast carcinoma. *Genes Chromosomes Cancer* 2004; **40**; 152–157.

14. Brill LB II, Kanner WA, Fehr A *et al*. Analysis of MYB expression and MYB–NFIB gene fusions in adenoid cystic carcinoma and other salivary neoplasms. *Mod. Pathol.* 2011; **24**; 1169–1176.

15. Lu X, Lu X, Wang ZC, Iglehart JD, Zhang X, Richardson AL. Predicting features of breast cancer with gene expression patterns. *Breast Cancer Res. Treat.* 2008; **108**; 191–201.

16. Coudray N, Ocampo PS, Sakellaropoulos T *et al*. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat. Med.* 2018; **24**; 1559–1567.

17. Kather JN, Pearson AT, Halama N *et al*. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat. Med.* 2019; **25**; 1054–1056.

18. Couture HD, Williams LA, Geradts J *et al*. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. *NPJ Breast Cancer* 2018; **4**; 30.

19. Heng YJ, Lester SC, Tse GM *et al*. The molecular basis of breast cancer pathological phenotypes. *J. Pathol.* 2017; **241**; 375–391.

20. Lester SC, Bose S, Chen YY *et al*. Protocol for the examination of specimens from patients with invasive carcinoma of the breast. *Arch. Pathol. Lab. Med.* 2009; **133**; 1515–1538.

21. Rakha EA, Ahmed MA, Aleskandarany MA *et al*. Diagnostic concordance of breast pathologists: lessons from the National Health Service Breast Screening Programme Pathology External Quality Assurance Scheme. *Histopathology* 2017; **70**; 632–642.

22. Rakha EA, Bennett RL, Coleman D, Pinder SE, Ellis IO. Review of the national external quality assessment (EQA) scheme for breast pathology in the UK. *J. Clin. Pathol.* 2017; **70**; 51–57.

23. Elsharawy KA, Toss MS, Abuelmaaty SR *et al*. Prognostic significance of nucleolar assessment in invasive breast cancer. *Histopathology* 2020; **76**; 635–787.

24. Rakha EA, Aleskandarany MA, Toss MS *et al*. Impact of breast cancer grade discordance on prediction of outcome. *Histopathology* 2018; **73**; 904–915.

25. Lohse M, Bolger AM, Nagel A *et al*. Robina: a user-friendly, integrated software solution for RNA-SEQ-based transcriptomics. *Nucleic Acids Res.* 2012; **40**; W622–W627.

26. Robinson MD, McCarthy DJ, Smyth GK. Edger: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010; **26**; 139–140.

27. Zhang B, Kirov S, Snoddy J. Webgestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res.* 2005; **33**; W741–W748.

28. Wang J, Vasaikar S, Shi Z, Greer M, Zhang B. Webgestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Res.* 2017; **45**; W130–W137.

29. Liao Y, Wang J, Jaehnig EJ, Shi Z, Zhang B. Webgestalt 2019: gene set analysis toolkit with revamped UIS and APIS. *Nucleic Acids Res* 2019; **47**; W199–W205.

30. Gyorffy B, Lanczky A, Eklund AC *et al*. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res. Treat.* 2010; **123**; 725–731.

31. Horlings HM, Weigelt B, Anderson EM *et al*. Genomic profiling of histological special types of breast cancer. *Breast Cancer Res. Treat.* 2013; **142**; 257–269.

32. Lacroix-Triki M, Suarez PH, MacKay A *et al*. Mucinous carcinoma of the breast is genomically distinct from invasive ductal carcinomas of no special type. *J. Pathol.* 2010; **222**; 282–298.

33. Lopez-Garcia MA, Geyer FC, Natrajan R *et al*. Transcriptomic analysis of tubular carcinomas of the breast reveals similarities and differences with molecular subtype-matched ductal and lobular carcinomas. *J. Pathol.* 2010; **222**; 64–75.

34. Riener MO, Nikolopoulos E, Herr A *et al*. Microarray comparative genomic hybridization analysis of tubular breast carcinoma shows recurrent loss of the cdh13 locus on 16q. *Hum. Pathol.* 2008; **39**; 1621–1629.

35. Thor AD, Eng C, Devries S *et al*. Invasive micropapillary carcinoma of the breast is associated with chromosome 8 abnormalities detected by comparative genomic hybridization. *Hum. Pathol.* 2002; **33**; 628–631.

36. Zhang Y, Toy KA, Kleer CG. Metaplastic breast carcinomas are enriched in markers of tumor-initiating cells and epithelial to mesenchymal transition. *Mod. Pathol.* 2012; **25**; 178–184.

37. Weigelt B, Geyer FC, Reis-Filho JS. Histological types of breast cancer: How special are they? *Mol. Oncol.* 2010; **4**; 192–208.

38. Rakha EA, Aleskandarani M, Toss MS *et al*. Breast cancer histologic grading using digital microscopy: concordance and outcome association. *J. Clin. Pathol.* 2018; **71**; 680–686.

39. Dziegiel P, Owczarek T, Plazuk E *et al*. Ceramide galactosyltransferase (UGT8) is a molecular marker of breast cancer malignancy and lung metastases. *Br. J. Cancer* 2010; **103**; 524–531.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Photo micrographic examples of nucleolar scores and grade discordance. (**A**) inconspicuous nucleoli, (**B**) occasional conspicuous nucleoli, (**C**) prominent nucleoli easily seen, (**D**) example of grade 1:2 discordance and (**E**) example of grade 2:3 discordance.

Table S1. Distribution of various histological breast cancer histological subtypes in the TCGA dataset and included in the study.

Table S2. Differentially expressed genes and the top regulated pathway in association with the grade components and nucleolar prominence (concordant cases) defined by differential gene expression, and Gene set enrichment analysis (GSEA)/ pathway analysis.