AMPLIFICATION METHODS AND SYSTEMS FOR MERFISH AND OTHER APPLICATIONS

RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Patent Application Serial No. 62/779,333, filed December 13, 2018, entitled "Amplification Methods and Systems for MERFISH and Other Applications," by Zhuang, *et al.*, incorporated herein by reference in its entirety.

GOVERNMENT FUNDING

This invention was made with government support under MH113094, MH111502, and MH114830 awarded by National Institutes of Health. The government has certain rights in the invention.

FIELD

The present invention generally relates to systems and methods for imaging or determining nucleic acids in cells or other samples.

BACKGROUND

Single-molecule fluorescence *in situ* hybridization (smFISH) provides both RNA abundance and spatial positions by the direct detection of individual RNA molecules in single cells. By using this technique, RNA localization has been shown to associate with multiple cell functions, like body patterning in development, cell fate determination during cell division, local translation, cell migration, and establishment of polarity. To study RNA localization at transcriptomic scale within intact cells and tissues, multiplexed error-robust fluorescence *in situ* hybridization (MERFISH) and *in situ* sequencing have been developed. Among these, MERFISH uses error-robust binary barcoding and sequential imaging to multiplex smFISH measurements with high detection efficiency. However, higher photon counts of the signal would improve throughput limited by imaging acquiring speed, light scattering caused by deep tissue imaging, and low signal to background due to sample autofluorescence. Thus, improvements in increasing photon counts are needed to allow substantial increases in imaging throughput, the targeting of short RNAs that may not allow the binding of sufficient probes to produce signals above the background signal levels, and/or to extend multiplexed measurements to sample types with high levels of background signal.

However, the performance of multiplexed single-molecule RNA imaging methods depends on many properties that signal amplification can degrade or challenge. For example, it is often important for the variation in the signal brightness from molecule to molecule to be

relatively small. Similarly, it is often important for the physical size or extent of the signal from individual molecules to be as small as possible to prevent substantial overlap of the signal from physically adjacent molecules. In parallel, amplification methods may not be high efficiency, and some molecules may not be amplified while others are, leading to a reduction in the fraction of detected molecules. In addition, it is often important to be able to amplify multiple, orthogonal, molecular signals and, critically, to be able to rapidly extend amplification methods to more distinct molecular signals. Finally, such amplification methods should be rapid to reduce sample preparation time. For these reasons, amplification methods that do not introduce signal brightness variation, increase the physical extent of the signal from individual molecules, which are highly efficient in the amplification of the signal from all target molecules, and which can be rapidly extended to multiple distinct targets are needed.

SUMMARY

The present invention generally relates to systems and methods for imaging or determining nucleic acids in cells or other samples. The subject matter of the present invention involves, in some cases, interrelated products, alternative solutions to a particular problem, and/or a plurality of different uses of one or more systems and/or articles.

In one aspect, the present invention is generally directed to a method. According to a first set of embodiments, the method comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where a maximum number of primary amplifier nucleic acids is able to bind to a nucleic acid probe, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, where a maximum number of secondary amplifier nucleic acids is able to bind to the primary amplifier nucleic acids, determining a distribution of the nucleic acid probes within the sample using fluorescence, creating codewords based on the fluorescence distribution within the sample, and for at least some of the codewords, matching the codeword to a valid codeword optionally where, if no match is found, applying error correction to the codeword to form a valid codeword.

The method, in another set of embodiments, comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, where the binding of primary amplifier nucleic acids and secondary amplifier nucleic acids to a target is saturatable, determining a distribution of the nucleic acid probes within the sample using

fluorescence, creating codewords based on the fluorescence distribution within the sample, and for at least some of the codewords, matching the codeword to a valid codeword optionally where, if no match is found, applying error correction to the codeword to form a valid codeword.

In yet another set of embodiments, the method comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, where the secondary amplifier nucleic acids bind to the primary amplifier nucleic acids within a fixed distance, determining a distribution of the nucleic acid probes within the sample using fluorescence, creating codewords based on the fluorescence distribution within the sample, and for at least some of the codewords, matching the codeword to a valid codeword optionally where, if no match is found, applying error correction to the codeword to form a valid codeword.

The method, according to still another set of embodiments, comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where the primary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, determining a distribution of the nucleic acid probes within the sample using fluorescence, creating codewords based on the fluorescence distribution within the sample, and for at least some of the codewords, matching the codeword to a valid codeword optionally where, if no match is found, applying error correction to the codeword to form a valid codeword.

The method, in yet another set of embodiments, comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, where the secondary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides, determining a distribution of the nucleic acid probes within the sample using fluorescence, creating codewords based on the fluorescence distribution within the sample, and for at least some of the codewords, matching the codeword to a valid codeword optionally where, if no match is found, applying error correction to the codeword to form a valid codeword.

In one set of embodiments, the method comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where a maximum number of primary amplifier nucleic acids is able to bind to a nucleic acid probe, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

The method, in another set of embodiments, comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where a maximum number of primary amplifier nucleic acids is able to bind to a nucleic acid probe, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample, and for at least some of the codewords, matching the codeword to a valid codeword where, if no match is found, potentially applying error correction to the codeword to form a valid codeword.

In yet another set of embodiments, the method comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where the binding of primary amplifier nucleic acids to a target is saturatable, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

The method, according to still another set of embodiments, comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where the binding of primary amplifier nucleic acids to a target is saturatable, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample, and for at least some of the codewords, matching the codeword to a valid codeword where, if no match is found, potentially applying error correction to the codeword to form a valid codeword.

In one set of embodiments, the method comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where the primary amplifier nucleic acids bind to the nucleic acid probes within a fixed distance, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

In another set of embodiments, the method comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where the primary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

The method, in accordance with yet another set of embodiments, comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where the primary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample and for at least some of the codewords, matching the codeword to a valid codeword where, if no match is found, potentially applying error correction to the codeword to form a valid codeword.

According to still another set of embodiments, the method comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where a maximum number of primary amplifier nucleic acids is able to bind to a nucleic acid probe, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, where a maximum number of secondary amplifier nucleic acids is able to bind to the primary amplifier nucleic acids, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

In another set of embodiments, the method comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, where the binding of primary amplifier nucleic acids and secondary amplifier nucleic acids to a target is saturatable, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

The method, in yet another set of embodiments, comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, exposing the primary amplifier nucleic acids to secondary

amplifier nucleic acids able to bind to the primary amplifier nucleic acids, where the secondary amplifier nucleic acids bind to the primary amplifier nucleic acids within a fixed distance, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

In still another set of embodiments, the method comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where the primary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

In yet another set of embodiments, the method comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, where the secondary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

The method, according to another set of embodiments, comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

In another set of embodiments, the method includes exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

According to still another set of embodiments, the method comprises exposing a sample to a binding entity conjugated to a nucleic acid probe, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, determining

a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

In yet another set of embodiments, the method comprises exposing a sample to a binding entity conjugated to a nucleic acid probe, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where the binding of primary amplifier nucleic acids to a target is saturatable, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

In one set of embodiments, the method comprises exposing a sample to a targeting entity conjugated to a nucleic acid probe, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

The method, in another set of embodiments, comprises exposing a sample to a targeting entity conjugated to a nucleic acid probe, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, where the binding of primary amplifier nucleic acids and secondary amplifier nucleic acids to a target is saturatable, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

According to yet another set of embodiments, the method comprises exposing a sample to a targeting entity conjugated to a nucleic acid probe, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, where the primary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

In still another set of embodiments, the method comprises exposing a sample to a targeting entity conjugated to a nucleic acid probe, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary

amplifier nucleic acids, where the secondary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides, determining a distribution of the nucleic acid probes within the sample using fluorescence, and creating codewords based on the fluorescence distribution within the sample.

In one set of embodiments, the method comprises exposing a sample to nucleic acid probes, where the nucleic acid probes comprise a first portion comprising a target sequence and a second portion comprising one or more read sequences, and where at least some of the plurality of nucleic acid probes comprises distinguishable nucleic acid probes formed from combinatorial combination of one or more read sequences taken from a plurality of read sequences, and exposing the nucleic acid probes to primary amplifier nucleic acids and secondary amplifier nucleic acids, where the primary amplifier nucleic acids are able to bind the nucleic acid probes and the secondary amplifier nucleic acids the primary nucleic acid probes, and where the maximum number of primary amplifier nucleic acids and secondary amplifier nucleic acids associated with a target within the sample is fixed.

In another set of embodiments, the method comprises exposing a sample to nucleic acid probes, exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, determining a distribution of secondary amplifier nucleic acids within the sample using fluorescence, creating codewords based on the fluorescence distribution within the sample, and for at least some of the codewords, matching the codeword to a valid codeword where, if no match is found, applying error correction to the codeword to form a valid codeword.

In another aspect, the present invention encompasses methods of making one or more of the embodiments described herein. In still another aspect, the present invention encompasses methods of using one or more of the embodiments described herein.

Other advantages and novel features of the present invention will become apparent from the following detailed description of various non-limiting embodiments of the invention when considered in conjunction with the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

Non-limiting embodiments of the present invention will be described by way of example with reference to the accompanying figures, which are schematic and are not intended to be drawn to scale. In the figures, each identical or nearly identical component illustrated is typically represented by a single numeral. For purposes of clarity, not every component is labeled in every figure, nor is every component of each embodiment of the

invention shown where illustration is not necessary to allow those of ordinary skill in the art to understand the invention. In the figures:

FIGs. 1A-1E provide a schematic illustration of an embodiment that uses primary and secondary amplifier nucleic acids to amplify a signal;

FIGs. 2A-2F illustrate that amplification can dramatically increase signal brightness without substantial changes in spot size, in accordance with another embodiment of the invention;

FIGs. 3A-3F illustrate MERFISH measurement of 130 RNAs, in yet another embodiment of the invention;

FIGs. 4A-4C illustrate amplification in tissue samples, in still another embodiment of the invention; and

FIGs. 5A-5C illustrate certain example sequences used for amplification, in one embodiment of the invention.

DETAILED DESCRIPTION

The present invention generally relates to systems and methods for imaging or determining nucleic acids in cells or other samples. In some cases, the transcriptome of a cell may be determined. Certain embodiments are generally directed to determining nucleic acids and other targets in a sample at relatively high resolutions. For instance, nucleic acid probes may be applied to sample, and binding of the nucleic acid probes to a target may be amplified using primary and secondary amplifier nucleic acids. In some cases, there is a maximum number of amplifier nucleic acids that can be bound to a target, e.g., the binding is saturatable, and cannot grow indefinitely, even in the presence of abundant reagents. This may be advantageous, for example, for controlling the brightness of each binding event, controlling the size of the amplified regions (e.g., during imaging), and/or for limiting the degree of amplification noise (i.e. the final variation in amplified signal from molecule to molecule), etc. In addition, in some embodiments, the primary and/or secondary amplifier nucleic acids may be formed from only 3 of the 4 naturally-occurring nucleotides, which may result in less secondary structure, faster binding rates, etc. These properties can in some cases facilitate the rapid design of multiple orthogonal amplification sequences, allowing the extension of such an approach to many distinct molecular targets.

In one aspect, the present invention is generally directed to systems and methods for amplifying the signal of targets (potentially tens, hundreds, thousands, or more) within a biological sample, e.g., for imaging using MERFISH or other techniques. For example, in some embodiments, these techniques provide a fast, simple, and/or efficient way to amplify

the signal of hundreds or thousands of RNA targets simultaneously, e.g., in the native environment of biological samples. Such amplification can be well-controlled by using a saturatable system, in certain embodiments, as discussed herein. Because of this, the variation in brightness from spot to spot can be minimized during the amplification, which can be useful in decoding using MERFISH or other techniques. In some embodiments, the sizes of the amplified spots also do not increase. This may improve the ability to identify targets, e.g., located relatively to close to each other. For example, the signal from one target may overlap with that from another target if the spot sizes increase too much. In addition, as discussed below, the amplifier nucleic acids in some embodiments do not contain hairpin structures, e.g., that may be involved in the amplification process, which may facilitate the creation of a saturatable system, and/or apply the design of multiple amplifier systems to a large number of targets. In addition, also as discussed below, the amplifier nucleic acids may be constructed using only three nucleotides. Three-letter nucleotides may have significantly less secondary structure than four-letter nucleotides, and faster binding rates. In addition, in some cases, the possibility that any given amplifier sequence will work reliably is increased, e.g., by reducing the possibility of unintentional secondary structure.

A non-limiting example of such a system is now illustrated in FIGs. 1A-1E. In FIG. 1A, a target 10 (RNA, in this example) is illustrated. There may be hundreds or thousands of targets distributed within a biological sample (e.g., a cell or a tissue), and the binding of nucleic acid probes to the targets may be used to determine their distribution, e.g., by using fluorescent probes and imaging the sample. Note, however, that only a single target is illustrated here for clarity.

In some embodiments, a plurality of nucleic acid probes are used that have different sequences, and the distribution of each of the nucleic acid probes is sequentially analyzed and used to create "codewords" for each location, based on the binding patterns of each of the nucleic acid probes. By selecting nucleic acid probes that defines a suitable code space, apparent errors in the observed binding patterns can be identified, and/or discarded and/or corrected to identify the correct codeword, and thus the correct target of the nucleic acid probes within the sample. This error-robustness and error-correction system was first introduced for multiplexed error-robust fluorescence *in situ* hybridization (MERFISH), and has also been subsequently used in various related techniques. See, e.g., Int. Pat. Apl. Pub. Nos. WO 2016/018960 and WO 2016/018963, each incorporated herein by reference in its entirety.

An example of an encoding nucleic acid probe is shown in FIG. 1A, where an encoding nucleic acid probe 15 (shown in a dotted box) has bound to a target 10, e.g., a target RNA. Other nucleic acid probes 16, 17 may also bind to the target RNA, and/or to other targets within the sample. Probe 15 may comprise a target sequence 11 that is able to bind to the target RNA (e.g., via specific binding), and a read sequence 12 (or "readout" sequence), i.e., a sequence that can be "read" to determine whether or not binding has occurred. One, two, three, or more read sequences may be present on a probe. For instance, in this example, two such read sequences are present in probe 15 (identified as read sequence 12 and read sequence 19). The read sequences may each independently be the same or different. In addition, probes such as 16 and 17 may have the same or different numbers of read sequences, and/or the same or different structures, as nucleic acid probe 15.

If no amplification is applied, then nucleic acid probe 15 may be exposed to a suitable secondary nucleic acid probe 32 containing a signaling entity 40, as is shown in FIG. 1E. In this example, the signaling entity is linked to the secondary nucleic acid probe via a disulfide linkage, although other techniques may be used in other embodiments. However, in this case, only one signaling entity can be linked to the target. It thus can be relatively difficult to detect the single signaling entity, and use it to determine binding of nucleic acid probe 15 to target 10, due to the low signal intensity produced after such a binding event.

Accordingly, in FIG. 1B, a primary amplifier nucleic acid 20 can be used, in accordance with certain embodiments. The primary amplifier nucleic acid may contain a first primary recognition sequence 22 able to bind (e.g., specifically) to a read sequence of nucleic acid probe 15, and one or more primary read sequences 23 able to bind (e.g., specifically) to one or more secondary amplifier nucleic acids, as discussed below. In this example, "N" such read sequences are shown schematically in the primary amplifier nucleic acid (N may be, for instance, 5, 7, 9, or other numbers as discussed herein). The primary read sequences may each have the same or different sequences, and may have the same or different lengths. In this example, each read sequence is 20 nucleotides long, although this is by way of example only. In addition, as previously noted, although two such primary amplifier nucleic acids are shown here, this is by way of example only, and other numbers of primary amplifier nucleic acids may be bound to the nucleic acid probe in other embodiments.

Next, as shown in FIG. 1C, secondary amplifier nucleic acids 30 may be bound to the primary amplifier nucleic acid. A secondary amplifier nucleic acid may contain a first recognition sequence 33 able to bind (e.g., specifically) to read sequence 23 of primary

amplifier nucleic acid 20, and one or more secondary read sequences 34 able to bind to a signaling entity, as discussed below.

As with the primary amplifier nucleic acids, any number of secondary read sequences may be present in the secondary amplifier nucleic acids, as is shown in this figure. The secondary read sequences may each have the same or different sequences, and may have the same or different lengths relative to each other. The secondary read sequences also may be the same or different than the read sequences of the primary amplifier nucleic acids. In this example, each secondary amplifier nucleic acid may have "M" read sequences. M may be, e.g., 5, 7, 9, or other numbers as discussed herein, and M may be the same or different than N.

In FIG. 1D, a plurality of signaling entities 40 has been bound to the read sequences of the secondary amplifier nucleic acids. In this example, the signaling entities are each bound via a disulfide linkage, although other techniques may be used in other embodiments, as discussed herein.

Additionally, in this case, there is expected to be a maximum, or a saturation limit, of signaling entities that have bound to each read sequence of the target sequence. In this particular example, there are NxM such positions available for each of the read sequences of the nucleic acid probe (2 such read sequences here), assuming both have substantially the same structure (although they do not necessarily have to have the same structure, i.e., the same number of NxM positions available, e.g., if they have amplifier nucleic acids with different structures). Thus, the number of signaling entities that can be associated with a given target is a finite, predictable number, and cannot grow indefinitely or without bound.

