# Design, Analytical Validation and Diagnostic Yield of a Novel Canine Cancer Gene Sequencing Panel

*Shukmei Wong[1], Manisha Warrier[1], Sara Byron[1], Victoria Zismann[1], Martin Boateng[1], Salvatore Facista[1], Timothy Whitsett[1], Colt Tallant[1], Natalia Briones[1], Tyler Izatt[1], Kathryn Banovich[1], David Haworth[1], Han-Yu Chuang[1], William PD Hendricks[1]*

*[1]The Translational Genomics Research Institute (TGen) and Vidium Animal Health, a Subsidiary of TGen*

## INTRODUCTION

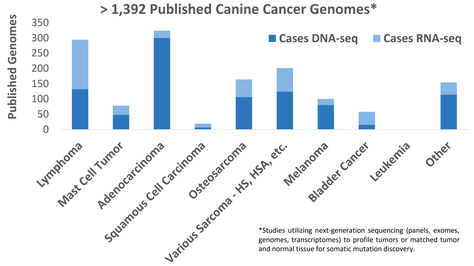
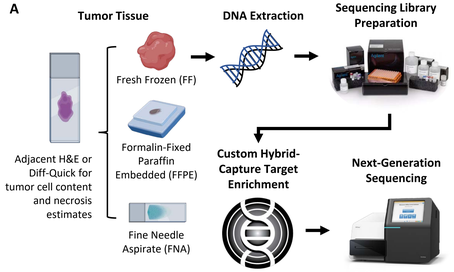



**Figure 1. Knowledge of Canine Cancer Genome Landscapes is Rapidly Accelerating.** Canine cancer genome sequencing studies are unearthing new candidate diagnostic, prognostic, and predictive biomarkers. Development and rigorous validation of genomic diagnostics is thus increasingly important for enabling new biomarkers to be confidently leveraged in research and the clinic.

## METHODS



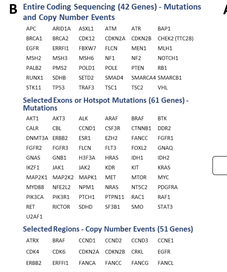
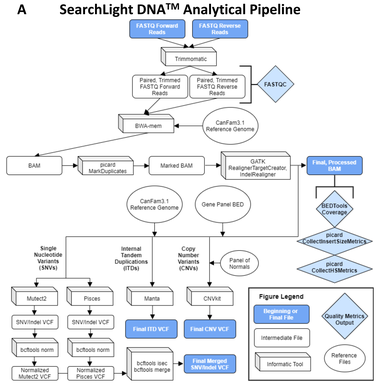




**Figure 2. SearchLight DNA™ Workflow and Content.** SearchLight DNA™ is a tumor-only, next-generation sequencing (NGS), hybrid-capture, canine gene panel covering 482,000 base pairs of 120 genes associated with canine or human cancer. (**A**) SearchLight DNA™ wet laboratory workflow for sequencing tumor tissue (FF, FFPE, FNA). (**B**) Genes and mutation types evaluated by SearchLight DNA™.



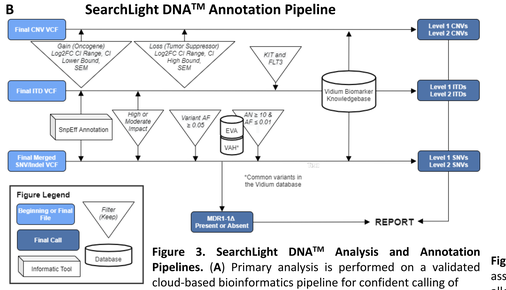




**Figure 3. SearchLight DNA™ Analysis and Annotation Pipelines.** (**A**) Primary analysis is performed on a validated cloud-based bioinformatics pipeline for confident calling of single nucleotide variants (SNVs), copy number variants (CNVs), and internal tandem duplications (ITDs). (**B**) Candidate pathogenic variants are annotated according to predicted impact and filtered to remove common non-pathogenic single-nucleotide polymorphisms (SNPs) from the European Variant Archive (EVA) and Vidium data, then annotated with human-to-canine-translated databases (COSMIC and cBioPortal) as well as biomarker associations from Vidium's knowledge database.

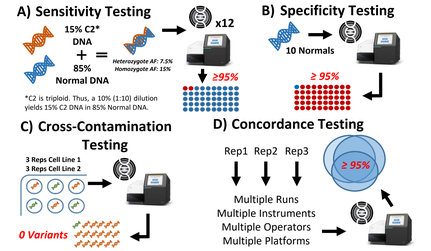



**Figure 4. SearchLight DNA™ Analytical Validation Experimental Design.** We assessed: (**A**) Sensitivity at a target Limit of Detection (LoD) of 7.5% heterozygous allele fraction. Replicate sequencing (12x) of a diluted co-mixture of the triploid C2 mast cell tumor cell line 1:10 (10%) with an unmatched normal sample was performed and frequency of true C2 variants measured. (**B**) Specificity via sequencing of normal tissues to identify background variants at low allele frequencies (0.03 < AF < 0.20). (**C**) Cross-contamination by preparing two different samples in a checkerboard. (**D**) Concordance via replicates run under varying conditions. Target metrics (acceptance criteria) for each study are shown in ***red.***

