NGS FAQ

1) **What types of sequencing kits are available for the NextSeq?**

There are five different kinds of NextSeq sequencing reagent kits: 75, 150 or 300 cycles in MID or HIGH read outputs as follows:

NextSeq's available sequencing kits

	75 cycles	150 cycles	300 cycles
MID output	-	Y	Y
HIGH output	Y	Y	Y

2) **What does "x cycles" mean for the NextSeq sequencing kits?**

The number of cycles refer to the maximum number of bases that can be sequenced with the particular kit.
The cycles required to read the barcode indices do not count against this number.
E.g. 300 cycles implies that a maximum of 300 bases in 1x300bp or 2x150bp format plus 2x8bp barcodes can be sequenced.
It is also important to note that you don't have to use all available cycles. E.g. one can run a 1x50bp experiment with the 75 cycles, HIGH output kit.

3) **How many reads can be obtained with the different output kits?**

Specs for the kits indicate that one can expect to get a maximum of 130 million and 400 million reads for the MID and HIGH output kits respectively.
Do note that this is the maximum optimal outputs that are available from the respective kits. Typically, one should expect to get less than maximum but it is common to get more reads than specs too. But read quality can suffers when the number of reads are way higher than specs.

4) **Why is it not possible to always get maximum number of reads on the NextSeq?**

Prepped libraries can be very different due to insert lengths, library prep kit manufacturers or even person-to-person variation, so it is very difficult to get consistent maximum number of reads for different libraries. In fact, we tend to be more conservative with new samples (new users/new organisms/new prep kit manufacturer), usually aiming for 100/300 million reads for MID/HIGH output respectively.
This is not a problem specific to the NextSeq; all Illumina models need to face this problem. A way to minimize variation is to be consistent with library quantification methods. E.g. Always using Qubit and TapeStation to quantify library concentration.

5) **Single reads vs paired-end reads**

In single reads, only one end of the nucleic acid fragments is sequenced.

In paired-end reads, the nucleic acid fragments are first sequenced from one end, and then from the other end. This allows generation of high quality, alignable sequence data. Depending on the application, paired-end reads usually do not overlap.

According to Illumina's website, "paired-end sequencing allows detection of genomic rearrangements and repetitive sequence elements, as well as gene fusions and novel transcripts".

6) **What gets sequenced on the NextSeq?**

Only the nucleic acid inserts/fragments are sequenced on the NextSeq. The P5 and P7 adapters added during library prep for hydridization to the flow cell does not get sequenced. The barcode sequences, needed for multiplexing, do get sequenced but separate from the inserts.

7) **What is multiplexing?**

Due to the large number of reads available from a single run of the NextSeq, it is often economical to combine multiple samples (multiplex) in a run to make full use of all available reads.

In such cases, there must be a way to distinguish the different samples from each other. This is by way of adding barcode indices to the different samples during library prep.

Barcode indices are 6 or 8 bp DNA sequences that are added to the samples during adapter ligations. After sequencing, NextSeq output reads are automatically sorted into different bins based on their barcode indices.

8) **Can any kind of nucleic samples/application be multiplexed?**

Any samples can be multiplexed but to get the best results, the following aspects need to be considered during multiplexing:
   i)   Fragment lengths from the different samples must match.
      - Different length libraries will hybridize to the flowcell at different efficiency so unmatched fragment lengths can lead to uneven coverage of samples
   ii) Same kind of barcode strategy must be used
      - All must be single indices or all must be dual indices
   iii) Barcode from the same type of kits must be used
      - E.g. TruSeq barcodes cannot be combined with Nextera barcodes
   iv) Illumina-equivalent indicies must be known if using non-Illumina library prep kits. The alternative is to know the sequence of the indices and use them as "custom barcode"

9) **How do I submit a sample to run on the NextSeq?**

   a) First, you will fill out the scheduling form on the BNL NGS website or by clicking this link.
   b) Next, you will fill out the Sample Barcode Template with the sequences of barcode used.
   c) Within 3 days, BNL will then assign you a date for sequencing.
   d) Finally, you will submit the sample to BNL at least two days before your scheduled date.

10) How do I get my data?

Sequencing data from the NextSeq is automatically uploaded to Illumina's cloud-based platform called BaseSpace.
Once a run is started, you will receive two links to BaseSpace: one will be a link to the Project folder which will contain the output data files once the run is completed. The other will be a link to the current run which can be used to check on the progress of the run as well as other metrics associated with the run.

11) Are DNA-Seq, RNA-Seq the only NGS techniques that can be ran on the NextSeq?

No, absolutely not. The NextSeq is "blind" to what NGS techniques/applications you are trying to do. As long as the library is prepped properly so that it will hybridize correctly to the flow-cell and there are correct binding sites for the enzymes, the NextSeq will sequence the library. Check out this Illumina poster for the >100 NGS experiments that can be performed on Illumina sequencers.