In this example, two read sequences 12 and 19 were discussed, each of which may have primary and secondary amplifier nucleic acids and associated signaling entities. These may or may not have the same or different structures, e.g., signaling entities and/or amplifier nucleic acids associated with read sequence 12 may not associate with read sequence 19, and vice versa. (However, as mentioned above, 2 read sequences are provided here by way of example only, and in other embodiments, there may be 1, 2, 3, 4, etc. distinct read sequences, which can be amplified in parallel, e.g., using similar approaches, including with distinct amplifier nucleic acids, etc.) The read sequences may be independently determined, e.g., sequentially or simultaneously, by determining signaling entities associated with each of the read sequences, which may be the same or different. For instance, as shown in FIG. 1D, signaling entities 40 are able to associate with primary amplifier nucleic acid 20 and secondary amplifier nucleic acid 30, ultimately to read sequence 12, but are not able to

associate with read sequence 19 or its associated primary and secondary amplifier nucleic acids 29 and 39, respectively.

In addition, in some embodiments, amplification can involve the binding of only one round of amplifier nucleic acids (producing a N-fold amplification), two rounds (producing NxM-fold amplification), three rounds (producing a NxMxO amplification where the third round of molecules contain O read sequences), or more in some cases. In some cases, any number of rounds of amplification can be applied.

Additionally, the sample may contain nucleic acid probes that have many different readout sequences, e.g., that can be recognized using different amplifier nucleic acids or signaling entities. For example, 8, 10, 12, 14, 16, 24, 32, 48, 64, or other numbers of readout sequences could be used, including more than 64 rounds. In some cases, unique amplifier nucleic acids can be used for the amplification of the readout sequences, e.g., such as that the original readout sequence present can be amplified into, e.g., NxM copies, using suitable amplifier nucleic acids. In certain cases, the amplifier nucleic acids may be efficiently designed. For example, by utilizing only three of the four nucleotides in the sequences of the amplifier nucleic acids, the probability of unanticipated secondary structure within the amplifier nucleic acids can be reduced. Without wishing to be bound by any theory, it is believed that because the effect of such secondary structure of binding of the amplifier nucleic acids themselves, or subsequent amplifier nucleic acids can be difficult to predict, reducing or eliminating secondary structure can increase the probability that a given amplifier nucleic acid will properly assemble, which can facilitate the design and use of large numbers of orthogonal amplifier nucleic acids. In addition, by reducing secondary structure, which can also inhibit the rate of binding of these amplifier nucleic acids, such design considerations can decrease the time required to assemble such structures, or decrease the amount of amplifier nucleic acids required for each sample, etc.

The controlled amplification provided by this approach is in contrast to techniques such as hairpin unfolding or rolling circle amplification, where amplification of a signal can effectively grow in an uncontrolled manner or indefinitely, i.e., when sufficient reagents are present. Such uncontrolled amplification can be difficult to accurately determine, as the amount of signal present may not be well-correlated to the number of targets, or the location of the target (for example, with larger amounts of signal created by uncontrolled amplification, the "spot size" appearing in a microscopic image may grow larger, and not necessarily centered around the target, thus impeding the resolution of the image, or interfering with signals from other, nearby targets). In contrast, the use of saturatable

amplification techniques, as discussed herein, may create a maximum number of signaling entities that can associate with a target, which may limit spot sizes, create uniformity in the brightness or intensities of the spots, improve detection, or the like.

As mentioned, in certain embodiments, such techniques may be combined with error correction, e.g., as is used in MERFISH or other similar techniques. For example, codewords may be based on the binding (or non-binding) of the plurality of nucleic acid probes, and in some cases, the codewords may define an error-correcting code to help reduce or prevent misidentification of the nucleic acid probes. In some cases, a relatively large number of different targets may be identified using a relatively small number of labels, e.g., by using various combinatorial approaches. Image acquisition techniques such as STORM can also be used to image such samples and facilitate determination of the nucleic acid probes. See, e.g., U.S. Pat. Nos. 9,712,805 or 10,073,035, or Int. Pat. Apl. Pub. Nos. WO 2008/091296 or WO 2009/085218, each incorporated herein by reference in its entirety, for additional details regarding techniques such as MERFISH.

The above discussion is a non-limiting example of one embodiment of the present invention that can be used to improve determination of targets in a sample using MERFISH or other techniques. However, other embodiments are also possible, in addition to the ones described above. Accordingly, more generally, various aspects of the invention are directed to various systems and methods for imaging or determining nucleic acids in cells or other samples.

Certain aspects thus are directed to determining a sample, which may include a cell culture, a suspension of cells, a biological tissue, a biopsy, an organism, or the like. The sample can also be cell-free but nevertheless contain nucleic acids in some cases. If the sample contains a cell, the cell may be a human cell, or any other suitable cell, e.g., a mammalian cell, a fish cell, an insect cell, a plant cell, or the like. More than one cell may be present in some cases.

Within the sample, the targets to be determined can include nucleic acids, proteins, or the like. Nucleic acids to be determined may include, for example, DNA (for example, genomic DNA), RNA, or other nucleic acids that are present within a cell (or other sample). The nucleic acids may be endogenous to the cell, or added to the cell. For instance, the nucleic acid may be viral, or artificially created. In some cases, the nucleic acid to be determined may be expressed by the cell. The nucleic acid is RNA in some embodiments. The RNA may be coding and/or non-coding RNA. For example, the RNA may encode a

protein. Non-limiting examples of RNA that may be studied within the cell include mRNA, siRNA, rRNA, miRNA, tRNA, lncRNA, snoRNAs, snRNAs, exRNAs, piRNAs, or the like.

In some cases, a significant portion of the nucleic acid within the cell may be studied. For instance, in some cases, enough of the RNA present within a cell may be determined so as to produce a partial or complete transcriptome of the cell. In some cases, at least 4 types of mRNAs are determined within a cell, and in some cases, at least 3, at least 4, at least 7, at least 8, at least 12, at least 14, at least 15, at least 16, at least 22, at least 30, at least 31, at least 32, at least 50, at least 63, at least 64, at least 72, at least 75, at least 100, at least 127, at least 128, at least 140, at least 255, at least 256, at least 500, at least 1,000, at least 1,500, at least 2,000, at least 2,500, at least 3,000, at least 4,000, at least 5,000, at least 7,500, at least 10,000, at least 12,000, at least 15,000, at least 20,000, at least 25,000, at least 30,000, at least 40,000, at least 50,000, at least 75,000, or at least 100,000 types of mRNAs may be determined within a cell.

In some cases, the transcriptome of a cell may be determined. It should be understood that the transcriptome generally encompasses all RNA molecules produced within a cell, not just mRNA. Thus, for instance, the transcriptome may also include rRNA, tRNA, siRNA, etc. in certain instances. In some embodiments, at least 5%, at least 10%, at least 15%, at least 20%, at least 25%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or 100% of the transcriptome of a cell may be determined.

In some embodiments, other targets to be determined can include targets that are linked to nucleic acids, proteins, or the like. For instance, in one set of embodiments, a binding entity able to recognize a target may be conjugated to a nucleic acid probe. The binding entity may be any entity that can recognize a target, e.g., specifically or non-specifically. Non-limiting examples include enzymes, antibodies, receptors, complementary nucleic acid strands, aptamers, or the like. For example, an oligonucleotide-linked antibody may be used to determine a target. The target may bind to the oligonucleotide-linked antibody, and the oligonucleotides determined as discussed herein.

The determination of targets, such as nucleic acids within the cell or other sample, may be qualitative and/or quantitative. In addition, the determination may also be spatial, e.g., the position of the nucleic acids, or other targets, within the cell or other sample may be determined in two or three dimensions. In some embodiments, the positions, number, and/or concentrations of nucleic acids, or other targets, within the cell or other sample may be determined.

In some cases, a significant portion of the genome of a cell may be determined. The determined genomic segments may be continuous or interspersed on the genome. For example, in some cases, at least 4 genomic segments are determined within a cell, and in some cases, at least 3, at least 4, at least 7, at least 8, at least 12, at least 14, at least 15, at least 16, at least 22, at least 30, at least 31, at least 32, at least 50, at least 63, at least 64, at least 72, at least 75, at least 100, at least 127, at least 128, at least 140, at least 255, at least 256, at least 500, at least 1,000, at least 1,500, at least 2,000, at least 2,500, at least 3,000, at least 4,000, at least 5,000, at least 7,500, at least 10,000, at least 12,000, at least 15,000, at least 20,000, at least 25,000, at least 30,000, at least 40,000, at least 50,000, at least 75,000, or at least 100,000 genomic segments may be determined within a cell.

In some cases, the entire genome of a cell may be determined. It should be understood that the genome generally encompasses all DNA molecules produced within a cell, not just chromosome DNA. Thus, for instance, the genome may also include, in some cases, mitochondria DNA, chloroplast DNA, plasmid DNA, etc., e.g., in addition to (or instead of) chromosome DNA. In some embodiments, at least about 5%, at least about 10%, at least about 15%, at least about 20%, at least about 25%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, or 100% of the genome of a cell may be determined.

As discussed herein, a variety of nucleic acid probes may be used to determine one or more targets within a cell or other sample. The probes may comprise nucleic acids (or entities that can hybridize to a nucleic acid, e.g., specifically) such as DNA, RNA, LNA (locked nucleic acids), PNA (peptide nucleic acids), and/or combinations thereof. In some cases, additional components may also be present within the nucleic acid probes, e.g., as discussed below. In addition, any suitable method may be used to introduce nucleic acid probes into a cell.

For example, in some embodiments, the cell is fixed prior to introducing the nucleic acid probes, e.g., to preserve the positions of the nucleic acids or other targets within the cell. Techniques for fixing cells are known to those of ordinary skill in the art. As non-limiting examples, a cell may be fixed using chemicals such as formaldehyde, paraformaldehyde, glutaraldehyde, ethanol, methanol, acetone, acetic acid, or the like. In one embodiment, a cell may be fixed using HEPES-glutamic acid buffer-mediated organic solvent (HOPE).

The nucleic acid probes may be introduced into the cell (or other sample) using any suitable method. In some cases, the cell may be sufficiently permeabilized such that the nucleic acid probes may be introduced into the cell by flowing a fluid containing the nucleic

acid probes around the cells. In some cases, the cells may be sufficiently permeabilized as part of a fixation process; in other embodiments, cells may be permeabilized by exposure to certain chemicals such as ethanol, methanol, Triton, or the like. In addition, in some embodiments, techniques such as electroporation or microinjection may be used to introduce nucleic acid probes into a cell or other sample.

Certain aspects are thus generally directed to nucleic acid probes that are introduced into a cell (or other sample). The probes may comprise any of a variety of entities that can hybridize to a nucleic acid, typically by Watson-Crick base pairing, such as DNA, RNA, LNA, PNA, etc., depending on the application. The nucleic acid probe typically contains a target sequence that is able to bind to at least a portion of a target, e.g., a target nucleic acid. In some cases, the binding may be specific binding (e.g., via complementary binding). When introduced into a cell or other system, the target sequence may be able to bind to a specific target (e.g., an mRNA, or other nucleic acids as discussed herein). The nucleic acid probe may also contain one or more read sequences, as discussed below.

In some cases, more than one type of nucleic acid probe may be applied to a sample, e.g., sequentially or simultaneously. For example, there may be at least 2, at least 5, at least 10, at least 25, at least 50, at least 75, at least 100, at least 300, at least 1,000, at least 3,000, at least 10,000, or at least 30,000 distinguishable nucleic acid probes that are applied to a sample. In some cases, the nucleic acid probes may be added sequentially. However, in some cases, more than one nucleic acid probe may be added simultaneously.

The nucleic acid probe may include one or more target sequences, which may be positioned anywhere within the nucleic acid probe. The target sequence may contain a region that is substantially complementary to a portion of a target, e.g., a target nucleic acid. For instance, in some cases, the portions may be at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary, e.g., to produce specific binding. Typically, complementarity is determined on the basis of Watson-Crick nucleotide base pairing.

In some cases, the target sequence may be at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 50, at least 60, at least 65, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, at least 350, at least 400, or at least 450 nucleotides in length. In some cases, the target sequence may be no more than 500, no more than 450, no more than 400, no more than 350, no more than 300, no more than 250, no more than 200, no more than 175, no more than 150, no more

than 125, no more than 100, be no more than 75, no more than 60, no more than 65, no more than 60, no more than 55, no more than 50, no more than 45, no more than 40, no more than 35, no more than 30, no more than 20, or no more than 10 nucleotides in length. Combinations of any of these are also possible, e.g., the target sequence may have a length of between 10 and 30 nucleotides, between 20 and 40 nucleotides, between 5 and 50 nucleotides, between 10 and 200 nucleotides, or between 25 and 35 nucleotides, between 10 and 300 nucleotides, etc.

The target sequence of a nucleic acid probe may be determined with reference to a target suspected of being present within a cell or other sample. For example, a target nucleic acid to a protein may be determined using the protein's sequence, e.g., by determining the nucleic acids that are expressed to form the protein. In some cases, only a portion of the nucleic acids encoding the protein are used, e.g., having the lengths as discussed above. In addition, in some cases, more than one target sequence that can be used to identify a particular target may be used. For instance, multiple probes can be used, sequentially and/or simultaneously, that can bind to or hybridize to the same or different regions of the same target. Hybridization typically refers to an annealing process by which complementary single-stranded nucleic acids associate through Watson-Crick nucleotide base pairing (e.g., hydrogen bonding, guanine-cytosine and adenine-thymine) to form double-stranded nucleic acid.

In some embodiments, a nucleic acid probe may also comprise one or more "read" sequences. The read sequences may be used, to identify the nucleic acid probe, e.g., through association with signaling entities, as discussed below. In some embodiments, the nucleic acid probe may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 or more, 20 or more, 24 or more, 32 or more, 40 or more, 48 or more, 50 or more, 64 or more, 75 or more, 100 or more, 128 or more read sequences. The read sequences may be positioned anywhere within the nucleic acid probe. If more than one read sequence is present, the read sequences may be positioned next to each other, and/or interspersed with other sequences.

The read sequences may be of any length. If more than one read sequence is used, the read sequences may independently have the same or different lengths. For instance, the read sequence may be at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 50, at least 60, at least 65, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, at least 350, at least 400, or at least 450 nucleotides in length. In some cases, the read sequence may be no more than 500, no more than 450, no more than 400, no more than 350, no more than 300, no more than 250, no

more than 200, no more than 175, no more than 150, no more than 125, no more than 100, be no more than 75, no more than 60, no more than 65, no more than 60, no more than 55, no more than 50, no more than 45, no more than 40, no more than 35, no more than 30, no more than 20, or no more than 10 nucleotides in length. Combinations of any of these are also possible, e.g., the read sequence may have a length of between 10 and 30 nucleotides, between 20 and 40 nucleotides, between 5 and 50 nucleotides, between 10 and 200 nucleotides, or between 25 and 35 nucleotides, between 10 and 300 nucleotides, etc.

The read sequence may be arbitrary or random in some embodiments. In certain cases, the read sequences are chosen so as to reduce or minimize homology with other components of the cell or other sample, e.g., such that the read sequences do not themselves bind to or hybridize with other nucleic acids suspected of being within the cell or other sample. In some cases, the homology may be less than 10%, less than 8%, less than 7%, less than 6%, less than 5%, less than 4%, less than 3%, less than 2%, or less than 1%. In some cases, there may be a homology of less than 20 basepairs, less than 18 basepairs, less than 15 basepairs, less than 14 basepairs, less than 13 basepairs, less than 12 basepairs, less than 11 basepairs, or less than 10 basepairs. In some cases, such basepairs are sequential.

In one set of embodiments, a population of nucleic acid probes may contain a certain number of read sequences, which may be less than the number of targets of the nucleic acid probes in some cases. Those of ordinary skill in the art will be aware that if there is one signaling entity and *n* read sequences, then in general 2^n-1 different nucleic acid targets may be uniquely identified. However, not all possible combinations need be used. For instance, a population of nucleic acid probes may target 12 different nucleic acid sequences, yet contain no more than 8 read sequences. As another example, a population of nucleic acids may target 140 different nucleic acid species, yet contain no more than 16 read sequences. Different nucleic acid sequence targets may be separately identified by using different combinations of read sequences within each probe. For instance, each probe may contain 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, etc. or more read sequences. In some cases, a population of nucleic acid probes may each contain the same number of read sequences, although in other cases, there may be different numbers of read sequences present on the various probes.

As a non-limiting example, a first nucleic acid probe may contain a first target sequence, a first read sequence, and a second read sequence, while a second, different nucleic acid probe may contain a second target sequence, the same first read sequence, but a third read sequence instead of the second read sequence. Such probes may thereby be distinguished by determining the various read sequences present or associated with a given

probe or location, as discussed herein. For example, the probes can be sequentially identified and encoded using "codewords," as discussed below. Optionally, the codewords may also be subjected to error detection and/or correction.

