Introduction

Histological grade (HG) describes the aggressive potential of solid tumors. The classical and widely adopted grading system for hepatocellular carcinoma is Edmondson-Steiner (ES), which is based on microscopic evaluation of the tubule formation, mitotic count, and nuclear pleomorphism. According to the ES grading system, tumors can be classified into three or four grades. A tumor of a higher grade tends to grow and spread at a faster pace, which needs more urgent and aggressive treatment.

Needle biopsies and histopathological evaluation works as a gold standard for HG diagnosis. Collectively, clinical physicians have accumulated a large amount of experience in using this method. However, it has two major problems, diagnostic subjectivity and biopsy inaccessibility (1), which might hinder its full efficacy. A stricter tumor grading requires two or more pathologists with expertise in a specific cancer to reduce diagnostic subjectivity. Much effort has been made with non-invasive methods such as magnetic resonance and contrast computed tomography (CT) to avoid biopsy unavailability (2). In contrast to diverse imaging methods, molecular biomarkers could overcome the two problems mentioned above. For example, miR-1290 could work as a biomarker of high-grade serous ovarian carcinoma (3), and tumor tissue protein signatures could predict the HG of breast cancer (4). Apart from the expression of biomarkers as an indicator
of HG, the genomic variation could also be used. Many gene mutations have been recognized to be associated with HG, such as TP53 (5), IDH1/2 (6) and ACVR2 (7). We are interested to know how far genomic variations are associated with HG in HCC because it can help us to understand HG as an important clinical measure.

With the development of precision medicine, DNA sequencing provides rich information for disease diagnosis and precision treatment. By using a liquid biopsy, genomic variations could prove to be more useful in predicting HG and could perfectly overcome the two major problems in the traditional ES grading system. Additionally, this method could provide necessary information for precise treatment in one-shot sequencing. However, ctDNA concentration is more easily affected by cancer development and clinical therapy, and the standard to detect ctDNA from liquid biopsy has not been well established.

This study sequenced 487 tissue samples from 459 Chinese HCC patients to build a solid connection between genomic variation and HG. Genomic variation, including nucleotide substitution/indel, truncation, gene homozygous deletion and fusion, were called. Association of HG with factors including genomic variation types, substitution types, mutational frequency related scores and biological processes was studied. Among the factors, those found to be significant were compared to those of the Western population.

**Methods**

**Patients and samples**

This study was approved by Shandong Provincial Hospital Affiliated to Shandong University and The Affiliated Hospital of Qingdao University. A total of 602 patients were enrolled. Each participant provided written informed consent. Samples were collected from surgery after diagnosis or relapse. HG was scored by a specialist in hepatobiliary pathology according to the Edmondson and Steiner method (8). Grade 1 was defined as well differentiated (WD), grade 2 and 3 as moderate differentiated (MD), grade 4 as poor differentiated (PD). The patients were staged according to the seventh edition of the tumor-node-metastasis (TNM) classification system for lung cancer from the American Joint Committee. We also collected another public dataset to validate our analysis. This dataset, MSKCC, containing 360 samples, was downloaded from cBioportal (https://www.cbioportal.org/, accessed on March 5, 2019).

**Library preparation and next-generation sequencing**

Tissue samples (40 μm section) were collected for each patient. KAPA Hyper Prep Kit (#07962363001, Roche, Basel, Switzerland) was used to extract DNA. PBS (phosphate-buffered saline) was added to those samples with volumes of less than 5 mL in order to make each sample volume equivalent to 5 mL. They were centrifuged (2 times at 1,600 g for 10 and 15 min, respectively) for extraction of DNA and the supernatant was separated. Invitrogen Qubit® DNA HS Assay Kit (#Q32854) was used to measure the DNA concentration. Single strand DNA and protein contamination were excluded. Library construction was only applied in samples with at least 50 ng of double-stranded DNA extracted. Molecular identifiers (MIDs) were added to the DNA segment ends for DNA libraries to reduce the false discovery rate (FDR). Barcodes were also added to the reads for multiplex sequencing. Sequencing was performed on an Illumina Novaseq 6000 (Illumina, San Diego, CA) for 151 bp read length from both ends. The average sequencing depth was about 3,000x.

