The following supplement accompanies the article

Changes in immune functions in bottlenose dolphins in the northern Gulf of Mexico associated with the *Deepwater Horizon* oil spill

Sylvain De Guise*, Milton Levin, Erika Gebhard, Lindsay Jasperse, Leslie Burdett Hart, Cynthia R. Smith, Stephanie Venn-Watson, Forrest Townsend, Randall Wells, Brian Balmer, Eric Zolman, Teresa Rowles, Lori Schwacke

*Corresponding author: sylvain.deguise@uconn.edu


Animals

More than 90% of the potential pair-wise among-year comparisons of lymphocyte proliferation test results from SB were not statistically different. Therefore, dolphins sampled in SB during different years were pooled for use as reference.

Mice were used as quality control for functional assays, in order to discriminate between daily variability and true differences between dolphins analyzed at different times. Female B6C3F1 mice (*M. musculus*; Charles River Laboratories, MA) that were approximately 30 days-old were maintained at 18-26°C with relative humidity between 40 and 70%, and a light/dark cycle at 12-h intervals. Animals were housed five mice per cage containing sawdust (hardwood) bedding.

SB captures were conducted under NMFS Permit Nos. 522-1785 and 15543, while BB and MS captures were conducted under NMFS Permit No. 932-1905/MA-009526. Protocols were reviewed and approved by Mote Marine Laboratory Institutional Animal Care and Use Committee (IACUC) (SB) and NOAA Animal Care and Use Committees (BB and MS). All mouse procedures were approved by the IACUC at the University of Connecticut.

Blood sampling

Dolphin blood from the fluke blade periarterial venous rete was collected into BD Vacutainer® tubes with sodium heparin (Becton, Dickinson and Company, Franklin Lakes, NJ) as part of the physical examinations, kept cool and shipped overnight for functional immunological assays. In addition, blood was collected in Serum Separator tubes (SST), allowed to clot, and centrifuged to collect serum. One ml aliquots of serum were collected and immediately frozen prior to shipping on dry ice for cytokine analysis.

Mice were euthanized using CO₂ inhalation, followed by cervical dislocation to ensure death; the spleen was removed and stored in Dulbecco’s modified Eagle medium (DMEM; Gibco BRL, Grand Island, NY) until processing (below).

Isolation of lymphocytes

For the assessment of lymphocyte proliferation, peripheral blood mononuclear cells (PBMCs) were isolated by density gradient centrifugation on Ficoll-Paque Plus (GE Healthcare, Uppsala, Sweden) for 35 min at 900 g. PBMCs were then re-suspended in DMEM supplemented with 1 mM sodium pyruvate, 100 µM non-essential amino acids, 25 mM HEPES, 2 mM L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, and 0.25 µg/ml fungizone (all from Life Technologies, Grand Island, NY), along with 10% fetal bovine serum (Hyclone, Logan, UT),
thereafter referred to as complete DMEM. Cells were then washed three times, and cell counts and viability determined using trypan blue and light microscopy.

From each individual mouse spleen, a single cell suspension was prepared using two pairs of forceps in complete DMEM. PBMCs were isolated by density gradient centrifugation on Ficoll-Paque Plus for 15 min at 720 g for mice. Cell counts and viability were determined using trypan blue and light microscopy.

**Immune functions**

T and B cell proliferation are part of the battery of tests validated by the National Toxicology Program (Luster, Munson et al. 1988) and determined to be sensitive and predictive of immunomodulatory effects (Luster, Portier et al. 1992) and of increases in disease susceptibility (Luster, Portier et al. 1993). These assays are representative of acquired (T and B cell proliferation) immunity, can be evaluated *in vitro*, and that have been validated in marine mammals (De Guise, Bernier et al. 1996), and used in bottlenose dolphins (Schwacke, Twiner et al. 2010, Schwacke, Zolman et al. 2012).

Lymphocyte proliferation was evaluated as previously described (De Guise, Bernier et al. 1996). Briefly, lymphocytes were incubated with mitogens for 66 hours in flat-bottom 96-well plates (Falcon, Becton Dickinson, Lincoln Park, NJ) at 37°C with 5% CO₂. Mitogens chosen included two T cell mitogens (concanavalin A or Con-A, and phytohemagglutinin A or PHA) and a B cell mitogen (lipopolysaccharide or LPS). Mitogens were used at optimal as well as suboptimal concentrations, since suboptimal concentrations of mitogens allowed for higher sensitivity to subtle deficits when optimal concentrations of mitogens did not reveal differences (Mori, Morsey et al. 2006). Lymphocyte proliferation was evaluated by the incorporation of 5-bromo-2'-deoxyuridine (BrdU), a thymidine analogue, detected with a monoclonal antibody and colorimetric enzymatic reaction (Cell Proliferation ELISA BrdU (colorimetric), Roche Diagnostics GmbH, Mannheim, Germany) as per manufacturer’s instructions using an ELISA plate reader (Multiskan EX v.1.0) at 450 nm with a reference wavelength of 690 nm. This assay is a safe non-radioactive alternative to, and was shown to yield concordant results with, the classical radioactive thymidine incorporation (Messele, Roos et al. 2000). Results were expressed as optical density (OD) or as a stimulation index (SI), the ratio of OD upon mitogen stimulation to that of unstimulated cells.