In addition, the population of nucleic acid probes (and their corresponding, complimentary sites on the encoding probes), in certain embodiments, may be made using only 2 or only 3 of the 4 naturally occurring nucleotide bases, such as leaving out all the "G"s or leaving out all of the "C"s within the population of probes. Sequences lacking either "G"s or "C"s may form very little secondary structure in certain embodiments, and can contribute to more uniform, faster hybridization. Thus, in some cases, the nucleic acid probes may contain only A, T, and G; only A, T, and C; only A, C, and G; or only T, C, and G.

In one aspect, the read sequences on the nucleic acid probes may be able to bind (e.g., specifically) to corresponding recognition sequences on the primary amplifier nucleic acids. Thus, when a nucleic acid probe recognizes a target within a biological sample, e.g., a DNA or RNA target, the primary amplifier nucleic acid are also able to associate with the target via the nucleic acid probe, with interactions between the read sequences of the nucleic acid probes and corresponding recognition sequences on the primary amplifier nucleic acids, e.g., complementary binding. For instance, the recognition sequence may be able to recognize a target read sequence, but not substantially recognize or bind to other, non-target read sequence. The primary amplifier nucleic acids may also comprise any of a variety of entities able to hybridize a nucleic acid, e.g., DNA, RNA, LNA, and/or PNA, etc., depending on the application. For instance, such entities may form some or all of the recognition sequence.

In some cases, the recognition sequence may be substantially complementary to the target read sequence. In some cases, the sequences may be at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary. Typically, complementarity is determined on the basis of Watson-Crick nucleotide base pairing. The structures of the target read sequence may include those previously described.

In some cases, the recognition sequence may be at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 50, at least 60, at least 65, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, at least 350, at least 400, or at least 450 nucleotides in length. In some cases, the recognition sequence may be no more than 500, no more than 450, no more than 400, no more than 350, no more than 300, no more than 250, no more than 200, no more than 175, no more than 150, no more than 125, no more than 100, be no more than 75, no more than 60, no

more than 65, no more than 60, no more than 55, no more than 50, no more than 45, no more than 40, no more than 35, no more than 30, no more than 20, or no more than 10 nucleotides in length. Combinations of any of these are also possible, e.g., the recognition sequence may have a length of between 10 and 30 nucleotides, between 20 and 40 nucleotides, between 5 and 50 nucleotides, between 10 and 200 nucleotides, or between 25 and 35 nucleotides, between 10 and 300 nucleotides, etc.

In some embodiments, a primary amplifier nucleic acid may also comprise one or more read sequences able to bind to secondary amplifier nucleic acids, as discussed below. For example, a primary amplifier nucleic acid may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 or more, 20 or more, 32 or more, 40 or more, 50 or more, 64 or more, 75 or more, 100 or more, 128 or more read sequences. The read sequences may be positioned anywhere within the primary amplifier nucleic acid. If more than one read sequence is present, the read sequence may be positioned next to each other, and/or interspersed with other sequences. In one embodiment, the primary amplifier nucleic acid comprises a recognition sequence at a first end and a plurality of read sequences at a second end.

In some cases, a read sequence within the primary amplifier nucleic acid may be at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 50, at least 60, at least 65, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, at least 350, at least 400, or at least 450 nucleotides in length. In some cases, the read sequence may be no more than 500, no more than 450, no more than 400, no more than 350, no more than 300, no more than 250, no more than 200, no more than 175, no more than 150, no more than 125, no more than 100, be no more than 75, no more than 60, no more than 65, no more than 60, no more than 55, no more than 50, no more than 45, no more than 40, no more than 35, no more than 30, no more than 20, or no more than 10 nucleotides in length. Combinations of any of these are also possible, e.g., the read sequence may have a length of between 10 and 20 nucleotides, between 10 and 30 nucleotides, between 20 and 40 nucleotides, between 5 and 50 nucleotides, between 10 and 200 nucleotides, or between 25 and 35 nucleotides, between 10 and 300 nucleotides, etc.

There may be any number of read sequences within a primary amplifier nucleic acid. For example, there may be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more read sequences present within a primary amplifier nucleic acid. If more than one read sequence is present within a primary amplifier nucleic acid, the read sequences may be the same or different. In some cases, for example, the read sequences may all be identical.

In some embodiments, the population of primary amplifier nucleic acids may be made using only 2 or only 3 of the 4 naturally occurring nucleotide bases, such as leaving out all the "G"s or leaving out all of the "C"s within the population of nucleic acids. Sequences lacking either "G"s or "C"s may form very little secondary structure in certain embodiments, and can contribute to more uniform, faster hybridization. Thus, in some cases, the primary amplifier nucleic acids may contain only A, T, and G; only A, T, and C; only A, C, and G; or only T, C, and G.

In some cases, more than one type of primary amplifier nucleic acid may be applied to a sample, e.g., sequentially or simultaneously. For example, there may be at least 2, at least 5, at least 10, at least 25, at least 50, at least 75, at least 100, at least 300, at least 1,000, at least 3,000, at least 10,000, or at least 30,000 distinguishable primary amplifier nucleic acids that are applied to a sample. In some cases, the primary amplifier nucleic acids may be added sequentially. However, in some cases, more than one primary amplifier nucleic acid may be added simultaneously.

In one set of embodiments, the read sequences on the primary amplifier nucleic acids may be able to bind (e.g., specifically) to corresponding recognition sequences on the secondary amplifier nucleic acids. Thus, when a nucleic acid probe recognizes a target within a biological sample, e.g., a DNA or RNA target, the secondary amplifier nucleic acids are also able to associate with the target, via the primary amplifier nucleic acids, with interactions between the read sequences of the primary amplifier nucleic acids and corresponding recognition sequences on the secondary amplifier nucleic acids, e.g., complementary binding. For instance, the recognition sequence on a secondary amplifier nucleic acid may be able to recognize a read sequence on a primary amplifier nucleic acid, but not substantially recognize or bind to other, non-target read sequence. The secondary amplifier nucleic acids may also comprise any of a variety of entities able to hybridize a nucleic acid, e.g., DNA, RNA, LNA, and/or PNA, etc., depending on the application. For instance, such entities may form some or all of the recognition sequence.

In some cases, the recognition sequence on the secondary amplifier nucleic acid may be substantially complementary to a read sequence on a primary amplifier nucleic acid. In some cases, the sequences may be at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 92%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% complementary.

In some cases, the recognition sequence on the secondary amplifier nucleic acid may be at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40,

at least 50, at least 60, at least 65, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, at least 350, at least 400, or at least 450 nucleotides in length. In some cases, the recognition sequence may be no more than 500, no more than 450, no more than 400, no more than 350, no more than 300, no more than 250, no more than 200, no more than 175, no more than 150, no more than 125, no more than 100, be no more than 75, no more than 60, no more than 65, no more than 60, no more than 55, no more than 50, no more than 45, no more than 40, no more than 35, no more than 30, no more than 20, or no more than 10 nucleotides in length. Combinations of any of these are also possible, e.g., the recognition sequence may have a length of between 10 and 30 nucleotides, between 20 and 40 nucleotides, between 5 and 50 nucleotides, between 10 and 200 nucleotides, or between 25 and 35 nucleotides, between 10 and 300 nucleotides, etc.

In some embodiments, a secondary amplifier nucleic acid may also comprise one or more read sequences able to bind to a signaling entity, as discussed herein. For example, a secondary amplifier nucleic acid may comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 or more, 20 or more, 32 or more, 40 or more, 50 or more, 64 or more, 75 or more, 100 or more, 128 or more read sequences able to bind to a signaling entity. The read sequences may be positioned anywhere within the secondary amplifier nucleic acid. If more than one read sequences is present, the read sequences may be positioned next to each other, and/or interspersed with other sequences. In one embodiment, the secondary amplifier nucleic acid comprises a recognition sequence at a first end and a plurality of read sequences at a second end. This structure may also be the same or different than the structure of the primary amplifier nucleic acid.

In some cases, the read sequence within the secondary amplifier nucleic acid may be at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 50, at least 60, at least 65, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, at least 350, at least 400, or at least 450 nucleotides in length. In some cases, the read sequence may be no more than 500, no more than 450, no more than 400, no more than 350, no more than 300, no more than 250, no more than 200, no more than 175, no more than 150, no more than 125, no more than 100, be no more than 75, no more than 60, no more than 65, no more than 60, no more than 55, no more than 50, no more than 45, no more than 40, no more than 35, no more than 30, no more than 20, or no more than 10 nucleotides in length. Combinations of any of these are also possible, e.g., the read sequence within the secondary amplifier nucleic acid may have a length of between 10 and 20 nucleotides, between 10 and 30 nucleotides, between 20 and 40 nucleotides, between

5 and 50 nucleotides, between 10 and 200 nucleotides, or between 25 and 35 nucleotides, between 10 and 300 nucleotides, etc.

There may be any number of read sequences within a secondary amplifier nucleic acid. For example, there may be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, or more read sequences present within a secondary amplifier nucleic acid. If more than one read sequence is present within a secondary amplifier nucleic acid, the read sequences may be the same or different. In some cases, for example, the read sequences may all be identical. In addition, there may independently be the same or different numbers of read sequences in the primary and in the secondary amplifier nucleic acids.

The population of secondary amplifier nucleic acids may be made using only 2 or only 3 of the 4 naturally occurring nucleotide bases, in certain embodiments such as leaving out all the "G"s or leaving out all of the "C"s within the population of nucleic acids. Sequences lacking either "G"s or "C"s may form very little secondary structure in certain embodiments, and can contribute to more uniform, faster hybridization. Thus, in some cases, the secondary amplifier nucleic acids may contain only A, T, and G; only A, T, and C; only A, C, and G; or only T, C, and G.

In some cases, more than one type of secondary amplifier nucleic acid may be applied to a sample, e.g., sequentially or simultaneously. For example, there may be at least 2, at least 5, at least 10, at least 25, at least 50, at least 75, at least 100, at least 300, at least 1,000, at least 3,000, at least 10,000, or at least 30,000 distinguishable secondary amplifier nucleic acids that are applied to a sample. In some cases, the secondary amplifier nucleic acids may be added sequentially. However, in some cases, more than one secondary amplifier nucleic acid may be added simultaneously.

In addition, in certain embodiments, this pattern can instead be repeated prior to the signaling entity, e.g., with tertiary amplifier nucleic acids, quaternary nucleic acids, etc., similar to the above discussion. The signaling entities may thus be bound to the ending amplifier nucleic acid. Thus, as non-limiting examples, to a target may be bound an encoding nucleic acid probe, to which a primary amplifier nucleic acid is bound, to which a secondary amplifier nucleic acid is bound, to which a tertiary amplifier nucleic acid is bound, to which a signaling entity is bound, or to a target may be bound an encoding nucleic acid probe, to which a primary amplifier nucleic acid is bound, to which a secondary amplifier nucleic acid is bound, to which a tertiary amplifier nucleic acid is bound, to which a quaternary amplifier nucleic acid is bound, to which a signaling entity is bound, etc. Accordingly, the ending

amplifier nucleic acid need not necessarily be the secondary amplifier nucleic acid in all embodiments.

Other components may also be present within a nucleic acid probe or an amplifier nucleic acid as well. For example, in one set of embodiments, one or more primer sequences may be present, e.g., to facilitate enzymatic amplification. Those of ordinary skill in the art will be aware of primer sequences suitable for applications such as amplification (e.g., using PCR or other suitable techniques). Many such primer sequences are available commercially. Other examples of sequences that may be present within a primary nucleic acid probe include, but are not limited to promoter sequences, operons, identification sequences, nonsense sequences, or the like.

Typically, a primer is a single-stranded or partially double-stranded nucleic acid (e.g., DNA) that serves as a starting point for nucleic acid synthesis, allowing polymerase enzymes such as nucleic acid polymerase to extend the primer and replicate the complementary strand. A primer is (e.g., is designed to be) complementary to and to hybridize to a target nucleic acid. In some embodiments, a primer is a synthetic primer. In some embodiments, a primer is a non-naturally-occurring primer. A primer typically has a length of 10 to 50 nucleotides. For example, a primer may have a length of 10 to 40, 10 to 30, 10 to 20, 25 to 50, 15 to 40, 15 to 30, 20 to 50, 20 to 40, or 20 to 30 nucleotides. In some embodiments, a primer has a length of 18 to 24 nucleotides.

In some embodiments, one or more signaling entities may be bound to the recognition entities on the secondary amplifier nucleic acids (or other ending amplifier nucleic acid). Non-limiting examples of signaling entities include fluorescent entities (fluorophores) or phosphorescent entities, e.g., as discussed below. The signaling entities may then be determined, e.g., to determine the nucleic acid probes or the targets. In some cases, the determination may be spatial, e.g., in two or three dimensions. In addition, in some cases, the determination may be quantitative, e.g., the amount or concentration of signaling entity and/or of a target may be determined.

In one set of embodiments, the signaling entities may be attached to the secondary amplifier nucleic acid (or other ending amplifier nucleic acid). The signaling entities may be attached to the secondary amplifier nucleic acid (or other ending amplifier nucleic acid) before or after association of the secondary amplifier nucleic acid to targets within the sample. For example, the signaling entities may be attached to the secondary amplifier nucleic acid initially, or after the secondary amplifier nucleic acids have been applied to a

sample. In some cases, the signaling entities are added, then reacted to attach them to the amplifier nucleic acids.

In one set of embodiments, the signaling entities may be attached to a nucleotide sequence via a bond that can be cleaved to release the signaling entity. For example, after determine the distribution of nucleic acid probes within a sample, the signaling entities may be released or inactivated, prior to another round of nucleic acid probes and/or amplifier nucleic acids. Thus, in some embodiments, the bond may be a cleavable bond, such as a disulfide bond or a photocleavable bond. Examples of photocleavable bonds are discussed in detail herein. In some cases, such bonds may be cleaved, for example, upon exposure to reducing agents or light (e.g., ultraviolet light). See below for additional details. Other examples of systems and methods for inactivating and/or removing the signaling entity are discussed in more detail herein.

In certain embodiments, the use of primary and secondary amplifier nucleic acids suggests that there is a maximum number of signaling entities that can be bound to a given nucleic acid probe. For instance, there may be a maximum number of primary amplifier nucleic acids is able to bind to a nucleic acid probe, e.g., due to a maximum number of secondary amplifier nucleic acids that are able to bind to a finite number of primary amplifier nucleic acids, and/or due to a maximum number of primary amplifier nucleic acids that are able to bind to the finite number of read sequences on the nucleic acid probes. While each potential location need not actually be filled with a signaling entity, this structure suggests that there is a saturation limit of signaling entities, beyond which any additional signaling entities that may happen to be present are unable to associate with a nucleic acid probe or its target.

Accordingly, certain embodiments of the invention are generally directed to systems and methods of amplifying a signal indicating a nucleic acid probe or its target that are saturatable, i.e., such that there is an upper, saturation limit of how many signaling entities can associate with the nucleic acid probe or its target. Typically, that number is greater than 1. For instance, the upper limit of signaling entities may be at least 2, at least 3, at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, at least 75, at least 100, at least 125, at least 150, at least 175, at least 200, at least 250, at least 300, at least 400, at least 500, etc. In some cases, the upper limit may be less than 500, less than 400, less than 300, less than 250, less than 200, less than 175, less than 150, less than 125, less than 100, less than 75, less than 50, less than 40, less than 30, less than 25, less than 20, less than 15, less than 10, less than 5, etc. In some cases, the upper limit may be determined as the

maximum number of signaling entities that can bind to a secondary amplifier nucleic acid, multiplied by the maximum number of secondary amplifier nucleic acids that can bind to a primary amplifier nucleic acid, multiplied by the maximum number of primary amplifier nucleic acids that can bind to a nucleic acid probe that binds to a target. In contrast, techniques such as rolling circle amplification or hairpin unfolding allow the amplification of a signal in an uncontrolled manner, i.e., when sufficient reagents are present, amplification can continue without a predetermined endpoint or saturation limit. Thus, such techniques have no theoretical upper limit as to the number of signaling entities that can associate with the nucleic acid probe or its target.

It should be understood, however, that the average number of signaling entities actually bound to a nucleic acid probe or its target need not actually be the same as its upper limit, i.e., the signaling entities may not actually be at full saturation (although they can be). For instance, the amount of saturation (or the number of signaling entities bound, relative to the maximum number that can bind) may be less than 97%, less than 95%, less than 90%, less than 85%, less than 80%, less than 75%, etc., and/or at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, etc. In some cases, allowing more time for binding to occur and/or increasing the concentration of reagents may increase the amount of saturation.

Because of the potential upper limit on the number of signaling entities actually bound to a nucleic acid probe or its target, the binding events distributed within a sample, e.g., spatially, may present substantially uniform sizes and/or brightnesses, in contrast to uncontrolled amplifications, such as those discussed above. For instance, due to the specific number of secondary amplifier nucleic acids that can bind to a primary amplifier nucleic acids, the secondary amplifier nucleic acids cannot be found greater than a fixed distance from the nucleic acid probe or its target, which may limit the "spot size" or diameter of fluorescence from the signaling entities, indicating binding.