## RESULTS

| Analytical Sensitivity | Median Sensitivity all SNVs (Range) | Median Sensitivity at Target AFs 7.5% - 10.5% (Range) | Sensitivity for KIT ITD (Positive Replicates) |
|---|---|---|---|
| C2:Normal Dilution | 90.28% (83.33% - 97.22%) | 95.24% (95.24% - 100%) | 100% (12/12 Replicates) |

**Table 1. SearchLight DNA™ Analytical Sensitivity.** Sensitivity for detection of unique Single Nucleotide Variants (SNVs) and an Internal Tandem Duplication (ITD) in the C2 cell line when diluted to a target heterozygous allele frequency (AF) of 7.5%. 36 unique C2 "truth" SNVs were determined by sequencing 10 undiluted replicates of C2 and selecting passing SNVs at AF ≥ 0.03 that occurred in all replicates. Sensitivity was calculated based on the frequency at which these variants were detected at AF ≥ 0.03 in 12 replicates of diluted C2.

| Analytical Specificity | Median Specificity SNVs (Range) | Specificity for KIT and FLT3 ITDs (Negative Replicates) |
|---|---|---|
| Normal Tissues | 99.9995% (99.9981 - 100) | 100% (KIT: 10/10 Replicates) (FLT3: 10/10 Replicates) |

**Table 2. SearchLight DNA™ Analytical Specificity.** Specificity was determined by sequencing a panel of 10 constitutional normal DNAs from peripheral blood mononuclear cells (PBMCs). These genomes were expected not to contain SNVs with low AFs. AFs between 0.03 and 0.20 were thus considered noise. Specificity is calculated as the frequency of these SNVs among all sequenced nucleotides that pass analysis filters.

| Cross-Contamination | Contaminating SNVs and Indels from Cell Line 2 | | Contaminating SNVs and Indels from Cell Line 1 |
|---|---|---|---|
| Cell Line 1 Rep 1 | 0 | Cell Line 2 Rep 1 | 0 |
| Cell Line 1 Rep 2 | 0 | Cell Line 2 Rep 2 | 0 |
| Cell Line 1 Rep 3 | 0 | Cell Line 2 Rep 3 | 0 |

**Table 3. SearchLight DNA™ Cross-Contamination.** To evaluate cross-contamination rates when preparing samples in parallel in our clinical laboratory, we prepared 6 libraries in an adjacent checkerboard pattern from 2 distinct cell lines. Contamination was considered the presence of any SNV unique to one cell line that was detected in a passing call with AF ≥ 0.03 in the opposite cell line.

| Reproducibility | Shared Mutations (Percent Concordance) (95% Confidence Interval) |
|---|---|
| Run 1 v Run 2<br>Inter-Instrument Variability<br>*Sequencer 1 v Sequencer 2*<br>*Operator 1* | 428 (99.07%) (98.63% - 99.52%) |
| Run 1 v Run 3<br>Inter-Run Variability<br>*Instrument 1*<br>*Operator 1* | 425 (98.73%) (97.60% - 99.85%) |
| Run 1 v Run 4<br>Inter-Operator Variability<br>*Instrument 1*<br>*Operator 1 v Operator 2* | 427 (98.50%) (98.28% - 98.72%) |

**Table 4. SearchLight DNA™ Reproducibility.** In order to determine reproducibility, we prepared 12 replicate libraries of a canine pulmonary adenocarcinoma cell line which were sequenced across 4 sequencing runs varying the instrument and operator. Concordance was calculated for SNVs detected in a passing call with AF ≥ 0.03 across replicates.

| FFPE | Total # SNVs and Indels | # Common Variants | Concordance |
|---|---|---|---|
| FFPE | 475 | 422 | 88.84% |
| Cell Line | 428 | | 98.60% |

**Table 5. SearchLight DNA™ Interfering Substances: FFPE.** To determine the impact of FFPE preservation on sequencing results, we analyzed a fresh biopsy that was sub-divided into a formalin-fixed, paraffin-embedded section (FFPE) and a separate section grown in culture (cell line). Concordance was calculated for SNVs detected in a passing call with AF ≥ 0.03 across replicates.

| Average Tumor-Normal Unique SNVs | Average Tumor-Only Unique SNVs | Average Shared SNVs | Average Total SNVs | Concordance (95% CI) |
|---|---|---|---|---|
| 0.8 | 0.4 | 5.2 | 6.4 | 77.2% (61.53% - 92.86%) |

**Table 6. SearchLight DNA™ Tumor-Normal and Tumor-Only Concordance.** We compared matched tumor-normal sequencing to tumor-only analysis for 10 tumors with matched normal DNA. Concordance was calculated for SNVs detected in a passing call with AF ≥ 0.03 across replicates.