**Variants calling**

A pan-cancer panel (Yuansuo®, Origimed, Shanghai, China) comprising 588 genes was captured with targeted amplification. Adaptors were trimmed from raw DNA reads by cutadapt (version 1.18) (9). MID-labeled reads were de-duplicated with an in-house pipeline. BWA MEM (version 0.7.9a) (10) mapped the high-quality reads to the UCSC hg19 reference sequences. Base quality was recalibrated by the BaseRecalibrator tool from GATK (version 3.8) (11). Mutect2 with a tumor-only mode (12) and Varscan (version 2.3.9) (13) with the default parameters were used to call variants.

For each sample, the germline variants having variant allele frequency (VAF) <0.1% were filtered according to the databases of ExAC (14), gnomAD (15), 1000 Genomes (16), and ESP6500 (17). Somatic variants that had not been filtered were further annotated by ANNOVAR (2017/07/17) (18) with RefSeq (version 2017/06/01).

Fermi-lite (19) was used to identify gene fusion and rearrangement. The breakpoints were further checked by BLAT (http://genome.ucsc.edu, version 3.50). Those reads uniquely mapping to the reference genome constituted rearrangement supported reads.

CNVKit (20) was used to estimate the logR scores. Copy number was assigned 1 for logR values below −0.25, 3 for
Bioinformatics analysis

The mutant allele tumor heterogeneity (MATH) score for a tumor was calculated as the median absolute deviation divided by the median MAF of all somatic mutations detected in the tumor sample. As suggested by Jiang et al. (22), the calculation of MATH used somatic mutation calls with MAF of 0.075 or greater. Clonal mutation burden (CMB) (22) was defined as the number of mutations per clone, and divided into low (low TMB, high MATH), high (high TMB, low MATH), or intermediate (others).

Statistical analysis

The Mann-Whitney U test was used to compare TMB, MSAF, MATH, and CMB between different HGs. Fisher’s exact test was performed to compare the count number of nucleotide mutations for different HGs.

Survival analysis was conducted with R software. Samples were classified by a cutoff at the median mutational frequency. The survival time was plotted against overall survival probability by the Kaplan-Meier method. The log-rank test was applied to calculate the P value between the two groups.

Results

Patients and genomic variation detection

The analysis workflow of this study is illustrated in Figure S1. Initially, a total of 602 patients were enrolled in this study. Of these, only 459 patients had histological grading information available. The other patients’ samples were thus filtered out from the following study. Their clinicopathologic characteristics are summarized in Table S1. The median age of patients was 55 years old (range, 16 to 82 years old). Most of the patients were male (87.6%). According to the TNM classification system (23), the number of patients in the early stage (I/II/III) and late-stage (IV) were 418 and 41, respectively. Patients who consumed alcohol more than 200 days per year were classified as “drinking”, and those who had an immediate family member with any type of cancer were labeled as “family history”. HGs were divided into three categories: poorly differentiated (PD), moderately differentiated (MD) and well differentiated (WD).

Tissue samples were prepared by surgery and enriched with a pan-cancer panel of genes (YuanSuo®, Origimed Co., Ltd, Shanghai, China) (Figure S2). For the 459 patients, 487 samples were collected. Somatic genomic variations (SGVs) were called by Mutec2 (12) and Varscan (13) for each sample. Gene amplifications were called by CNVKit (20). Fermi-lite (19) was used to identify gene fusion and rearrangement. The genomic variations at top high frequency are depicted in Figure 1. Genomic variations from multiple samples of each patient were merged under the same patient.

The bias of SGV types in different HGs

There are multiple types of SGVs deriving from different mechanisms. Those SGVs were classified into five types (fusion/rearrangement, gene amplification, gene homozygous deletion, substitution/indel, and truncation). The percentage of those groups was summarized according to the HG groups (Figure 2A). From poorly to moderately to well-differentiated HG, the percentage of the truncation and substitution/indel group increased but that of the gene amplification group dropped. The poorly differentiated group had the lowest percentage of fusion/rearrangement but the highest percentage of gene amplification variations.