In certain embodiments, at least 50%, at least 60%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, or at least 95% of the binding events may exhibit substantially the same brightnesses, sizes (e.g., apparent diameters), colors, or the like, which may make it easier to distinguish binding events from other events, such as nonspecific binding, noise, or the like.

In addition, as previously discussed, certain aspects of the invention use code spaces that encode the various binding events, and optionally can use error detection and/or correction to determine the binding of nucleic acid probes to their targets. In some cases, a

population of nucleic acid probes may contain certain "read sequences" which can bind certain amplifier nucleic acids, as discussed above, and the locations of the nucleic acid probes or targets can be determined within the sample using signaling entities associated with the amplifier nucleic acids, for example, within a certain code space, e.g., as discussed herein. See also Int. Pat. Apl. Pub. Nos. WO 2016/018960 and WO 2016/018963, each incorporated herein by reference in its entirety. As mentioned, in some cases, a population of read sequences within the nucleic acid probes may be combined in various combinations, e.g., such that a relatively small number of read sequences may be used to determine a relatively large number of different nucleic acid probes, as discussed herein.

Thus, in some cases, a population of nucleic acid probes may each contain a certain number of read sequences, some of which are shared between different nucleic acid probes such that the total population of nucleic acid probes may contain a certain number of read sequences. A population of nucleic acid probes may have any suitable number of read sequences. For example, a population of nucleic acid probes may have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, etc. read sequences. More than 20 are also possible in some embodiments. In addition, in some cases, a population of nucleic acid probes may, in total, have 1 or more, 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more, 11 or more, 12 or more, 13 or more, 14 or more, 15 or more, 16 or more, 20 or more, 24 or more, 32 or more, 40 or more, 50 or more, 60 or more, 64 or more, 100 or more, 128 or more, etc. of possible read sequences present, although some or all of the probes may each contain more than one read sequence, as discussed herein. In addition, in some embodiments, the population of nucleic acid probes may have no more than 100, no more than 80, no more than 64, no more than 60, no more than 50, no more than 40, no more than 32, no more than 24, no more than 20, no more than 16, no more than 15, no more than 14, no more than 13, no more than 12, no more than 11, no more than 10, no more than 9, no more than 8, no more than 7, no more than 6, no more than 5, no more than 4, no more than 3, or no more than two read sequences present. Combinations of any of these are also possible, e.g., a population of nucleic acid probes may comprise between 10 and 15 read sequences in total.

As a non-limiting example of an approach to combinatorially identifying a relatively large number of nucleic acid probes from a relatively small number of read sequences contained within the nucleic acid probes, in a population of 6 different types of nucleic acid probes, each comprising one or more read sequences, the total number of read sequences within the population may be no greater than 4. It should be understood that although 4 read

sequences are used in this example for ease of explanation, in other embodiments, larger numbers of nucleic acid probes may be realized, for example, using 5, 8, 10, 16, 32, etc. or more read sequences, or any other suitable number of read sequences described herein, depending on the application. For example, if each of the nucleic acid probes contains two different read sequences, then by using 4 such read sequences (A, B, C, and D), up to 6 probes may be separately identified. It should be noted that in this example, the ordering of read sequences on a nucleic acid probe is not essential, i.e., "AB" and "BA" may be treated as being synonymous (although in other embodiments, the ordering of read sequences may be essential and "AB" and "BA" may not necessarily be synonymous). Similarly, if 5 read sequences are used (A, B, C, D, and E) in the population of nucleic acid probes, up to 10 probes may be separately identified (e.g., AB, AC, AD, AE, BC, BD, BE, CD, CE, DE). For example, one of ordinary skill in the art would understand that, for k read sequences in a population with n read sequences on each probe, up to $\binom{n}{k}$ different probes may be produced, assuming that the ordering of read sequences is not essential; because not all of the probes need to have the same number of read sequences and not all combinations of read sequences need to be used in every embodiment, either more or less than this number of different probes may also be used in certain embodiments. In addition, it should also be understood that the number of read sequences on each probe need not be identical in some embodiments. For instance example, some probes may contain 2 read sequences while other probes may contain 3 read sequences.

In some aspects, the read sequences and/or the pattern of binding of nucleic acid probes within a sample may be used to define an error-detecting and/or an error-correcting code, for example, to reduce or prevent misidentification or errors of the nucleic acids. Thus, for example, if binding is indicated (e.g., as determined using a signaling entity), then the location may be identified with a "1"; conversely, if no binding is indicated, then the location may be identified with a "0" (or vice versa, in some cases). Multiple rounds of binding determinations, e.g., using different nucleic acid probes, can then be used to create a "codeword," e.g., for that spatial location. In some embodiments, the codeword may be subjected to error detection and/or correction. For instance, the codewords may be organized such that, if no match is found for a given set of read sequences or binding pattern of nucleic acid probes, then the match may be identified as an error, and optionally, error correction may be applied sequences to determine the correct target for the nucleic acid probes. In some

cases, the codewords may have fewer "letters" or positions that the total number of nucleic acids encoded by the codewords, e.g. where each codeword encodes a different nucleic acid.

Such error-detecting and/or the error-correction code may take a variety of forms. A variety of such codes have previously been developed in other contexts such as the telecommunications industry, such as Golay codes or Hamming codes. In one set of embodiments, the read sequences or binding patterns of the nucleic acid probes are assigned such that not every possible combination is assigned.

For example, if 4 read sequences are possible and a nucleic acid probe contains 2 read sequences, then up to 6 nucleic acid probes could be identified; but the number of nucleic acid probes used may be less than 6. Similarly, for k read sequences in a population with n read sequences on each nucleic acid probe, $\binom{n}{k}$ different probes may be produced, but the number of nucleic acid probes that are used may be any number more or less than $\binom{n}{k}$. In addition, these may be randomly assigned, or assigned in specific ways to increase the ability to detect and/or correct errors.

As another example, if multiple rounds of nucleic acid probes are used, the number of rounds may be arbitrarily chosen. If in each round, each target can give two possible outcomes, such as being detected or not being detected, up to 2^n different targets may be possible for n rounds of probes, but the number of targets that are actually used may be any number less than 2^n. For example, if in each round, each target can give more than two possible outcomes, such as being detected in different color channels, more than 2^n (e.g. 3^n, 4^n, …) different targets may be possible for n rounds of probes. In some cases, the number of targets that are actually used may be any number less than this number. In addition, these may be randomly assigned, or assigned in specific ways to increase the ability to detect and/or correct errors.

The codewords may be used to define various code spaces. For example, in one set of embodiments, the codewords or nucleic acid probes may be assigned within a code space such that the assignments are separated by a Hamming distance, which measures the number of incorrect "reads" in a given pattern that cause the nucleic acid probe to be misinterpreted as a different valid nucleic acid probe. In certain cases, the Hamming distance may be at least 2, at least 3, at least 4, at least 5, at least 6, or the like. In addition, in one set of embodiments, the assignments may be formed as a Hamming code, for instance, a Hamming(7, 4) code, a Hamming(15, 11) code, a Hamming(31, 26) code, a Hamming(63, 57) code, a Hamming(127, 120) code, etc. In another set of embodiments, the assignments

may form a SECDED code, e.g., a SECDED(8,4) code, a SECDED(16,4) code, a SCEDED(16, 11) code, a SCEDED(22, 16) code, a SCEDED(39, 32) code, a SCEDED(72, 64) code, etc. In yet another set of embodiments, the assignments may form an extended binary Golay code, a perfect binary Golay code, or a ternary Golay code. In another set of embodiments, the assignments may represent a subset of the possible values taken from any of the codes described above.

For example, an error-correcting code may be formed by using only binary words that contain a fixed or constant number of "1" bits (or "0" bits) to encode the targets. For example, the code space may only include 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, etc. "1" bits (or "0" bits), e.g., all of the codes have the same number of "1" bits or "0" bits, etc. In another set of embodiments, the assignments may represent a subset of the possible values taken from codes described above for the purpose of addressing asymmetric readout errors. For example, in some cases, a code in which the number of "1" bits may be fixed for all used binary words may eliminate the biased measurement of words with different numbers of "1"s when the rate at which "0" bits are measured as "1"s or "1" bits are measured as "0"s are different.

Accordingly, in some embodiments, once the codeword is determined (e.g., as discussed herein), the codeword may be compared to the known nucleic acid codewords. If a match is found, then the nucleic acid target can be identified or determined. If no match is found, then an error in the reading of the codeword may be identified. In some cases, error correction can also be applied to determine the correct codeword, and thus resulting in the correct identity of the nucleic acid target. In some cases, the codewords may be selected such that, assuming that there is only one error present, only one possible correct codeword is available, and thus, only one correct identity of the nucleic acid target is possible. In some cases, this may also be generalized to larger codeword spacings or Hamming distances; for instance, the codewords may be selected such that if two, three, or four errors are present (or more in some cases), only one possible correct codeword is available, and thus, only one correct identity of the nucleic acid targets is possible.

The error-correcting code may be a binary error-correcting code, or it may be based on other numbering systems, e.g., ternary or quaternary error-correcting codes. For instance, in one set of embodiments, more than one type of signaling entity may be used and assigned to different numbers within the error-correcting code. Thus, as a non-limiting example, a first signaling entity (or more than one signaling entity, in some cases) may be assigned as "1" and a second signaling entity (or more than one signaling entity, in some cases) may be

assigned as "2" (with "0" indicating no signaling entity present), and the codewords distributed to define a ternary error-correcting code. Similarly, a third signaling entity may additionally be assigned as "3" to make a quaternary error-correcting code, etc.

As discussed herein, in certain aspects, signaling entities are determined, e.g., by imaging, to determine nucleic acid probes and/or to create codewords. Examples of signaling entities include those discussed herein. In some cases, signaling entities within a sample may be determined, e.g., spatially, using a variety of techniques. In some embodiments, the signaling entities may be fluorescent, and techniques for determining fluorescence within a sample, such as fluorescence microscopy or confocal microscopy, may be used to spatially identify the positions of signaling entities within a cell. In some cases, the positions of entities within the sample may be determined in two or even three dimensions. In addition, in some embodiments, more than one signaling entity may be determined at a time (e.g., signaling entities with different colors or emissions), and/or sequentially.

In addition, in some embodiments, a confidence level for an identified target, e.g., a nucleic acid target, may be determined. For example, the confidence level may be determined using a ratio of the number of exact matches to the number of matches having one or more one-bit errors. In some cases, only matches having a confidence ratio greater than a certain value may be used. For instance, in certain embodiments, matches may be accepted only if the confidence ratio for the match is greater than about 0.01, greater than about 0.03, greater than about 0.05, greater than about 0.1, greater than about 0.3, greater than about 0.5, greater than about 1, greater than about 3, greater than about 5, greater than about 10, greater than about 30, greater than about 50, greater than about 100, greater than about 300, greater than about 500, greater than about 1000, or any other suitable value. In addition, in some embodiments, matches may be accepted only if the confidence ratio for the identified target is greater than an internal standard or false positive control by about 0.01, about 0.03, about 0.05, about 0.1, about 0.3, about 0.5, about 1, about 3, about 5, about 10, about 30, about 50, about 100, about 300, about 500, about 1000, or any other suitable value

In some embodiments, the spatial positions of the entities (and thus, nucleic acid probes that the entities may be associated with) may be determined at relatively high resolutions. For instance, the positions may be determined at spatial resolutions of better than about 100 micrometers, better than about 30 micrometers, better than about 10 micrometers, better than about 3 micrometers, better than about 1 micrometer, better than about 800 nm, better than about 600 nm, better than about 500 nm, better than about 400 nm, better than about 300 nm, better than about 200 nm, better than about 100 nm, better than about 90 nm,

better than about 80 nm, better than about 70 nm, better than about 60 nm, better than about 50 nm, better than about 40 nm, better than about 30 nm, better than about 20 nm, or better than about 10 nm, etc.

There are a variety of techniques able to determine or image the spatial positions of entities optically, e.g., using fluorescence microscopy. More than one color can be used in some embodiments. In some cases, the spatial positions may be determined at super resolutions, or at resolutions better than the wavelength of light or the diffraction limit. Non-limiting examples include STORM (stochastic optical reconstruction microscopy), STED (stimulated emission depletion microscopy), NSOM (Near-field Scanning Optical Microscopy), 4Pi microscopy, SIM (Structured Illumination Microscopy), SMI (Spatially Modulated Illumination) microscopy, RESOLFT (Reversible Saturable Optically Linear Fluorescence Transition Microscopy), GSD (Ground State Depletion Microscopy), SSIM (Saturated Structured-Illumination Microscopy), SPDM (Spectral Precision Distance Microscopy), Photo-Activated Localization Microscopy (PALM), Fluorescence Photoactivation Localization Microscopy (FPALM), LIMON (3D Light Microscopical Nanosizing Microscopy), Super-resolution optical fluctuation imaging (SOFI), or the like. See, e.g., U.S. Pat. No. 7,838,302, issued November 23, 2010, entitled "Sub-Diffraction Limit Image Resolution and Other Imaging Techniques," by Zhuang, *et al.*; U.S. Pat. No. 8,564,792, issued October 22, 2013, entitled "Sub-diffraction Limit Image Resolution in Three Dimensions," by Zhuang, *et al.*; or Int. Pat. Apl. Pub. No. WO 2013/090360, published June 20, 2013, entitled "High Resolution Dual-Objective Microscopy," by Zhuang, *et al.*, each incorporated herein by reference in their entireties.

As an illustrative non-limiting example, in one set of embodiments, the sample may be imaged with a high numerical aperture, oil immersion objective with 100X magnification and light collected on an electron-multiplying CCD camera. In another example, the sample could be imaged with a high numerical aperture, oil immersion lens with 40X magnification and light collected with a wide-field scientific CMOS camera. With different combinations of objectives and cameras, a single field of view may correspond to no less than 40 x 40 microns, 80 x 80 microns, 120 x 120 microns, 240 x 240 microns, 340 x 340 microns, or 500 x 500 microns, etc. in various non-limiting embodiments. Similarly, a single camera pixel may correspond, in some embodiments, to regions of the sample of no less than 80x80 nm, 120x120 nm, 160x160 nm, 240x240 nm, or 300x300 nm, etc. In another example, the sample may be imaged with a low numerical aperture, air lens with 10X magnification and light collected with a sCMOS camera. In additional embodiments, the sample may be

optically sectioned by illuminating it via a single or multiple scanned diffraction limited foci generated either by scanning mirrors or a spinning disk and the collected passed through a single or multiple pinholes. In another embodiment, the sample may also be illuminated via thin sheet of light generated via any one of multiple methods known to those versed in the art.

In one embodiment, the sample may be illuminated by single Gaussian mode laser lines. In some embodiments, the illumination profiled may be flattened by passing these laser lines through a multimode fiber that is vibrated via piezo-electric or other mechanical means. In some embodiments, the illumination profile may be flattened by passing single-mode, Gaussian beams through a variety of refractive beam shapers, such as the piShaper or a series of stacked Powell lenses. In yet another set of embodiments, the Gaussian beams may be passed through a variety of different diffusing elements, such as ground glass or engineered diffusers, which may be spun in some cases at high speeds to remove residual laser speckle. In yet another embodiment, laser illumination may be passed through a series of lenslet arrays to produce overlapping images of the illumination that approximate a flat illumination field.

In some embodiments, the centroids of the spatial positions of the entities may be determined. For example, a centroid of a signaling entity may be determined within an image or series of images using image analysis algorithms known to those of ordinary skill in the art. In some cases, the algorithms may be selected to determine non-overlapping single emitters and/or partially overlapping single emitters in a sample. Non-limiting examples of suitable techniques include a maximum likelihood algorithm, a least squares algorithm, a Bayesian algorithm, a compressed sensing algorithm, or the like. Combinations of these techniques may also be used in some cases.

In addition, the signaling entity may be inactivated in some cases. For example, in some embodiments, a first secondary nucleic acid probe that can associate with a signaling entity (e.g., using amplifier nucleic acids) may be applied to a sample that can recognize a first read sequence (e.g., on the nucleic acid probe), then the signaling entity can be inactivated before a second secondary nucleic acid probe is applied to the sample, e.g., that can associate with a signaling entity (e.g., using amplifier nucleic acids). If multiple signaling entities are used, the same or different techniques may be used to inactivate the signaling entities, and some or all of the multiple signaling entities may be inactivated, e.g., sequentially or simultaneously.

Inactivation may be caused by removal of the signaling entity (e.g., from the sample, or from the nucleic acid probe, etc.), and/or by chemically altering the signaling entity in

some fashion (e.g., by photobleaching the signaling entity, bleaching or chemically altering the structure of the signaling entity, for example, by reduction, etc.). For instance, in one set of embodiments, a fluorescent signaling entity may be inactivated by chemical or optical techniques such as oxidation, photobleaching, chemically bleaching, stringent washing or enzymatic digestion or reaction by exposure to an enzyme, dissociating the signaling entity from other components (e.g., a probe), chemical reaction of the signaling entity (e.g., to a reactant able to alter the structure of the signaling entity) or the like. For instance, bleaching may occur by exposure to oxygen, reducing agents, or the signaling entity could be chemically cleaved from the nucleic acid probe and washed away via fluid flow.