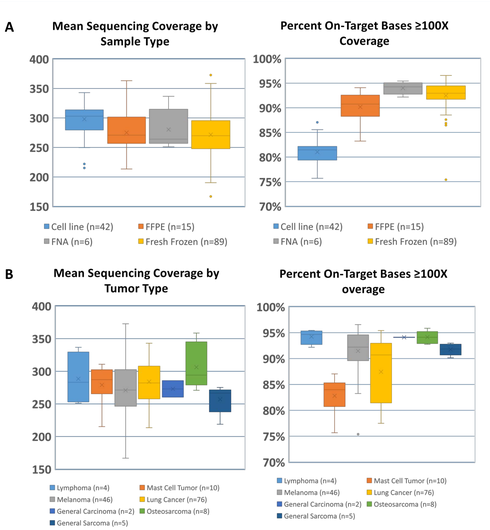



**Figure 7. SearchLight DNA™ Performance by Tumor and Tissue Type.** We have performed SearchLight DNA™ sequencing of >152 samples in our clinical laboratory on multiple tumor and tissue types. (**A**) Mean sequencing target coverage and percent of on-target sequenced reads with coverage ≥100x across tissue types. (**B**) Mean sequencing coverage and percent of on-target sequenced reads with coverage ≥100x across tumor types.

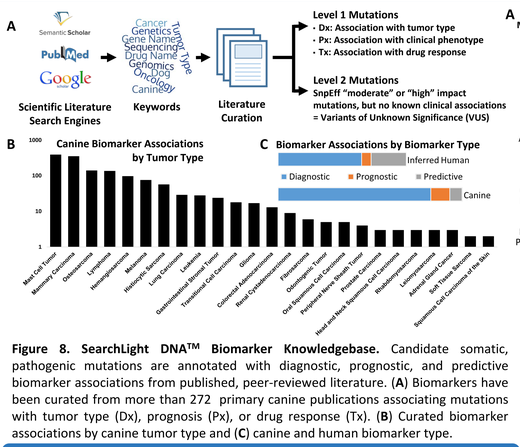



**Figure 8. SearchLight DNA™ Biomarker Knowledgebase.** Candidate somatic, pathogenic mutations are annotated with diagnostic, prognostic, and predictive biomarker associations from published, peer-reviewed literature. (**A**) Biomarkers have been curated from more than 272 primary canine publications associating mutations with tumor type (Dx), prognosis (Px), or drug response (Tx). (**B**) Curated biomarker associations by canine tumor type and (**C**) canine and human biomarker type.

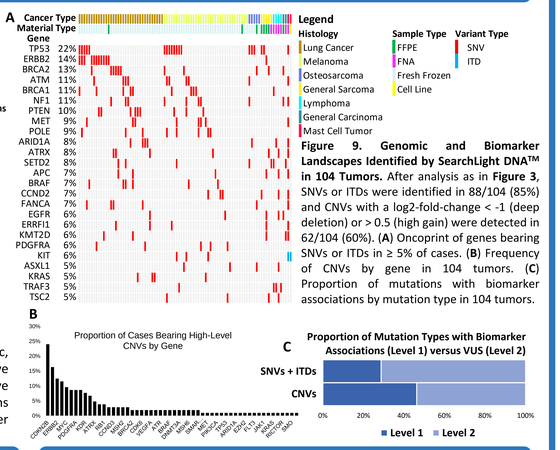



**Figure 9. Genomic and Biomarker Landscapes Identified by SearchLight DNA™ in 104 Tumors.** After analysis as in **Figure 3**, SNVs or ITDs were identified in 88/104 (85%) and CNVs with a log2-fold-change < -1 (deep deletion) or > 0.5 (high gain) were detected in 62/104 (60%). (**A**) Oncoprint of genes bearing SNVs or ITDs in ≥ 5% of cases. (**B**) Frequency of CNVs by gene in 104 tumors. (**C**) Proportion of mutations with biomarker associations by mutation type in 104 tumors.

## CONCLUSIONS

- For SearchLight DNA™ analytical validation, we have achieved target metrics including: 95% sensitivity and 99% specificity for detection of low AF SNVs and ITDs, no cross-contamination, reproducibility ≥ 95% and high performance across tissue and tumor types.
- SearchLight DNA™ is robustly designed and strictly validated. It represents the state-of-the-art in genomic diagnostics.
- Clinical applicability is the aggregate effect of many factors, but understanding the genomic identity of a tumor in a dog can be a valuable part of understanding the biology of that tumor.

## FUNDING & ACKNOWLEDGEMENTS

We thank Drs. Clifford, Crawford-Jakubiak, Gillings, Hamilton, Hershey, Overley, and Ward for their engagement and guidance. We thank Dr. Amy LeBlanc and Christina Mazcko of the NCI Comparative Oncology Program, and Drs. Holly Borghese and Gwendolen Lorch from the OSU-CVM Biospecimen Repository for assistance with sample procurement and annotation. Some samples described in these studies were also procured through the Canine Comparative Oncology Genomics Consortium. We utilized BioRender for generation of some images. *Financial disclosures: These studies were funded by Vidium Animal Health (VAH), a TGen subsidiary. WPDH, KB, and DH are VAH executive team members and TGen employees.*