Single nucleotide variants (SNVs) can be classified into transversion substitution and transition substitution. Transition SNVs regularly had a higher frequency and caused no functional change because of codon “wobble”. To study the association between amino acid changes and HGs, the percentages of transversion and transition were compared (Figure 2A). In total, the percentage range of transversion and transition for different HGs was 65–72% and 28–34%, respectively. WD had higher transition than PD with percentages of 52% and 47%, respectively (P value =2.537e-11), but had lower transversion than PD with percentages of 48% and 53%, respectively (P value =2.2e-16). For specific substitutions, WD had less C>G transversion (P value =0.0058) and more G>A transition (P value =0.026) than non-WD. PD had higher C>T transition (P value =0.01) than non-PD.

Except for mutational occurrence, we also studied the association between HGs and mutational frequency related scores including maximum somatic allele frequency (MSAF) (24), MATH and CMB (22). MSAF was regularly used as a measure of cellular tumor prevalence. Higher MSAF denoted higher tumor content. The MATH score denoted allele heterogeneity among each sample, which reflected the diversity of mutational clones. CMB score combined
Figure 1 The landscape of genomic variations. From top to bottom, the bar plot indicates the tumor mutation burdens (TMBs) and the below heat map indicated the clinicopathological characteristics. HG (histological grades) includes WD (well-differentiated), MD (moderately differentiated), and PD (poorly differentiated). The bottom left bar plot indicates the percentage of genomic variation for each gene in the patients. The bottom right heatmap shows genomic variation types.

Figure 2 Variation distribution for hepatocellular carcinoma (HCC). (A) The upper plot shows the distribution of five types of genomic variation for three groups of HGs (histological grades). The lower plot is the distribution of 12 substitution types, which are grouped into transition and transversion. The x-axis indicates the patient percentage and the y-axis indicated the HGs. (B) The upper and the lower plots show MSAF (maximum somatic allele frequency) and MATH (mutant allele tumor heterogeneity) distributions for the three HG groups, respectively.
the tumor mutation burden (TMB) and MATH score (22). High CMB was defined as high TMB and low MATH. These scores were compared among different HGs. WD had significantly lower MSAF than MD and PD with P values equal to 0.031 and 0.038, respectively (Figure 2B). As for the MATH score, MD was highest among HGs, but only MD transversion. PD had a P value of less than 0.05. We also tested the CMB score, but no significant difference was found among HGs (result not shown).

The functional bias of genomic variations for different HGs

Driver genes play a big part in cancer. Their specific effect on HG was also studied. The driver genes of HCC were collected from the literature (25,26). The top 10 variable genes are displayed in Figure 3A. Genes TP53, TERT, CTNNB1, RB1, AXIN1, and ARID1A were prone to substitution/indel/truncation variation, while CCND1, FGF19, FGF4, and FGF3 preferred gene amplification. Among those driver genes, there were three genes showing significantly different HGs after mutation (Figure 3B). Of these three genes, mutational TP53 (TP53+) had a higher average HG than non-mutational TP53 (TP53−). A non-parameter Wilcoxon’s rank-sum test showed significance at P value =3.8e-2. In contrast, mutational CTNNB1 (CTNNB1+) and FGF3 (FGF3+) showed significantly lower
HG with P value = 7.8e-4 and P value = 0.04, respectively. We also tested the association between gene amplification variation and HG for those driver genes, but no significant difference was found (Figure 3C).

Mutations, such as substitution, indel, and truncation, can modify the targeted gene functions, and amplification can modify their expression. To study their functional bias, three HG groups were intersected with each other as displayed by a Venn plot in Figure 4A. WD, MD, and PD had 14, 57 and 62 unique mutated genes, respectively. WD had fewer unique mutated genes than other HGs.

The unique genes for WD and PD were enriched with the biological processes of gene ontology. A hypergeometric test was performed for each biological process. The Bonferroni-Hochberg (BH) method was applied to correct for multiple testing errors. The top 10 enriched biological processes for WD were listed in Figure 4B. WD was enriched in the regulation of cell differentiation and cell proliferation. There were 1,168 biological processes enriched for PD specific mutations with multiple testing corrected P values less than 0.05, such as cell proliferation, protein phosphorylation,
and cellular response to a stimulus.