In some embodiments, various nucleic acid probes may be associated with one or more signaling entities, e.g., using amplifier nucleic acids as discussed herein. If more than one nucleic acid probe is used, the signaling entities may each by the same or different. In certain embodiments, a signaling entity is any entity able to emit light. For instance, in one embodiment, the signaling entity is fluorescent. In other embodiments, the signaling entity may be phosphorescent, radioactive, absorptive, etc. In some cases, the signaling entity is any entity that can be determined within a sample at relatively high resolutions, e.g., at resolutions better than the wavelength of visible light or the diffraction limit. The signaling entity may be, for example, a dye, a small molecule, a peptide or protein, or the like. The signaling entity may be a single molecule in some cases. If multiple secondary nucleic acid probes are used, the nucleic acid probes may associate with the same or different signaling entities.

Non-limiting examples of signaling entities include fluorescent entities (fluorophores) or phosphorescent entities, for example, cyanine dyes (e.g., Cy2, Cy3, Cy3B, Cy5, Cy5.5, Cy7, etc.), Alexa Fluor dyes, Atto dyes, photoswitchable dyes, photoactivatable dyes, fluorescent dyes, metal nanoparticles, semiconductor nanoparticles or "quantum dots," fluorescent proteins such as GFP (Green Fluorescent Protein), or photoactivabale fluorescent proteins, such as PAGFP, PSCFP, PSCFP2, Dendra, Dendra2, EosFP, tdEos, mEos2, mEos3, PAmCherry, PAtagRFP, mMaple, mMaple2, and mMaple3. Other suitable signaling entities are known to those of ordinary skill in the art. See, e.g., U.S. Pat. No. 7,838,302 or Int. Pat Apl. Pub. No. WO 2015/160690, each incorporated herein by reference in its entirety.

In one set of embodiments, the signaling entity may be attached to an oligonucleotide sequence via a bond that can be cleaved to release the signaling entity. In one set of embodiments, a fluorophore may be conjugated to an oligonucleotide via a cleavable bond, such as a photocleavable bond. Non-limiting examples of photocleavable bonds include, but

are not limited to, 1-(2-nitrophenyl)ethyl, 2-nitrobenzyl, biotin phosphoramidite, acrylic phosphoramidite, diethylaminocoumarin, 1-(4,5-dimethoxy-2-nitrophenyl)ethyl, cyclododecyl (dimethoxy-2-nitrophenyl)ethyl, 4-aminomethyl-3-nitrobenzyl, (4-nitro-3-(1-chlorocarbonyloxyethyl)phenyl)methyl-S-acetylthioic acid ester, (4-nitro-3-(1-thlorocarbonyloxyethyl)phenyl)methyl-3-(2-pyridyldithiopropionic acid) ester, 3-(4,4'-dimethoxytrityl)-1-(2-nitrophenyl)-propane-1,3-diol-[2- cyanoethyl-(N,N-diisopropyl)]-phosphoramidite, 1-[2-nitro-5-(6-trifluoroacetylcaproamidomethyl)phenyl]-ethyl-[2-cyanoethyl-(N,N-diisopropyl)]-phosphoramidite, 1-[2-nitro-5-(6-(4,4'-dimethoxytrityloxy)butyramidomethyl)phenyl]-ethyl-[2-cyanoethyl-(N,N-diisopropyl)]-phosphoramidite, 1-[2-nitro-5-(6-(N-(4,4'-dimethoxytrityl))-biotinamidocaproamidomethyl)phenyl]-ethyl-[2-cyanoethyl-(N,N-diisopropyl)]-phosphoramidite, or similar linkers. The oligonucleotide sequence may be, for example, a primary or secondary (or other) amplifier nucleic acid, such as those discussed herein.

In another set of embodiments, the fluorophore may be conjugated to an oligonucleotide via a disulfide bond. The disulfide bond may be cleaved by a variety of reducing agents such as, but not limited to, dithiothreitol, dithioerythritol, beta-mercaptoethanol, sodium borohydride, thioredoxin, glutaredoxin, trypsinogen, hydrazine, diisobutylaluminum hydride, oxalic acid, formic acid, ascorbic acid, phosphorous acid, tin chloride, glutathione, thioglycolate, 2,3-dimercaptopropanol, 2-mercaptoethylamine, 2-aminoethanol, tris(2-carboxyethyl)phosphine, bis(2-mercaptoethyl) sulfone, N,N'-dimethyl-N,N'-bis(mercaptoacetyl)hydrazine, 3-mercaptoproptionate, dimethylformamide, thiopropyl-agarose, tri-*n*-butylphosphine, cysteine, iron sulfate, sodium sulfite, phosphite, hypophosphite, phosphorothioate, or the like, and/or combinations of any of these. The oligonucleotide sequence may be, for example, a primary or secondary (or other) amplifier nucleic acid, such as those discussed herein.

In another embodiment, the fluorophore may be conjugated to an oligonucleotide via one or more phosphorothioate modified nucleotides in which the sulfur modification replaces the bridging and/or non-bridging oxygen. The fluorophore may be cleaved from the oligonucleotide, in certain embodiments, via addition of compounds such as but not limited to iodoethanol, iodine mixed in ethanol, silver nitrate, or mercury chloride. In yet another set of embodiments, the signaling entity may be chemically inactivated through reduction or oxidation. For example, in one embodiment, a chromophore such as Cy5 or Cy7 may be reduced using sodium borohydride to a stable, non-fluorescence state. In still another set of embodiments, a fluorophore may be conjugated to an oligonucleotide via an azo bond, and

the azo bond may be cleaved with 2-[(2-N-arylamino)phenylazo]pyridine. In yet another set of embodiments, a fluorophore may be conjugated to an oligonucleotide via a suitable nucleic acid segment that can be cleaved upon suitable exposure to DNAse, e.g., an exodeoxyribonuclease or an endodeoxyribonuclease. Examples include, but are not limited to, deoxyribonuclease I or deoxyribonuclease II. In one set of embodiments, the cleavage may occur via a restriction endonuclease. Non-limiting examples of potentially suitable restriction endonucleases include BamHI, BsrI, NotI, XmaI, PspAI, DpnI, MboI, MnlI, Eco57I, Ksp632I, DraIII, AhaII, SmaI, MluI, HpaI, ApaI, BclI, BstEII, TaqI, EcoRI, SacI, HindII, HaeII, DraII, Tsp509I, Sau3AI, PacI, etc. Over 3000 restriction enzymes have been studied in detail, and more than 600 of these are available commercially. In yet another set of embodiments, a fluorophore may be conjugated to biotin, and the oligonucleotide conjugated to avidin or streptavidin. An interaction between biotin and avidin or streptavidin allows the fluorophore to be conjugated to the oligonucleotide, while sufficient exposure to an excess of addition, free biotin could "outcompete" the linkage and thereby cause cleavage to occur. In addition, in another set of embodiments, the probes may be removed using corresponding "toe-hold-probes," which comprise the same sequence as the probe, as well as an extra number of bases of homology to the encoding probes (e.g., 1-20 extra bases, for example, 5 extra bases). These probes may remove the labeled readout probe through a strand-displacement interaction. The oligonucleotide sequence may be, for example, a primary or secondary (or other) amplifier nucleic acid, such as those discussed herein.

As used herein, the term "light" generally refers to electromagnetic radiation, having any suitable wavelength (or equivalently, frequency). For instance, in some embodiments, the light may include wavelengths in the optical or visual range (for example, having a wavelength of between about 400 nm and about 700 nm, i.e., "visible light"), infrared wavelengths (for example, having a wavelength of between about 300 micrometers and 700 nm), ultraviolet wavelengths (for example, having a wavelength of between about 400 nm and about 10 nm), or the like. In certain cases, as discussed in detail below, more than one entity may be used, i.e., entities that are chemically different or distinct, for example, structurally. However, in other cases, the entities may be chemically identical or at least substantially chemically identical.

In one set of embodiments, the signaling entity is "switchable," i.e., the entity can be switched between two or more states, at least one of which emits light having a desired wavelength. In the other state(s), the entity may emit no light, or emit light at a different wavelength. For instance, an entity may be "activated" to a first state able to produce light

having a desired wavelength, and "deactivated" to a second state not able to emit light of the same wavelength. An entity is "photoactivatable" if it can be activated by incident light of a suitable wavelength. As a non-limiting example, Cy5, can be switched between a fluorescent and a dark state in a controlled and reversible manner by light of different wavelengths, i.e., 633 nm (or 642nm, 647nm, 656 nm) red light can switch or deactivate Cy5 to a stable dark state, while 405 nm green light can switch or activate the Cy5 back to the fluorescent state. In some cases, the entity can be reversibly switched between the two or more states, e.g., upon exposure to the proper stimuli. For example, a first stimuli (e.g., a first wavelength of light) may be used to activate the switchable entity, while a second stimuli (e.g., a second wavelength of light) may be used to deactivate the switchable entity, for instance, to a non-emitting state. Any suitable method may be used to activate the entity. For example, in one embodiment, incident light of a suitable wavelength may be used to activate the entity to emit light, i.e., the entity is "photoswitchable." Thus, the photoswitchable entity can be switched between different light-emitting or non-emitting states by incident light, e.g., of different wavelengths. The light may be monochromatic (e.g., produced using a laser) or polychromatic. In another embodiment, the entity may be activated upon stimulation by electric field and/or magnetic field. In other embodiments, the entity may be activated upon exposure to a suitable chemical environment, e.g., by adjusting the pH, or inducing a reversible chemical reaction involving the entity, etc. Similarly, any suitable method may be used to deactivate the entity, and the methods of activating and deactivating the entity need not be the same. For instance, the entity may be deactivated upon exposure to incident light of a suitable wavelength, or the entity may be deactivated by waiting a sufficient time.

Typically, a "switchable" entity can be identified by one of ordinary skill in the art by determining conditions under which an entity in a first state can emit light when exposed to an excitation wavelength, switching the entity from the first state to the second state, e.g., upon exposure to light of a switching wavelength, then showing that the entity, while in the second state can no longer emit light (or emits light at a much reduced intensity) when exposed to the excitation wavelength.

In one set of embodiments, as discussed, a switchable entity may be switched upon exposure to light. In some cases, the light used to activate the switchable entity may come from an external source, e.g., a light source such as a laser light source, another light-emitting entity proximate the switchable entity, etc. The second, light emitting entity, in some cases, may be a fluorescent entity, and in certain embodiments, the second, light-emitting entity may itself also be a switchable entity.

In some embodiments, the switchable entity includes a first, light-emitting portion (e.g., a fluorophore), and a second portion that activates or "switches" the first portion. For example, upon exposure to light, the second portion of the switchable entity may activate the first portion, causing the first portion to emit light. Examples of activator portions include, but are not limited to, Alexa Fluor 405 (Invitrogen), Alexa Fluor 488 (Invitrogen), Cy2 (GE Healthcare), Cy3 (GE Healthcare), Cy3B (GE Healthcare), Cy3.5 (GE Healthcare), or other suitable dyes. Examples of light-emitting portions include, but are not limited to, Cy5, Cy5.5 (GE Healthcare), Cy7 (GE Healthcare), Alexa Fluor 647 (Invitrogen), Alexa Fluor 680 (Invitrogen), Alexa Fluor 700 (Invitrogen), Alexa Fluor 750 (Invitrogen), Alexa Fluor 790 (Invitrogen), DiD, DiR, YOYO-3 (Invitrogen), YO-PRO-3 (Invitrogen), TOT-3 (Invitrogen), TO-PRO-3 (Invitrogen) or other suitable dyes. These may linked together, e.g., covalently, for example, directly, or through a linker, e.g., forming compounds such as, but not limited to, Cy5-Alexa Fluor 405, Cy5-Alexa Fluor 488, Cy5-Cy2, Cy5-Cy3, Cy5-Cy3.5, Cy5.5-Alexa Fluor 405, Cy5.5-Alexa Fluor 488, Cy5.5-Cy2, Cy5.5-Cy3, Cy5.5-Cy3.5, Cy7-Alexa Fluor 405, Cy7-Alexa Fluor 488, Cy7-Cy2, Cy7-Cy3, Cy7-Cy3.5, Alexa Fluor 647-Alexa Fluor 405, Alexa Fluor 647-Alexa Fluor 488, Alexa Fluor 647-Cy2, Alexa Fluor 647-Cy3, Alexa Fluor 647-Cy3.5, Alexa Fluor 750-Alexa Fluor 405, Alexa Fluor 750-Alexa Fluor 488, Alexa Fluor 750-Cy2, Alexa Fluor 750-Cy3, or Alexa Fluor 750-Cy3.5. Those of ordinary skill in the art will be aware of the structures of these and other compounds, many of which are available commercially. The portions may be linked via a covalent bond, or by a linker, such as those described in detail below. Other light-emitting or activator portions may include portions having two quaternized nitrogen atoms joined by a polymethine chain, where each nitrogen is independently part of a heteroaromatic moiety, such as pyrrole, imidazole, thiazole, pyridine, quinoine, indole, benzothiazole, etc., or part of a nonaromatic amine. In some cases, there may be 5, 6, 7, 8, 9, or more carbon atoms between the two nitrogen atoms.

In certain cases, the light-emitting portion and the activator portions, when isolated from each other, may each be fluorophores, i.e., entities that can emit light of a certain, emission wavelength when exposed to a stimulus, for example, an excitation wavelength. However, when a switchable entity is formed that comprises the first fluorophore and the second fluorophore, the first fluorophore forms a first, light-emitting portion and the second fluorophore forms an activator portion that switches that activates or "switches" the first portion in response to a stimulus. For example, the switchable entity may comprise a first fluorophore directly bonded to the second fluorophore, or the first and second entity may be

connected via a linker or a common entity. Whether a pair of light-emitting portion and activator portion produces a suitable switchable entity can be tested by methods known to those of ordinary skills in the art. For example, light of various wavelength can be used to stimulate the pair and emission light from the light-emitting portion can be measured to determined wither the pair makes a suitable switch.

As a non-limiting example, Cy3 and Cy5 may be linked together to form such an entity. In this example, Cy3 is an activator portion that is able to activate Cy5, the light-emission portion. Thus, light at or near the absorption maximum (e.g., near 532 nm light for Cy3) of the activation or second portion of the entity may cause that portion to activate the first, light-emitting portion, thereby causing the first portion to emit light (e.g., near 647 nm for Cy5). See, e.g., U.S. Pat. No. 7,838,302, incorporated herein by reference in its entirety. In some cases, the first, light-emitting portion can subsequently be deactivated by any suitable technique (e.g., by directing 647 nm red light to the Cy5 portion of the molecule).