Dolphin serum cytokines were quantified using the Bio-Plex Pro™ Human Cytokine Th1/Th2 Panel (Bio-Rad, Hercules, CA) and the Millipore Porcine 5-plex Panel (Millipore, Billerica, MA), according to the manufacturers’ instruction. Preliminary data in our lab has demonstrated the cross reactivity of those reagents in bottlenose dolphins. The Th1/Th2 cytokine kit included antibodies to detect IL-2, IL-4, IL-5, IL-10, IL-12, IL-13, IFNγ, TNFα, and GM-CSF. The porcine kit included antibodies to IL-4, IL-12, IL-8, IL-1ß, and IFNγ. After the incubation and conjugation process, the plates were measured on the Bio-Plex 200 system (Bio-Rad, Hercules, CA), and analyzed using Bio-Rad Manager 5.0. The observed concentration (pg/ml) of each analyte for each sample was calculated using a curve fit generated for each analyte from the eight standards. If a sample concentration was extrapolated outside the standard curve and designated as “Value extrapolated beyond standard range” by the software, that sample concentration was accepted as the calculated value. If a sample concentration was reported as “out of range” by the software, that sample concentration was given a 0 pg/ml value or the highest value on the standard curve, depending on whether it was below or above the measurable range. We obtained measurements from both the porcine and human kits for IL-4, IL-12, and IFNγ. We selected, and reported herein, the results from the reagents for which values were highest, with the lowest proportion of non-detectable concentrations.
Quality control

For functional assays (T and B lymphocyte proliferation), cells from mice (one mouse for each experimental day) were assayed concurrently with dolphin samples for quality control, as previously described (Schwacke, Zolman et al. 2012). After each field sampling effort, mouse data were assessed for the presence of outliers using the SPSS software (IBM SPSS Statistics version 21, Armonk, NY). If outliers were detected, it was assumed that normal daily variability for the assay was exceeded, and the corresponding dolphin data for that assay on that day were eliminated from the dataset. Further, all mouse data for the 2011-14 sampling seasons, from which this dataset was extracted, were pooled and once again assessed for the presence of outliers. If outliers were detected, it was assumed that normal variability for the assay as part of that dataset was exceeded, and the corresponding dolphin data for that assay on that day were eliminated from the dataset. The quality control process resulted in the elimination of 5% of the lymphocyte proliferation dolphin data points.

Prior to each use of the Bio-Plex 200 system, an instrument calibration and validation procedure using the Bio-Rad Validation and Calibration kit (Bio-Rad, Hercules, CA) was performed to assure the instrument was performing properly, as per manufacturers' instruction. For the Millipore kits, two quality control samples were included with each kit to assure accurate results. For the Bio-Rad kits, aliquots from archived tissue culture supernatant samples from the same humans were used across several kits to assure reproducibility. The instrument passed both calibration and validation tests prior to each use.

Immune Function Comparisons between Sites

Differences between groups (locations and time) were assessed using one-way analysis of variance (ANOVA), with the Dunnett’s post-hoc test used to compare groups to the SB reference group. When the assumptions of the test (normal distribution and equal variance) were not met, data were transformed (log and ln, in that order). If assumptions were still not met, Kruskal-Wallis one-way ANOVA on ranks was used with Dunn's Method for multiple comparisons versus the reference group. All analyses were performed using SigmaStat 3.5 (Systat, San Jose, CA), with p < 0.05 for statistical significance.

T-Cell Proliferation Reference Intervals

T-cell proliferation data (SubConA, OptConA, SubPHA and OptPHA) from dolphins sampled in SB between 2011 and 2014 were used to construct 95th percentile reference intervals (RI) using methods previously reported for bottlenose dolphin hematological and serum chemical parameters (Schwacke, Hall et al. 2009) and guidelines provided by the American Society for Veterinary Clinical Pathology (Friedrichs, Harr et al. 2012). For dolphins repeatedly sampled between these years, RI calculations used the first observation. Sample sizes varied for each parameter due to the exclusion of some observations that did not meet laboratory quality control standards. Statistical analyses and RI calculations were conducted using R, version 2.15.3 (R Core Team, R Foundation for Statistical Computing, 2013). Histograms were visually inspected for each parameter and a Shapiro-Wilks test was used to determine if the data were normally distributed. Annual variation was tested using a one-way ANOVA or a Kruskal-Wallis test, differences between sex were evaluated by a Student’s t-test or a Mann-Whitney U test, and correlations with age were computed with a Pearson or Spearman’s correlation test to determine the need for data partitioning by these demographic variables. Statistical significance for partitioning was determined using α=0.05. Non-Gaussian distributed data were transformed via Box-Cox methods, and outliers were evaluated using Tukey’s interquartile fences (Horn and Pesce 2003, Schwacke, Hall et al. 2009, Friedrichs, Harr et al. 2012). For all parameters, non-parametric bootstrap methods described by Schwacke et al. (Schwacke, Hall et al. 2009) were used to construct 95th percentile RIs and associated 90th percentile confidence intervals.
SubConA, OptConA, SubPHA and OptPHA observations from dolphins sampled in BB (2011, 2013, and 2014) and MS (2013) were compared to respective 95th percentile RIs. The prevalence and 95% confidence intervals of measurements that fell outside of the 95th RI were calculated for each parameter. Given that we would expect 5% of observations to fall outside of a 95th percentile RI, if the confidence interval for the calculated prevalence of out of range values did not contain 0.05 (our expected value), the prevalence of observations that fell outside the RI was considered significantly different than expected.

LITERATURE CITED