Apart from substitution/indel/truncation mutations, copy number variation can also disrupt cellular function by modifying gene regulation. The amplified genes were intersected with each other (Figure S3A) to obtain the HG-specific genes. The specific genes had similar distribution as substitution/indel/truncation for gene amplification. WD had less specific gene amplification than other HGs, while PD had the highest number of specific genes. WD was enriched in the regulation of fibroblast migration and the negative regulation of transport (Figure S3B); MD was enriched in the positive regulation of cellular processes and the regulation of cell proliferation (Figure S3C); and PD was enriched in the positive regulation of metabolic processes (Figure S3D).

Comparison to the Western population

The findings above were compared against the Western population. An MSKCC dataset from the Western population was downloaded from cbioportal (https://www.cbioportal.org/, accessed on March 5, 2019). With this dataset, nucleotide usage, TMB, driver genes, and biological processes were analyzed using the same procedures as in our dataset. Results showed that WD possessed a higher percentage of transition mutation than PD in the Western population. Meanwhile, PD held a higher percentage of transversion than WD. Such results were in line with those from our dataset. As for the nucleotide usage, only C>G transversion showed higher frequency in PD than in WD (P value =0.017), matching the result from our dataset. Specifically to the Western population, WD had higher A>G mutation than non-WD with P value =0.031. PD had higher A>C and lower A>G substitution than non-PD with P values =0.036 and 7.5e-4, respectively. Among the driver genes, only TP53 mutation was consistently associated with higher HG (P value =1.3e-3, Mann-Whitney U test). Additionally, in the Western population, the RB1 mutation tended to be enriched in the high-grade samples.

Further investigation of the similarity between the functional biases for WD- and PD-specific genes revealed an extraordinary consistency. The top significantly enriched biological processes in our dataset showed similar significance in the MSKCC dataset (Figure 5A,B,C). For example, for both our dataset and the MSKCC dataset, PD-specific genes took part in cell proliferation, protein phosphorylation, and regulation of cell proliferation; MD-specific genes took part in the cellular protein modification process and cellular response to stimulus; and WD-specific genes taking part in the regulation of developmental processes and cell differentiation.

It was noteworthy that HG-specific genes may share common enriched biological processes (Figure 5D). To extract the consistent HG-specific biological processes between the two datasets, we first extracted the HG-specific biological processes taken by HG-specific genes for both datasets. Then an intersection was conducted between HG-specific biological processes for both datasets. Through these means, we identified the HG-specific biological processes commonly taken by both datasets. There were 3 WD-specific, 150 MD-specific and 64 PD-specific common biological processes (Table S2). These gene lists were applied for gene ontology enrichment analysis. The PD-specific common biological processes included angiogenesis, phosphatidylinositol-3-phosphate biosynthetic process, glycerophospholipid metabolic process and development of primary male sexual characteristics; the MD-specific common biological processes included response to hydrogen peroxide, response to peptide hormone and protein localization to the nucleus; and the three WD-specific common biological processes were regulation of epithelial to mesenchymal transition involved in endocardial cushion formations, epithelial to mesenchymal transition involved in endocardial cushion formations and regulation of homotypic cell-cell adhesion.

Discussion

Although there have been many studies on the association between gene expression and HG, information on the association between genomic variation and HG is still scarce. The intention of this study was to understand the association between genomic variation and HG and explore the potential of genomic variation as an indicator of HG.

A stable genomic variation pattern should be associated with a hidden molecular mechanism. For example, C>T and C>G substitution could come from DNA editing catalyzed by apolipoprotein B mRNA catalytic subunit-like (APOBEC) and activation-induced deaminase (AID) family, which can bind to both RNA and single-stranded (ss) DNA. DNA deamination by these proteins results in the C>U conversion in single-stranded DNA. Such mutations could result in C>T transition and C>G transversion by different DNA repair polymerases (27). In lung cancer, different cancer subtypes also showed a large difference in C>T transition and C>G transversion (28). Due to the existence
of such molecular mechanisms, those stable genomic variation patterns could be stable predictors of HG. In this study, we have analyzed the association of HGs with genomic variation and mutational frequency in the Chinese population and the Western population, and have found a higher C>G transversion mutated in patients with PD HCC for both populations. This association was meaningful in the treatment of such a subset of HCC patients. As reported, APOBEC-related mutagenesis was found to be highly correlated with immunotherapy response (29). Thus, detected C>G transversion could be a good indicator of immunotherapy efficacy. In spite of high C>G transversion being found in HCC and believed as an etiology of HCC by Morishita et al. (30), they did not associate it with any biological significance. Our results revealed that patients with high C>G transversion were strongly associated with poorly differentiated HCC, involving in APOBEC-related mutagenesis.