Other non-limiting examples of potentially suitable activator portions include 1,5 IAEDANS, 1,8-ANS, 4-Methylumbelliferone, 5-carboxy-2,7-dichlorofluorescein, 5-Carboxyfluorescein (5-FAM), 5-Carboxynapthofluorescein, 5-Carboxytetramethylrhodamine (5-TAMRA), 5-FAM (5-Carboxyfluorescein), 5-HAT (Hydroxy Tryptamine), 5-Hydroxy Tryptamine (HAT), 5-ROX (carboxy-X-rhodamine), 5-TAMRA (5-Carboxytetramethylrhodamine), 6-Carboxyrhodamine 6G, 6-CR 6G, 6-JOE, 7-Amino-4-methylcoumarin, 7-Aminoactinomycin D (7-AAD), 7-Hydroxy-4-methylcoumarin, 9-Amino-6-chloro-2-methoxyacridine, ABQ, Acid Fuchsin, ACMA (9-Amino-6-chloro-2-methoxyacridine), Acridine Orange, Acridine Red, Acridine Yellow, Acriflavin, Acriflavin Feulgen SITSA, Alexa Fluor 350, Alexa Fluor 405, Alexa Fluor 430, Alexa Fluor 488, Alexa Fluor 500, Alexa Fluor 514, Alexa Fluor 532, Alexa Fluor 546, Alexa Fluor 555, Alexa Fluor 568, Alexa Fluor 594, Alexa Fluor 610, Alexa Fluor 633, Alexa Fluor 635, Alizarin Complexon, Alizarin Red, AMC, AMCA-S, AMCA (Aminomethylcoumarin), AMCA-X, Aminoactinomycin D, Aminocoumarin, Aminomethylcoumarin (AMCA), Anilin Blue, Anthrocyl stearate, APTRA-BTC, APTS, Astrazon Brilliant Red 4G, Astrazon Orange R, Astrazon Red 6B, Astrazon Yellow 7 GLL, Atabrine, ATTO 390, ATTO 425, ATTO 465, ATTO 488, ATTO 495, ATTO 520, ATTO 532, ATTO 550, ATTO 565, ATTO 590, ATTO 594, ATTO 610, ATTO 611X, ATTO 620, ATTO 633, ATTO 635, ATTO 647, ATTO 647N, ATTO 655, ATTO 680, ATTO 700, ATTO 725, ATTO 740, ATTO-TAG CBQCA, ATTO-TAG FQ, Auramine, Aurophosphine G, Aurophosphine, BAO 9 (Bisaminophenyloxadiazole), BCECF (high pH), BCECF (low pH), Berberine Sulphate,

Bimane, Bisbenzamide, Bisbenzimide (Hoechst), bis-BTC, Blancophor FFG, Blancophor SV, BOBO -1, BOBO -3, Bodipy 492/515, Bodipy 493/503, Bodipy 500/510, Bodipy 505/515, Bodipy 530/550, Bodipy 542/563, Bodipy 558/568, Bodipy 564/570, Bodipy 576/589, Bodipy 581/591, Bodipy 630/650-X, Bodipy 650/665-X, Bodipy 665/676, Bodipy Fl, Bodipy FL ATP, Bodipy Fl-Ceramide, Bodipy R6G, Bodipy TMR, Bodipy TMR-X conjugate, Bodipy TMR-X, SE, Bodipy TR, Bodipy TR ATP, Bodipy TR-X SE, BO-PRO -1, BO-PRO -3, Brilliant Sulphoflavin FF, BTC, BTC-5N, Calcein, Calcein Blue, Calcium Crimson, Calcium Green, Calcium Green-1 Ca^{2+} Dye, Calcium Green-2 Ca^{2+}, Calcium Green-5N Ca^{2+}, Calcium Green-C18 Ca^{2+}, Calcium Orange, Calcofluor White, Carboxy-X-rhodamine (5-ROX), Cascade Blue, Cascade Yellow, Catecholamine, CCF2 (GeneBlazer), CFDA, Chromomycin A, Chromomycin A, CL-NERF, CMFDA, Coumarin Phalloidin, CPM Methylcoumarin, CTC, CTC Formazan, Cy2, Cy3.1 8, Cy3.5, Cy3, Cy5.1 8, cyclic AMP Fluorosensor (FiCRhR), Dabcyl, Dansyl, Dansyl Amine, Dansyl Cadaverine, Dansyl Chloride, Dansyl DHPE, Dansyl fluoride, DAPI, Dapoxyl, Dapoxyl 2, Dapoxyl 3' DCFDA, DCFH (Dichlorodihydrofluorescein Diacetate), DDAO, DHR (Dihydorhodamine 123), Di-4-ANEPPS, Di-8-ANEPPS (non-ratio), DiA (4-Di-16-ASP), Dichlorodihydrofluorescein Diacetate (DCFH), DiD - Lipophilic Tracer, DiD (DiIC18(5)), DIDS, Dihydorhodamine 123 (DHR), DiI (DiIC18(3)), Dinitrophenol, DiO (DiOC18(3)), DiR, DiR (DiIC18(7)), DM-NERF (high pH), DNP, Dopamine, DTAF, DY-630-NHS, DY-635-NHS, DyLight 405, DyLight 488, DyLight 549, DyLight 633, DyLight 649, DyLight 680, DyLight 800, ELF 97, Eosin, Erythrosin, Erythrosin ITC, Ethidium Bromide, Ethidium homodimer -1 (EthD-1), Euchrysin, EukoLight, Europium (III) chloride, Fast Blue, FDA, Feulgen (Pararosaniline), FIF (Formaldehyd Induced Fluorescence), FITC, Flazo Orange, Fluo-3, Fluo-4, Fluorescein (FITC), Fluorescein Diacetate, Fluoro-Emerald, Fluoro-Gold (Hydroxystilbamidine), Fluor-Ruby, FluorX, FM 1-43, FM 4-46, Fura Red (high pH), Fura Red/Fluo-3, Fura-2, Fura-2/BCECF, Genacryl Brilliant Red B, Genacryl Brilliant Yellow 10GF, Genacryl Pink 3G, Genacryl Yellow 5GF, GeneBlazer (CCF2), Gloxalic Acid, Granular blue, Haematoporphyrin, Hoechst 33258, Hoechst 33342, Hoechst 34580, HPTS, Hydroxycoumarin, Hydroxystilbamidine (FluoroGold), Hydroxytryptamine, Indo-1, high calcium, Indo-1, low calcium, Indodicarbocyanine (DiD), Indotricarbocyanine (DiR), Intrawhite Cf, JC-1, JO-JO-1, JO-PRO-1, LaserPro, Laurodan, LDS 751 (DNA), LDS 751 (RNA), Leucophor PAF, Leucophor SF, Leucophor WS, Lissamine Rhodamine, Lissamine Rhodamine B, Calcein/Ethidium homodimer, LOLO-1, LO-PRO-1, Lucifer Yellow, Lyso Tracker Blue, Lyso Tracker Blue-White, Lyso Tracker Green, Lyso Tracker Red, Lyso

Tracker Yellow, LysoSensor Blue, LysoSensor Green, LysoSensor Yellow/Blue, Mag Green, Magdala Red (Phloxin B), Mag-Fura Red, Mag-Fura-2, Mag-Fura-5, Mag-Indo-1, Magnesium Green, Magnesium Orange, Malachite Green, Marina Blue, Maxilon Brilliant Flavin 10 GFF, Maxilon Brilliant Flavin 8 GFF, Merocyanin, Methoxycoumarin, Mitotracker Green FM, Mitotracker Orange, Mitotracker Red, Mitramycin, Monobromobimane, Monobromobimane (mBBr-GSH), Monochlorobimane, MPS (Methyl Green Pyronine Stilbene), NBD, NBD Amine, Nile Red, Nitrobenzoxadidole, Noradrenaline, Nuclear Fast Red, Nuclear Yellow, Nylosan Brilliant Iavin E8G, Oregon Green, Oregon Green 488-X, Oregon Green, Oregon Green 488, Oregon Green 500, Oregon Green 514, Pacific Blue, Pararosaniline (Feulgen), PBFI, Phloxin B (Magdala Red), Phorwite AR, Phorwite BKL, Phorwite Rev, Phorwite RPA, Phosphine 3R, PKH26 (Sigma), PKH67, PMIA, Pontochrome Blue Black, POPO-1, POPO-3, PO-PRO-1, PO-PRO-3, Primuline, Procion Yellow, Propidium Iodid (PI), PyMPO, Pyrene, Pyronine, Pyronine B, Pyrozal Brilliant Flavin 7GF, QSY 7, Quinacrine Mustard, Resorufin, RH 414, Rhod-2, Rhodamine, Rhodamine 110, Rhodamine 123, Rhodamine 5 GLD, Rhodamine 6G, Rhodamine B, Rhodamine B 200, Rhodamine B extra, Rhodamine BB, Rhodamine BG, Rhodamine Green, Rhodamine Phallicidine, Rhodamine Phalloidine, Rhodamine Red, Rhodamine WT, Rose Bengal, S65A, S65C, S65L, S65T, SBFI, Serotonin, Sevron Brilliant Red 2B, Sevron Brilliant Red 4G, Sevron Brilliant Red B, Sevron Orange, Sevron Yellow L, SITS, SITS (Primuline), SITS (Stilbene Isothiosulphonic Acid), SNAFL calcein, SNAFL-1, SNAFL-2, SNARF calcein, SNARF1, Sodium Green, SpectrumAqua, SpectrumGreen, SpectrumOrange, Spectrum Red, SPQ (6-methoxy-N-(3-sulfopropyl)quinolinium), Stilbene, Sulphorhodamine B can C, Sulphorhodamine Extra, SYTO 11, SYTO 12, SYTO 13, SYTO 14, SYTO 15, SYTO 16, SYTO 17, SYTO 18, SYTO 20, SYTO 21, SYTO 22, SYTO 23, SYTO 24, SYTO 25, SYTO 40, SYTO 41, SYTO 42, SYTO 43, SYTO 44, SYTO 45, SYTO 59, SYTO 60, SYTO 61, SYTO 62, SYTO 63, SYTO 64, SYTO 80, SYTO 81, SYTO 82, SYTO 83, SYTO 84, SYTO 85, SYTOX Blue, SYTOX Green, SYTOX Orange, Tetracycline, Tetramethylrhodamine (TAMRA), Texas Red, Texas Red-X conjugate, Thiadicarbocyanine (DiSC3), Thiazine Red R, Thiazole Orange, Thioflavin 5, Thioflavin S, Thioflavin TCN, Thiolyte, Thiozole Orange, Tinopol CBS (Calcofluor White), TMR, TO-PRO-1, TO-PRO-3, TO-PRO-5, TOTO-1, TOTO-3, TRITC (tetramethylrodamine isothiocyanate), True Blue, TruRed, Ultralite, Uranine B, Uvitex SFC, WW 781, X-Rhodamine, XRITC, Xylene Orange, Y66F, Y66H, Y66W, YO-PRO-1, YO-PRO-3, YOYO-1, YOYO-3, SYBR Green, Thiazole orange (interchelating dyes), or combinations thereof.

Another aspect of the invention is directed to a computer-implemented method. For instance, a computer and/or an automated system may be provided that is able to automatically and/or repetitively perform any of the methods described herein. As used herein, "automated" devices refer to devices that are able to operate without human direction, i.e., an automated device can perform a function during a period of time after any human has finished taking any action to promote the function, e.g. by entering instructions into a computer to start the process. Typically, automated equipment can perform repetitive functions after this point in time. The processing steps may also be recorded onto a machine-readable medium in some cases.

For example, in some cases, a computer may be used to control imaging of the sample, e.g., using fluorescence microscopy, STORM or other super-resolution techniques such as those described herein. In some cases, the computer may also control operations such as drift correction, physical registration, hybridization and cluster alignment in image analysis, cluster decoding (e.g., fluorescent cluster decoding), error detection or correction (e.g., as discussed herein), noise reduction, identification of foreground features from background features (such as noise or debris in images), or the like. As an example, the computer may be used to control activation and/or excitation of signaling entities within the sample, and/or the acquisition of images of the signaling entities. In one set of embodiments, a sample may be excited using light having various wavelengths and/or intensities, and the sequence of the wavelengths of light used to excite the sample may be correlated, using a computer, to the images acquired of the sample containing the signaling entities. For instance, the computer may apply light having various wavelengths and/or intensities to a sample to yield different average numbers of signaling entities in each region of interest (e.g., one activated entity per location, two activated entities per location, etc.). In some cases, this information may be used to construct an image and/or determine the locations of the signaling entities, in some cases at high resolutions, as noted above.

In some aspects, the sample is positioned on a microscope. In some cases, the microscope may contain one or more channels, such as microfluidic channels, to direct or control fluid to or from the sample. For instance, in one embodiment, nucleic acid probes such as those discussed herein may be introduced and/or removed from the sample by flowing fluid through one or more channels to or from the sample. In some cases, there may also be one or more chambers or reservoirs for holding fluid, e.g., in fluidic communication with the channel, and/or with the sample. Those of ordinary skill in the art will be familiar with channels, including microfluidic channels, for moving fluid to or from a sample.

U.S. Provisional Patent Application Serial No. 62/779,333, filed December 13, 2018, entitled "Amplification Methods and Systems for MERFISH and Other Applications," by Zhuang, *et al.*, is incorporated herein by reference in its entirety. The following documents are also incorporated herein by reference in their entireties: U.S. Pat. Nos. 2017/0220733 and 2017/0212986; Int. Pat. Apl. Pub. Nos. WO 2016/018960, WO 2016/018963, WO 2018/089445, and WO 2018/089438; and Int. Pat. Apl. Ser. No. PCT/US18/34651.

The following examples are intended to illustrate certain embodiments of the present invention, but do not exemplify the full scope of the invention.

EXAMPLE 1

Multiplexed error-robust fluorescence *in situ* hybridization (MERFISH) directly images and profiles hundreds to thousands of RNA species in their native cellular environment, thus allowing spatially resolved single-cell transcriptomics. The following examples use a controlled or saturatable amplification approach for MERFISH measurements. These examples demonstrate a substantial signal increase in RNA FISH samples, without increasing spot sizes or increasing the variation in brightness from spot to spot. 130 RNA species were accurately measured and identified with a near 100% detection efficiency. These advances should substantially extend the range of biological questions that can be addressed by MERFISH in both cell culture and tissues.

In particular, these examples show an approach that combines MERFISH amplification to dramatically increase the MERFISH signal brightness. These examples demonstrate that this approach amplified RNA smFISH signals up to 30.6x without spot size change in a constant coefficient-of-variation manner. This approach provided a fast, simple, and efficient way to simultaneously amplify the signal of hundreds RNAs in same single cells with a nearly 100% detection efficiency.

Design of incorporating DNA amplification into MERFISH. These examples show that this amplification strategy has certain advantages when combined with MERFISH, using certain modifications as described herein. The scaffolds that determine the amplification factor were found to be predictable and designable. Even if multiple amplifiers had different binding rates, the saturation states were reached with sufficient time for all amplifiers tested, making the amplification controllable, reproducible and with low variance, benefitting MERFISH decoding process.

This labeling approach used in these examples is illustrated in FIGs. 1A-1E. A complex library of oligonucleotide probes that had a 30-mer target region sequence complimentary to the target RNA templates and multiple 20-mer readout sequences,

termed encoding probes, were stained using MERFISH systems (see, e.g., Int. Pat. Apl. Pub. Nos. WO 2016/018960 and WO 2016/018963, each incorporated herein by reference in its entirety).

Amplifier sets were named according to the amplifier staining round as primary amplifiers and secondary amplifiers. Each of the primary amplifiers had a complimentary sequence that binds to one of the readout sequences of encoding probes and N 20-mer repeating sequences. Each of the secondary amplifier had a complimentary sequence to one of the 20-mer repeating sequences on primary amplifiers and N 20-mer repeating sequences that were complimentary to the final readout. After primary and secondary amplifiers staining, the original readout signal was one-to-one amplified N^2-fold to the final readout signal. This amplification was termed as NxN amplification (FIGs. 1A-1E).

There were several notable advances in these amplifications. First, the amplifiers were designed using only three nucleotides. Probe sequences that contain only three of the four nucleotides had faster hybridization rates than sequences that use all four nucleotides. This is believed to be due to significantly less secondary structure in such sequences. In addition, to preserve the encoding probes binding to RNAs during amplifier staining and washing step, 20-mer binding sequences were used for amplifiers that stained and washed with 10% formamide at 37 °C. Encoding probes containing 30-mer targeting regions to RNAs were washed with 30% formamide at 47 °C. The amplification was performed on gel embedded and cleared MERFISH samples that inherited the ability to remove fluorescence background. See, e.g., Int. Pat. Apl. Pub. No. WO 2018/089445, incorporated herein by reference in its entirety.

FIGs. 1A-1E show a schematic illustration of amplification for MERFISH imaging. FIG. 1A shows specimens were stained with MERFISH encoding probe library and cleared in polyacrylamide gel. Each encoding probe has a 30-mer target sequence and multiple readout sequences. FIG. 1B shows that primary amplifiers were hybridized to the original readouts on encoding probes following with a washing step to remove extra primary amplifiers for the boxed region in FIG. 1A. The primary amplifiers had a complimentary sequence to the readout sequence on encoding probes and N 20-mer repeating sequences. FIG. 1C shows that secondary amplifiers bind to the primary amplifiers repeating regions. The secondary amplifiers had a complimentary sequence to one of the 20-mer repeating sequences on primary amplifiers and N 20-mer repeating sequences that are complimentary to the final readout. After washing, the specimens were used in multi-round MERFISH imaging, with final readouts using disulfide-linked

fluorophores. FIG. 1D shows that the first round of MERFISH final readout staining with NxN amplified specimens. In comparison, FIG. 1E shows the first round of MERFISH readout staining with unamplified specimens.

EXAMPLE 2

Properties of DNA amplification. In this example, to implement the design, encoding probes containing readout sequences were hybridized to filamin A (FLNA) mRNA and 5x5 amplification was performed on the encoding probes in U-2 OS cells with 15 minutes per amplifier staining round. The average FLNA smFISH spots brightness of either direct labeling readout on encoding probes or labeling 5x5 amplified with the readout on the encoding probe were measured. A 10.5 fold signal increase was observed, which achieved 42% of the theoretical value. See FIGs. 2A and 2B.

It was reasoned that one possibility that the theoretical amplification factor was not reached was that the amplifier staining time was not sufficient. A time series of the same 5x5 amplification was conducted, with each round of amplifier staining with 15 minutes, 30 minutes, 60 minutes, and 180 minutes (FIG. 2C). The results showed that the amplified signal was saturated with hybridized amplifiers within 15 minutes in each round. Saturation at 15 minutes indicates that full assembly of the amplifiers was very rapid, illustrating an additional advantage of the use of the three-letter alphabet for the design of the amplifier probes in this example. In comparison, other methods can often take 45 minutes or more per round of amplification.

Next, whether the length of the amplifiers affected amplification performance was studied. 4x4 amplifiers and 9x9 amplifiers were studied. 5.5x and 30.6x brightness increases with 4x4 amplification and 9x9 amplification were observed, respectively. The results suggested all three amplifier lengths provided ~40% of the theoretical amplification values (FIG. 2D).

One potential consequence of amplification is increasing FISH spot sizes, which may diminish the ability to identify these RNAs, due to the signal from one molecule overlapping with that from other molecules. It is believed that the 9x9 amplifier scaffold has a length of about 132 nm (200 nucleotides x 2 x 0.33 nm per nucleotide), which is within the diffraction limit. The spot size measurements of 4x4 amplification, 5x5 amplification and 9x9 amplification were identical, as expected (FIG. 2E). Based on this design, the variation in brightness from spot to spot was not expected to increase substantially. To confirm this, the coefficient of variations of spot brightness was measured with unamplified, 4x4 amplified, 5x5 amplified, and 9x9 amplified samples. It was found

that the variation in brightness from spot to spot did not increase with amplification (FIG. 2F).

FIGs. 2A-2F show that DNA amplification dramatically increased signal brightness without substantial changes of spot size and variation in brightness from spot to spot. FIG. 2A shows images of U-2 OS cells stained with FLNA smFISH probes, an unamplified sample labeled with readout probes (left). The contrast has been increased 10x relative to the left image to better illustrate the fluorescence signal (middle), or after 5x5 amplification and stained with readout probes (right). (Scale bar, 10 micrometers.) FIG. 2B is an expanded view of the boxed region in FIG. 2A. (Scale bar, 2 micrometers.)

FIG. 2C shows the average brightness of individual FLNA mRNA spots with a time series of amplifier staining. The average brightness has been normalized to unamplified control sample that directly stained with readout probes. FIG. 2D shows the average brightness of individual FLNA mRNA spots in 4x4 amplified and 9x9 amplified samples. The average brightness has been normalized to unamplified control sample that directly stained with readout probes. FIG. 2E shows the average FLNA mRNA spot sizes in unamplified, 4x4 amplified, 5x5 amplified, and 9x9 amplified cells. The width (full width at half maximum) was determined by Gaussian fitting of point spread function (PSF) of RNA spots. FIG. 2F shows the coefficient of variations for RNA spot brightness in unamplified, 4x4 amplified, 5x5 amplified, and 9x9 amplified samples.

EXAMPLE 3

Amplifier screening for MERFISH imaging. In this example, to extend amplification to MERFISH measurements, with a 16-bit, modified Hamming distance-4 (MHD4) encoding scheme, 16 pairs of primary amplifiers and secondary amplifiers were used. Briefly, 20-mer, three-letter repeating sequences were designed by generating a random set of sequences with the per-base probability of 25% for A, 25% for T, and 50% for C or 25% for A, 25% for T, and 50% for G for primary amplifiers and secondary amplifiers, respectively. The sequences were crosschecked within the sequences themselves and blasted to the human transcriptome to avoid homology regions longer than 11 nt. 5x5 amplifiers provide an order of magnitude change in spot brightness. Thus, 5x5 amplification was used in subsequent MERFISH measurements. It was also found that about 20% of the amplifiers had either strong secondary amplifier induced, RNA-dependent, or background or low amplification efficiency. These could be corrected by changing the amplifier pairs). 16 pairs were screened from 20 pairs of 5x5 amplifiers for MERFISH imaging (FIGs. 5A-5C).

EXAMPLE 4

MERFISH measurements with DNA amplification. To validate if the 16 pair amplifiers described in Example 3 would work with MERIFSH, 5x5 amplification was performing using a 130 RNA MERFISH library in U-2 OS cells. 8 rounds of two-color imaging were used to read out 16 bits. In addition, reductive cleavage of disulfide bonds was used to remove the fluorophores linked to the readout probes between consecutive rounds of imaging with both amplified and unamplified samples. FIG. 3A shows that individual RNA molecules could be detected in each of the eight hybridization and imaging rounds of 5x5 amplified samples, allowing their identity to be decoded.

To determine the RNA decoding quality of 5x5 amplified MERFISH measurement, the nonzero misidentification error caused by misidentifying some RNAs as the wrong species and the non-100% calling rate caused by missing some RNAs during detection and decoding were considered from several aspects. First, the RNA counts per cell of the blank barcodes, which correlated with the misidentification rate, and 130 real RNA barcodes were examined. It was found that 121 of the 130 RNA species in the 5x5 amplified MERFISH measurement had a higher copy number per cell than the maximum copy number per cell observed with the blank barcodes (FIG. 3B).

Next, the average 1-to-0 error and 0-to-1 error rate for each bit was determined. If these error rates were low, a potentially obtain high calling rate in theory would be observed. A 1-to-0 error rate of ~1.7% and a 0-to-1 error rate of ~0.6% were observed (FIG. 3C). The 5x5 amplified data was further compared with previously unamplified MERFISH data that had a near 100% detection efficiency. FIG. 3D shows that the copy number per cell observed in the amplified sample correlated strongly with those measured in an unamplified sample with a Pearson correlation coefficient of 0.98 (ρ_{10} (rho) = 0.98 for the 121 RNA species whose measured copy numbers were larger than that observed for the largest blank barcode count). The ratio between the copy numbers measured in the amplified sample to those measured in the unamplified sample was 0.98 +/- 0.12 (SEM, n = 121 RNAs); thus, the 5x5 amplification was kept near 100% detection efficiency.

The average copy number per cell detected for these RNAs by MERFISH was compared with 5x5 amplification with the RNA abundance measured by RNA-seq (FIG. 3E). The Pearson correlation coefficient was 0.90. Also, the copy number per cell results were highly reproducible between replicates of MERFISH experiments with amplification (FIG. 3F).

FIGs. 3A-3F show MERFISH measurement of 130 RNA with 5x5 amplification in U-2 OS cells. FIG. 3A, left, all identified RNAs detected in a 5x5 amplified MERFISH measurement of 130 RNA species with the barcodes of the RNAs represented by the colors (represented here by shading) of the markers. (Scale bar, 10 micrometers.) Right, two-color smFISH images from each of the eight rounds of hybridization and imaging using readout probes labeled with Cy5 (green) or Alexa750 (red) for the boxed region of left picture. (Scale bars, 2 micrometers.) FIG. 3B shows the average RNA copy numbers per cell for real RNA barcodes (dark) and the blank control barcodes (light) detected with 5x5 amplification in about 1200 cells, sorted from largest to smallest value. FIG. 3C shows the error rates, the fraction of measured barcodes that contain a given bit flip, for each bit with 5x5 amplified MERFISH measurements. 1-to-0 error rates (left) and 0-to-1 error rates (right) are shown for each bit. FIG. 3D shows the average copy numbers per cell observed for these RNA species in 5x5 amplified U-2 OS cells versus the copy numbers obtained from an unamplified sample. The Pearson correlation coefficient was 0.98. FIG. 3E shows the average RNA copy numbers per cell determined by MERFISH with 5x5 amplification versus the abundances determined by RNAseq. The Pearson correlation coefficient for the log10 values of RNA abundances was 0.9. FIG. 3F shows the average copy numbers per cell detected in one 5x5 amplified sample versus a replicate sample. The Pearson correlation coefficient was 0.96.

EXAMPLE 5

These examples present a combination of MERFISH and three-letter amplification approach to measure hundreds RNAs in single cells. It was shown that the three-letter amplifiers bound to targets and reached a saturation state within 15 minutes per amplifier staining round. This approach was used to amplify a RNA smFISH signal to 5.5x 10.5x and 30.6x without spot size changes by using 4x4 amplification, 5x5 amplification and 9x9 amplification, respectively. The coefficient of variations of spot brightness in unamplified, 4x4 amplified, 5x5 amplified and 9x9 amplified samples were very close, indicating that the variation of the spot brightness did not increase during DNA amplification, which was important for MERFISH decoding. Finally, accurate identification and counting of 130 RNA species was demonstrated with a near 100% detection efficiency in amplified samples according to the ratio between the copy numbers measured in the amplified sample to those measured in the unamplified sample was 0.98.

It was observed that the amplification factors were around 40% of theoretical number for all three amplifications, including 4x4, 5x5, and 9x9 amplification. It is

believed that the amplifier and readout binding efficiencies were not 100%. The average binding rate per round was around 73.5%. In addition, 80% of the screened amplifier sets worked well, producing substantial amplification. The remaining 20% had either high background or low amplification efficiency. This probability of a amplifier set working without further refinement illustrates the ability to rapidly scale this approach to the amplification of a large number of readout sequences, which, in turn, will be useful for extending the multiplexing capacity.

Amplification was applied on an existing MERFISH library. The substantial increase in signal brightness generated by amplification should extend several additional applications of MERFISH. MERFISH measurements were performed using 92 encoding probes per RNA here. The increase in signal from this amplification approach should provide the possibility to detect RNAs that are much shorter, potentially RNAs that are as short as 300-mer length with as few as 10 encoding probes. The ability to detect RNA molecules with relatively few probes would also dramatically improve RNA isoforms discrimination. Second, MERFISH measurements used dynamic time, which is proportional to the imaged field of view (FOV) number, and a fixed readout staining, readout cleaving and washing time. Imaging time would dominate MERFISH throughput when the FOVs increased to a large number. It should be possible to substantially reduce the imaging duration with shorter exposure time by amplified signals, thus improving MERFISH throughput. In addition, spinning-disk confocal imaging gave higher contrast but less sensitive in detecting photons in thicker samples than wide-field microscopy. Thus, integration of amplification with MERFISH would facilitate the detection of FISH signals in tissues, e.g., with spinning-disk confocal microscopy.

EXAMPLE 6

Following are various materials and methods used in the above examples.

Design of the encoding probes. MERFISH measurements in human osteosarcoma cells (U-2 OS) (ATCC) were performed as follows. Briefly, a 16-bit MHD4 code were used to encode the RNAs. In this encoding scheme, each of the 140 possible barcodes has a constant Hamming weight (i.e., the number of "1" bits in each barcode) of 4 to avoid potential bias in the measurement of different barcodes due to a differential rate of "1" to "0" and "0" to "1" errors. In addition, all barcodes have a Hamming distance of at least 4 to allow for error detection and error correction. 130 of the 140 possible barcodes were used to encode cellular RNAs, and 10 barcodes were used to serve as blank controls. In the encoding

probe library, each RNA species contained 92 encoding probes, with each encoding probe containing three of the four readout sequences assigned to each RNA.

Construction of the encoding probes. The encoding probe library was amplified from complex oligonucleotide pools. Briefly, the oligopools (CustomArray) were amplified via limited-cycle PCR to make in vitro transcription templates first, converted and amplified into RNA through in vitro transcription (New England Biolabs), then converted the RNA back to DNA via reverse transcription (Maxima RT H-, Thermo Fisher). See, e.g., U.S. Pat. Apl. Pub. No. 2017/0212986, incorporated herein by reference. Excess NTPs or dNTPs were removed via desalting columns (Thermo Fisher) after both the in vitro transcription and the phenol-chloroform extraction to improve the purity and reaction yield. The final DNA probes were purified via alkaline hydrolysis to remove RNA templates, phenol-chloroform extraction to remove proteins, and ethanol precipitation to concentrate probes. The final probes were resuspended in RNAase-free water and stored at -20 °C.

Silanization of coverslips. To stabilize the polyacrylamide (PA) gel, coverslips were silanized. Briefly, 40-mm-diameter No. 1.5 coverslips (Bioptechs, 0420-0323-2) were washed for 30 min via immersion in a 1:1 mixture of 37% (vol/vol) methanol and hydrochloric acid at room temperature in a chemical hood. The coverslips were then washed three times in deionized water and once in 70% (vol/vol) ethanol. The coverslips were dried with nitrogen gas airflow and then immersed in chloroform with 0.1% (vol/vol) triethylamine (Sigma) and 0.2% (vol/vol) allyltrichlorosilane (Sigma) for 30 min at room temperature. The coverslips were washed once each with pure chloroform and pure ethanol and then blow dry with nitrogen gas. Silanized coverslips were then stored at room temperature in a desiccated chamber for weeks with no obvious reduction in the quality of the silane layer.

Cell culture and fixation. U-2 OS cells were cultured with Eagle's Minimum Essential Medium (ATCC) containing 10% (vol/vol) fetal bovine serum (FBS) (Thermo Fisher Scientific) and penicillin streptomycin (Thermo Fisher Scientific). To help cell adhesion, silanized coverslips were coated with 0.1 mg/mL poly-D-lysine (PDL) (Sigma) diluted in nuclease-free water for 1 h at room temperature. The coverslips were washed three times with water, incubated in water at room temperature overnight, and then dried and UV-sterilized before plating cells. U-2 OS cells were then at 250,000 cells per coverslip at 37 °C with 5% CO_2 for 48 h. The cells were fixed for 15 min with 4% (vol/vol) paraformaldehyde (PFA) (Electron Microscopy Sciences) in 1x PBS at room temperature, washed three times with 1x PBS, then permeabilized for 10 min with 0.5% (vol/vol) Triton X-100 (Sigma) in 1x PBS, and washed with 1x PBS three times again.

Encoding probe staining was performed as follows. Fixed and permeabilized U-2 OS cells were incubated for 5 min in encoding wash buffer comprising 2x saline-sodium citrate (SSC) (Ambion) and 30% (vol/vol) formamide (Ambion) in nuclease-free water. Then 30 microliters of ~300 micromolar MERFISH encoding probes or ~1 micromolar FLNA encoding probes and 3.3 micromolar of poly(dT) LNA anchor probe (having a sequence /5Acryd/ TTGAGTGGATGGAGTGTAATT+TT+TT+TT+TT+TT+TT+TT+TT+TT+T (SEQ ID NO: 65) , a 20-nt reverse complement of a readout sequence and a 20-nt sequence of alternating dT and thymidine-locked nucleic acid (dT+) with a 5'-acrydite modification (Integrated DNA Technologies) in encoding hybridization buffer was added to the surface of Parafilm and was covered with a cell-containing coverslip. The encoding hybridization buffer contained encoding wash buffer supplemented with 10% (wt/vol) dextran sulfate (Sigma), 0.1% (wt/vol) yeast tRNA (Life Technologies), and 1% (vol/vol) murine RNase inhibitor (New England Biolabs). Samples were incubated in a humid chamber inside a hybridization oven at 37 °C for 36 h. Cells then were washed twice with encoding wash buffer and incubated at 47 °C for 30 min each.

Sample embedding and clearing. 4% PA gels were used to anchor RNAs in place in the encoding probe-stained samples. Briefly, the encoding probe-stained samples on coverslips were first incubated for 2 min with a de-gassed PA solution, with 4% (vol/vol) of 19:1 acrylamide/bis-acrylamide (BioRad), 60 mM Tris·HCl pH 8 (Thermo Fisher), 0.3 M NaCl (Thermo Fisher), and a 1:100,000 dilution of 0.1-micrometer-diameter carboxylate-modified orange fluorescent beads (Life Technologies, F-8800). The beads were used as fiducial markers for the alignment of images taken across multiple rounds of MERFISH imaging. The coverslips containing cells were then incubated again for 2 min with the same PA gel solution containing polymerizing initiator ammonium persulfate (Sigma) and accelerator TEMED (Sigma) at final concentrations of 0.05% (wt/vol) and 0.05% (vol/vol), respectively. 50 microliters of this gel solution was added to the surface of a glass plate (TED Pella) that had been pretreated with 1 mL GelSlick (Lonza) so as not to stick to the PA. The sample was aspirated to remove excess PA gel solution, then inverted onto the 50-microliter droplet to form a thin layer of PA between the coverslip and the glass plate without any bubbles. The coverslip and the glass plate was then gently separated after the gel cast for 1.5 h at room temperature, and the coverslip was washed twice with a digestion buffer of 2% (wt/vol) sodium dodecyl sulfate (SDS) (Thermo Fisher Scientific), 50 mM Tris·HCl pH 8 (Ambion), 1 mM EDTA (Ambion), and 0.5% (vol/vol) Triton X-100 in nuclease-free water.

Then, the gel was immersed and digested with digestion buffer supplemented with 1% (vol/vol) proteinase K (New England Biolabs) in a humidified, 37 °C incubator for >12 h. The digested samples were washed three times with 2x SSC for 15 min each on a rocker. The samples were either stored in 2x SSC supplemented with 0.1% (vol/vol) murine RNase inhibitor at 4 °C for no longer than 48 h for the unamplified MERFISH measurements, or proceeded to amplifier staining.

Amplifier staining. Primary and secondary amplifiers were stained in gel embedded and cleared samples. Samples were incubated for 5 min in a 10% formamide wash buffer, containing 2x SSC (ThermoFisher) and 10% (vol/vol) formamide (ThermoFisher) in nuclease-free water. Next, 50 microliters of 5 nM primary amplifiers in amplifier hybridization buffer, containing 2x SSC, 10% (vol/vol) formamide, 0.1% (wt/vol) yeast tRNA (Life Technologies), 1% (vol/vol) murine RNase inhibitor (New England Biolabs), and 10% (wt/vol) dextran sulfate (Sigma), was added to the surface of Parafilm. The samples covered on the 50 microliter droplets on Parafilm after removing extra 10% formamide wash buffer with Kimwipes and incubated in a humidity-controlled 37 °C incubator for 30 minutes (MERFISH) or a time series of 15 minutes, 30 minutes, 60 minutes, and 180 minutes for binding rate measurements. Then the samples were washed three times in 10% formamide wash buffer at room temperature in Petri dishes for 5 minutes each. On new Parafilm, 50 microliters of 5 nM secondary amplifiers in the amplifier hybridization buffer was added. The washed samples were again covered on the droplets and incubated in 37 °C incubator with the same time as primary amplifier staining. Then the samples were washed twice in 10% formamide wash buffer at room temperature for 5 minutes each and a third wash in 10% formamide wash buffer in 37 °C incubator for 15 minutes. The samples were either imaged immediately or stored in 2x SSC supplemented with 0.1% (vol/vol) murine RNase inhibitor at 4 °C for no longer than 48 h.