Taking into account the important mutational scores in relation to survival, we also studied TMB, MSAF, MATH and CMB score, among which only MSAF is significantly associated with HG. As a measure of cellular tumor prevalence, MSAF has been used in many studies (24,31,32). Studies have shown that MSAF is also correlated with tumor burden (31) and several other research studies have revealed that tumor burden is strongly associated with HG (32). Therefore, it is reasonable that MSAF was significantly associated with HG.

Among the driver genes of HCC, TP53 mutation was a consistent biomarker of high HG in both populations, which agreed with the previous studies in ovarian cancer (5) and HCC (33). However, we also noticed difference between the Chinese and Western populations. In the Chinese population, mutations in CTNNB1, ARID2, and ACVR2A were associated with a lower HG, and in the western population, RB1 was associated with a high HG.

During the analysis of biological processes of substitution/indel/truncation and amplification for WD- and PD-specific genes, we found that the biological

![Figure 5 Biological processes could predict survival accurately. (A) The top enriched biological process in WD-specific genes from the Chinese population was validated in the Western population; (B) the top enriched biological process in MD-specific genes from the Chinese population was validated in the Western population; (C) the top enriched biological process in PD-specific genes from the Chinese population was validated in the Western population; (D) the intersection of enriched biological processed in the Chinese population and the Western population. WD, well-differentiated; MD, moderately differentiated; PD, poorly differentiated.](image)

processes were highly matched for mutation and amplification. For example, WD tumors showed higher substitution/indel/truncation and amplification in the cell differentiation, and PD tumor showed higher substitution/indel/truncation and amplification in the protein phosphorylation. These results demonstrated that a tumor could become WD or PD either through mutations or by amplification, or both. Comparisons between the Chinese and the Western populations also proved that the WD was most enriched in cell differentiation, and the PD was most enriched in phosphorylation. Furthermore, there were also genes for WD or PD involved in phosphorylation or cell differentiation, respectively. A further intersection of their biological processes disclosed the unique biological processes for different HGs. Although these biological processes have been well recognized in basic cancer research, they have not been systematically associated with genomic variations and HG in HCC before.

It should be noted that, instead of ctDNA (circulating tumor DNA) from blood, DNA from the solid tumor was extracted to detect gene mutations. Considering the instability of ctDNA detection, this should be a very important step for applying genomic variations as a predictor of HG. For example, ctDNA is easier to be detected in late-stage cancer and its concentration can be changed by many factors including clinical therapy and tumor development. How the instability of ctDNA detection affects its prediction is another issue to be discussed. The other limitation of this study is that the comparison with the Western population did not include CNV due to the missing information in the MSKCC dataset.

In summary, this pilot study has revealed multiple factors associated with HG. These findings improved our understanding of the molecular mechanism in different HGs of HCC. Further research using ctDNA to detect the genomic variation should be performed to verify this study.

Acknowledgments

Funding: None.

Footnote

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at http://dx.doi.org/10.21037/tcr.2020.03.32). The work was carried out as part of the employment of the corresponding author at the Affiliated Hospital of Qingdao University. The Affiliated Hospital of Qingdao University was not involved in the manuscript writing, editing, approval, or decision to publish. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was approved by Shandong Provincial Hospital Affiliated to Shandong University and The Affiliated Hospital of Qingdao University. Each participant provided written informed consent. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: https://creativecommons.org/licenses/by-nc-nd/4.0/.

References

1. Cillo U, Giuliani T, Polacco M, et al. Prediction of hepatocellular carcinoma biological behavior in patient selection for liver transplantation. World J Gastroenterol 2016;22:232-52.
2. Schadendorf D, Hodi FS, Robert C, et al. Pooled Analysis of Long-Term Survival Data From Phase II and Phase III Trials of Ipilimumab in Unresectable or Metastatic Melanoma. J Clin Oncol 2015;33:1889-94.
3. Kobayashi M, Sawada K, Nakamura K, et al. Exosomal miR-1290 is a potential biomarker of high-grade serous ovarian carcinoma and can discriminate patients from those with malignancies of other histological types. J Ovarian Res 2018;11:81.
4. Skoog P, Ohlsson M, Ferno M, et al. Tumor tissue protein signatures reflect histological grade of breast cancer. PLoS One 2017;12:e0179775.
5. Cole AJ, Dwight T, Gill AJ, et al. Assessing mutant p53 in primary high-grade serous ovarian cancer using immunohistochemistry and massively parallel sequencing. Sci Rep 2016;6:26191.
6. Deng L, Xiong P, Luo Y, et al. Association between
IDH1/2 mutations and brain glioma grade. Oncol Lett 2018;16:5405-9.
7. Wodziński D, Wosiak A, Pietrzak J, et al. Does the expression of the ACVR2A gene affect the development of colorectal cancer? Genet Mol Biol 2019;42:32-9.
8. Edmondson HA, Steiner PE. Primary carcinoma of the liver: a study of 100 cases among 48,900 necropsies. Cancer 1954;7:462-503.
9. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.
10. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:129-303.
11. Li H, Shi W, Wali VB, et al. Predictors of Chemosensitivity in Triple Negative Breast Cancer: An Integrated Genomic Analysis. PLoS Med 2016;13:e1002193.
12. Edge SB, Byrd D, Compton C, et al. Digestive system. In: Trotti A. editor. AJCC cancer staging manual, ed 7th. New York: Springer, 2010.
13. Kandoth C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. Nature 2013;502:333-9.
14. Boichard A, Pham TV, Yeerna H, et al. APOBEC-related mutagenesis and neo-peptide hydrophobicity: implications for response to immunotherapy. Oncoimmunology 2018;8:1550341.
15. Badowska-Kozakiewicz AM, Liszcz A, Sobol M, et al. Retrospective evaluation of histopathological examinations in invasive ductal breast cancer of no special type: an analysis of 691 patients. Arch Med Sci 2017;13:1408-15.
Figure S1 Workflow for this study.

Table S1 Clinical characteristics in each histological grade

| Characteristics | Poor/moderate | Well           |
|-----------------|---------------|----------------|
| Age (mean ± SD) | 54.9±15.0     | 60.7±12.5      |
| Gender          |               |                |
| Male            | 358           | 44             |
| Female          | 54            | 3              |
| Stage           |               |                |
| I/II/III        | 375           | 43             |
| IV              | 37            | 4              |
| Drink           |               |                |
| Yes             | 50            | 7              |
| No              | 362           | 40             |
| Family          |               |                |
| Yes             | 92            | 5              |
| No              | 320           | 42             |
The gene list of the targeted sequencing.
Figure S3 Gene amplification difference between different grades. (A) The gene amplification overlaps between three HGs for substitution/indel/truncation; (B) the enriched biological processes for well differentiated tumors; (C) the enriched biological processes for moderately differentiated tumors; (D) the enriched biological processes for poorly differentiated tumors. The length of the blue bar indicates the negative log transformed false discover rate (FDR). WD, well differentiated; MD, moderately differentiated; PD, poorly differentiated.
| GO:0008589 | Regulation of intracellular protein transport | MSKCC | MD | 0.0038896 |
|-------------|---------------------------------------------|-------|----|------------|
| GO:0014042  | Determination of heart left/right asymmetry  | MSKCC | WD | T cell receptor signaling pathway |
| ZB          | Positive regulation of granulocyte macrophage colony-biosynthetic process | 0.0343113 |
| GO:0014042  | DNA strand elongation                       | ZB    | PD | 0.0363016  |
| GO:0032612  | Negative regulation of cell-cell adhesion   | ZB    | PD | 0.0371894  |
| GO:0001541  | Response to monosaccharide                  | MSKCC | PD | 0.0114927  |
| GO:0010266  | Female gonad development                     | MSKCC | MD | Heart valve morphogenesis |
| GO:0010883  | Cellular response to vitamin B1             | MSKCC | MD | 0.0174305  |
| GO:0032611  | Positive regulation of gamma-delta T cell differentiation | 0.040316  |
| GO:0032496  | Positive regulation of fibroblast proliferation | MSKCC | MD | 0.0432135  |
| GO:0031648  | Positive regulation of myeloid leukocyte mediated immunity | 0.0437818  |
| GO:0009607  | Response to formaldehyde                     | MSKCC | PD | 0.0007089  |
| GO:0033365  | Development of primary male sexual characteristics | MSKCC | PD | 0.0130477  |
| GO:0009743  | Regulation of interleukin-3 production       | MSKCC | PD | 0.0136096  |
| GO:002237   | Positive regulation of phospholipid metabolic process | 0.0400926  |
| GO:0006284  | Histone H3-K9 methylation                   | MSKCC | MD | 0.0127762  |
| GO:0010883  | Protein modification by small protein conjugation | PD | MD | 0.0265122  |
| GO:0010431  | Phospholipid metabolic process              | ZB    | PD | 0.0243556  |
| GO:0090349  | Isoquinoline alkaloid metabolic process      | MSKCC | MD | 0.0282742  |
| GO:0014042  | Signal release                              | ZB    | PD | 0.0282742  |
| GO:0008654  | Response to other organism                   | MSKCC | MD | 0.0295013  |
| GO:0007257  | Response to reactive oxygen species          | ZB    | PD | 0.0359801  |
| GO:0032612  | Response to hydrogen peroxide                | MSKCC | MD | 0.0363016  |
| GO:0009409  | Cellular response to hydrogen peroxide       | ZB    | PD | 0.0364954  |
| GO:0031060  | Positive regulation of histone H3-K9 methylation | 0.0432135  |
| GO:0042223  | Response to monosaccharide                   | MSKCC | PD | 0.0414965  |
| GO:0032611  | Protein localization to nucleus              | MSKCC | MD | 0.0437818  |
| GO:0032611  | Protein modification by small protein conjugation | 0.0437818  |
| GO:0002237  | Interleukin-1 production                     | MSKCC | PD | 0.0012901  |
| GO:0018022  | Positive regulation of phospholipid metabolic process | 0.0400926  |
| GO:0038034  | Regulation of protein export from nucleus    | MSKCC | MD | 0.0015026  |
| GO:0030520  | Positive regulation of myeloid leukocyte mediated immunity | 0.0400926  |
| GO:1990874  | Regulation of response to DNA damage stimulus | MSKCC | MD | 0.0088126  |
| GO:0043367  | Development of primary female sexual characteristics | MSKCC | MD | 0.0139992  |
| GO:1903533  | Meiotic cell cycle process                   | ZB    | PD | 0.0002231  |
| GO:0006471  | Response to drug                             | MSKCC | MD | 0.0295013  |
| GO:0002220  | Positive regulation of interleukin-3 biosynthetic process | 0.0400926  |
| GO:0043207  | Response to formaldehyde                     | MSKCC | MD | 0.0282742  |
| GO:0003205  | Response to formaldehyde                     | MSKCC | MD | 0.0282742  |
| GO:0032446  | Positive regulation of fibroblast proliferation | MSKCC | MD | 0.0432135  |
| GO:0031060  | Positive regulation of histone H3-K9 methylation | 0.0432135  |
| GO:1904035  | Homotypic cell-cell adhesion                 | MSKCC | MD | 0.0363016  |
| GO:1903533  | Regulation of vascular smooth muscle cell proliferation | 0.0432135  |
| GO:0010883  | Response to hydrogen peroxide                | ZB    | PD | 0.0024086  |
| GO:0031648  | Negative regulation of low-density lipoprotein particle clearance | 0.0432135  |
| GO:0010883  | Response to hydrogen peroxide                | ZB    | PD | 0.0028681  |
| GO:0032612  | Response to formaldehyde                     | MSKCC | MD | 0.0002231  |
| GO:0002237  | Interleukin-1 production                     | MSKCC | PD | 0.0265122  |
| GO:1901652  | Response to formaldehyde                     | MSKCC | PD | 0.0265122  |
| GO:0043207  | Response to formaldehyde                     | MSKCC | MD | 0.0265122  |
| GO:0042493  | Regulation of response to DNA damage stimulus | MSKCC | MD | 0.0088126  |
| GO:1904705  | Development of primary female sexual characteristics | MSKCC | MD | 0.0139992  |