MERFISH imaging platforms. The samples were imaged on a home-built high-throughput imaging platform at Center for Advanced Imaging, Harvard University. Briefly, this microscope was constructed around an Olympus IX-71 microscope body and a Nikon, CFI Plan Apo Lambda 60x oil objective. Illumination in 750, 647, 560, 488 and 405 nm was provided using solid-state lasers (MBP Communications, 2RU-VFL-P-500-750-B1R; MBP Communications, 2RU-VFL-P-2000-647-B1R; MBP Communications, 2RU-VFL-P-2000-560-B1R; MBP Communications, 2RU-VFL-P-500-488-B1R; Coherent, Cube 405). These laser lines were used to excite readout probes labeled with Alexa750 and Cy5, orange fiducial beads, Poly dT readout and DAPI, respectively. The illumination profile was flattened with a

πShaper (Pishaper). The fluorescence emission from the sample was separated from the laser illumination using a penta-band dichroic (Chroma, zy405/488/561/647/752RP-UF1) and imaged with a scientific CMOS camera (sCMOS; Hamamatsu, C11440-22CU) after passing through two duplicate custom penta-notch filters (Chroma, ZET405/488/561/647-656/752m) to remove stray excitation light. The pixel size for the sCMOS camera was determined to correspond to 109 nm in the sample plane. The exposure time was 500 ms for each imaging frame. The sample position was controlled via a motorized microscope stage (Ludl) and focus was maintained via a custom focus-lock system realized through a feedback system between an objective nanopositioner (Mad City Labs, Nano-F100S) and the reflection of an IR laser (Thorlabs, LP980-SF15) onto a CMOS camera (Thorlabs, DCC1545M). The sample coverslip was held inside a flow chamber (Bioptechs, FCS2), and buffer exchange within this chamber was directed using a custom-built automated fluidics system, controlling three eight-way valves (Hamilton, MVP and HVXM 8-5) and a peristaltic pump (Gilison, Minipuls 3).

Sample imaging. Sequential MERFISH imaging was carried out on the high-throughput imaging platform described above. Briefly, each MERFISH round used readout probe staining (10 min), wash buffer washing (5 min), imaging buffer flowing (3 min), 100-400 FOVs imaging, readout fluorophore cleaving (15 min), and 2xSSC washing (5min). 8 rounds of two color MERFISH imaging were performed on each sample. Specifically, the readout probes staining was done by flowing 3 nM readout probes in hybridization buffer of 2x SSC, 10% (vol/vol) ethylene carbonate (Sigma-Aldrich), 0.1% (vol/vol) murine RNase inhibitor (NEB) in nuclease-free water. Wash buffer contained 2x SSC and 10% (vol/vol) ethylene carbonate in nuclease-free water. The imaging buffer used 2x SSC, 50 mM Tris·HCl pH 8, 10% (wt/vol) glucose, 2 mM Trolox (Sigma-Aldrich), 0.5 mg/mL glucose oxidase (Sigma-Aldrich), 40 micrograms/mL catalase (Sigma-Aldrich), and 0.1% (vol/vol) murine RNase inhibitor in nuclease-free water. Cleavage buffer comprising 2x SSC and 50 mM of Tris (2-carboxyethyl) phosphine (TCEP; Sigma) was used to cleave the disulfide bond conjugating dyes to the probes. SmFISH imaging was similar to single round MERFISH imaging but without cleaving and 2xSSC washing steps.

Image processing and decoding. For FLNA smFISH data, the brightness and PSF sizes of unamplied and amplified FISH spots were calculated via a Gaussian fitting routine. For the MERFISH data, registration of images of the same FOV across different imaging rounds as well as decoding of the RNA barcodes was conducted as follows. Briefly, the offsets between images in each imaging round were corrected using the localizations of the fiducial beads in each round of imaging. Next, background in the images were removed with

a high-pass filter, RNA spots were tightened by deconvolution, and RNA centroids that differed slightly in location between images were connected via low-pass filter. Then the brightness of RNA spot different color channels were normalized by equalizing their intensity histograms via an iterative process. By comparing the normalized intensity of each pixel across all the 16 images in 8 two-color rounds, the barcodes were assigned to individual pixels.

Computations were run on a desktop server that contained two 10-core Intel Xeon E5-2680 2.8-GHz CPUs and 256 GB of RAM.

EXAMPLE 7

To demonstrate the 5x5 amplifiers in tissue samples, in this example, the mouse medial preoptic area (MPOA) was chosen. This tissue was cut in 10 micrometer slices, and stained with a 135 gene MERFISH library following with gel embedding and clearing. Then 5x5 amplification was performed on the tissue samples, using techniques similar to those described above.

FIG. 4A shows the first round raw image of the 5x5 amplified bright spots. FIG. 4B is an expanded view of the white boxed region in FIG. 4A. (The scale bar in FIG. 4A is 10 micrometers and the scale bar in FIG. 4B is 2 micrometers)

After decoding, the RNA counts from the MERFISH measurements were compared with RNA abundant data from RNA-seq. They were correlated strongly with a Pearson correlation coefficient of 0.81. This can be seen in FIG. 4C, which shows a plot of FPKM (fragments per kilobase of transcript per million mapped reads) and the counts per field of view (FOV).

While several embodiments of the present invention have been described and illustrated herein, those of ordinary skill in the art will readily envision a variety of other means and/or structures for performing the functions and/or obtaining the results and/or one or more of the advantages described herein, and each of such variations and/or modifications is deemed to be within the scope of the present invention. More generally, those skilled in the art will readily appreciate that all parameters, dimensions, materials, and configurations described herein are meant to be exemplary and that the actual parameters, dimensions, materials, and/or configurations will depend upon the specific application or applications for which the teachings of the present invention is/are used. Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. It is, therefore, to

be understood that the foregoing embodiments are presented by way of example only and that, within the scope of the appended claims and equivalents thereto, the invention may be practiced otherwise than as specifically described and claimed. The present invention is directed to each individual feature, system, article, material, kit, and/or method described herein. In addition, any combination of two or more such features, systems, articles, materials, kits, and/or methods, if such features, systems, articles, materials, kits, and/or methods are not mutually inconsistent, is included within the scope of the present invention.

In cases where the present specification and a document incorporated by reference include conflicting and/or inconsistent disclosure, the present specification shall control. If two or more documents incorporated by reference include conflicting and/or inconsistent disclosure with respect to each other, then the document having the later effective date shall control.

All definitions, as defined and used herein, should be understood to control over dictionary definitions, definitions in documents incorporated by reference, and/or ordinary meanings of the defined terms.

The indefinite articles "a" and "an," as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to mean "at least one."

The phrase "and/or," as used herein in the specification and in the claims, should be understood to mean "either or both" of the elements so conjoined, i.e., elements that are conjunctively present in some cases and disjunctively present in other cases. Multiple elements listed with "and/or" should be construed in the same fashion, i.e., "one or more" of the elements so conjoined. Other elements may optionally be present other than the elements specifically identified by the "and/or" clause, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, a reference to "A and/or B", when used in conjunction with open-ended language such as "comprising" can refer, in one embodiment, to A only (optionally including elements other than B); in another embodiment, to B only (optionally including elements other than A); in yet another embodiment, to both A and B (optionally including other elements); etc.

As used herein in the specification and in the claims, "or" should be understood to have the same meaning as "and/or" as defined above. For example, when separating items in a list, "or" or "and/or" shall be interpreted as being inclusive, i.e., the inclusion of at least one, but also including more than one, of a number or list of elements, and, optionally, additional unlisted items. Only terms clearly indicated to the contrary, such as "only one of" or "exactly one of," or, when used in the claims, "consisting of," will refer to the inclusion of

exactly one element of a number or list of elements. In general, the term "or" as used herein shall only be interpreted as indicating exclusive alternatives (i.e. "one or the other but not both") when preceded by terms of exclusivity, such as "either," "one of," "only one of," or "exactly one of."

As used herein in the specification and in the claims, the phrase "at least one," in reference to a list of one or more elements, should be understood to mean at least one element selected from any one or more of the elements in the list of elements, but not necessarily including at least one of each and every element specifically listed within the list of elements and not excluding any combinations of elements in the list of elements. This definition also allows that elements may optionally be present other than the elements specifically identified within the list of elements to which the phrase "at least one" refers, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, "at least one of A and B" (or, equivalently, "at least one of A or B," or, equivalently "at least one of A and/or B") can refer, in one embodiment, to at least one, optionally including more than one, A, with no B present (and optionally including elements other than B); in another embodiment, to at least one, optionally including more than one, B, with no A present (and optionally including elements other than A); in yet another embodiment, to at least one, optionally including more than one, A, and at least one, optionally including more than one, B (and optionally including other elements); etc.

When the word "about" is used herein in reference to a number, it should be understood that still another embodiment of the invention includes that number not modified by the presence of the word "about."

It should also be understood that, unless clearly indicated to the contrary, in any methods claimed herein that include more than one step or act, the order of the steps or acts of the method is not necessarily limited to the order in which the steps or acts of the method are recited.

In the claims, as well as in the specification above, all transitional phrases such as "comprising," "including," "carrying," "having," "containing," "involving," "holding," "composed of," and the like are to be understood to be open-ended, i.e., to mean including but not limited to. Only the transitional phrases "consisting of" and "consisting essentially of" shall be closed or semi-closed transitional phrases, respectively, as set forth in the United States Patent Office Manual of Patent Examining Procedures, Section 2111.03.

What is claimed is:

CLAIMS

1. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein a maximum number of primary amplifier nucleic acids is able to bind to a nucleic acid probe;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, wherein a maximum number of secondary amplifier nucleic acids is able to bind to the primary amplifier nucleic acids;

determining a distribution of the nucleic acid probes within the sample using fluorescence;

creating codewords based on the fluorescence distribution within the sample; and

for at least some of the codewords, matching the codeword to a valid codeword optionally wherein, if no match is found, applying error correction to the codeword to form a valid codeword.

2. The method of claim 1, comprising exposing the sample to at least 5 different nucleic acid probes.

3. The method of any one of claims 1 or 2, comprising exposing the sample to at least 10 different nucleic acid probes.

4. The method of any one of claims 1-3, comprising exposing the sample to at least 100 different nucleic acid probes.

5. The method of any one of claims 1-4, comprising sequentially exposing the sample to a plurality of nucleic acid probes.

6. The method of any one of claims 1-5, wherein the plurality of nucleic acid probes comprises a combinatorial combination of nucleic acid probes with different sequences.

7. The method of claim 6, wherein the combinatorial combination of nucleic acid probes target a combinatorial combination of RNA species in the sample.

8. The method of any one of claims 6 or 7, wherein the combinatorial combination of nucleic acid probes target a combinatorial combination of DNA sequences in the sample.

9. The method of any one of claims 1-8, wherein at least some of the plurality of nucleic acid probes comprise DNA.

10. The method of any one of claims 1-9, wherein at least some of the plurality of nucleic acid probes comprise RNA.

11. The method of any one of claims 1-10, wherein the plurality of nucleic acid probes have an average length of between 10 and 300 nucleotides.

12. The method of any one of claims 1-11, wherein at least some of the plurality of nucleic acids probes are configured to bind to a nucleic acid within the sample.

13. The method of any one of claims 1-12, wherein at least some of the plurality of nucleic acid probes comprises a target sequence and one or more read sequences.

14. The method of claim 13, wherein the target sequence of the plurality of nucleic acid probes has an average length of between 10 and 200 nucleotides.

15. The method of any one of claims 13 or 14, wherein the target sequence is substantially complementary to a nucleic acid sequence encoding a protein.

16. The method of any one of claims 13-15, wherein the target sequence binds to a target via specific binding.

17. The method of claim 16, wherein the target comprises RNA.

18. The method of claim 17, wherein the RNA comprises non-coding RNA.

19. The method of any one of claims 17 or 18, wherein the RNA comprises mRNA.

20. The method of any one of claims 17-19, wherein the RNA comprises transfer RNA (tRNA).

21. The method of any one of claims 17-20, wherein the RNA comprises ribosomal RNA (rRNA).

22. The method of any one of claims 16-21, wherein the target comprises DNA.

23. The method of claim 22, wherein the DNA comprises genomic DNA.

24. The method of any one of claims 13-23, wherein the primary amplifier nucleic acids bind to the nucleic acid probes via the read sequences.

25. The method of any one of claims 13-24, wherein the plurality of nucleic acid probes comprises distinguishable nucleic acid probes formed from combinatorial combination of a pool of one or more read sequences.

26. The method of claim 25, wherein the pool has at least 8 possible read sequences.

27. The method of any one of claims 25 or 26, wherein the pool has at least 16 possible read sequences.

28. The method of any one of claims 25-27, wherein the pool has at least 24 possible read sequences.

29. The method of any one of claims 25-28, wherein the pool has at least 32 possible read sequences.

30. The method of any one of claims 25-29, wherein the pool has at least 48 possible read sequences.

31. The method of any one of claims 25-30, wherein the pool has at least 64 possible read sequences.

32. The method of any one of claims 25-31, wherein the pool has no more than 32 possible read sequences.

33. The method of any one of claims 25-32, wherein the pool has no more than 16 possible read sequences.

34. The method of any one of claims 25-33, wherein the plurality of read sequences are distributed on the plurality of nucleic acid probes so as to define an error-correcting code.

35. The method of any one of claims 13-34, wherein the read sequences have an average length of between 5 nucleotides and 50 nucleotides.

36. The method of any one of claims 13-35, wherein at least some of the plurality of nucleic acid probes comprise no more than 10 read sequences.

37. The method of any one of claims 13-36, wherein at least some of the plurality of nucleic acid probes comprise no more than 5 read sequences.

38. The method of any one of claims 1-37, wherein no more than 20 primary amplifier nucleic acids is able to bind to the nucleic acid probe.

39. The method of any one of claims 1-38, wherein no more than 10 primary amplifier nucleic acids is able to bind to the nucleic acid probe.

40. The method of any one of claims 1-39, wherein no more than 5 primary amplifier nucleic acids is able to bind to the nucleic acid probe.

41. The method of any one of claims 1-40, wherein the primary amplifier nucleic acids have an average length of less than 300 nucleotides.

42. The method of any one of claims 1-41, wherein the primary amplifier nucleic acids have an average length of less than 200 nucleotides.

43. The method of any one of claims 1-42, wherein the primary amplifier nucleic acids have a repeating sequence that the secondary amplifier nucleic acids are able to bind.

44. The method of any one of claims 1-43, wherein the secondary amplifier nucleic acids each target a sequence of less than 10 nucleotides on the primary amplifier nucleic acids.

45. The method of any one of claims 1-44, wherein the primary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides.

46. The method of any one of claims 1-45, wherein no more than 20 secondary amplifier nucleic acids is able to bind to a primary amplifier nucleic acid.

47. The method of any one of claims 1-46, wherein no more than 10 secondary amplifier nucleic acids is able to bind to a primary amplifier nucleic acid.

48. The method of any one of claims 1-47, wherein no more than 5 secondary amplifier nucleic acids is able to bind to a primary amplifier nucleic acid.

49. The method of any one of claims 1-48, wherein the secondary amplifier nucleic acids have an average length of less than 300 nucleotides.

50. The method of any one of claims 1-49, wherein the secondary amplifier nucleic acids have an average length of less than 200 nucleotides.

51. The method of any one of claims 1-50, wherein the secondary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides.

52. The method of any one of claims 1-51, wherein the secondary amplifier nucleic acids comprise a fluorescent signaling entity.

53. The method of any one of claims 1-52, further comprising exposing the secondary amplifier nucleic acids to a fluorescent signaling entity.

54. The method of any one of claims 52 or 53, wherein the secondary amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

55. The method of any one of claims 52-54, wherein the fluorescent signaling entity is linked to the secondary amplifier nucleic acids by a disulfide.

56. The method of any one of claims 1-51, further comprising exposing the secondary amplifier nucleic acids to tertiary amplifier nucleic acids able to bind to the secondary amplifier nucleic acids.

57. The method of claim 56, wherein the tertiary amplifier nucleic acids comprise a fluorescent signaling entity.

58. The method of claim 56, further comprising exposing the tertiary amplifier nucleic acids to a fluorescent signaling entity.

59. The method of any one of claims 57 or 58, wherein the tertiary amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

60. The method of claim 56, further comprising exposing the tertiary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

61. The method of claim 60, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

62. The method of claim 60, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

63. The method of any one of claims 61 or 62, wherein the ending round of amplifier nuclei acids has a repeating sequence that the fluorescent signaling entity is able to bind.

64. The method of any one of claims 52-55, 57-59, or 61-63, wherein the fluorescent signaling entity comprises a protein.

65. The method of any one of claims 52-55, 57-59, or 61-64, wherein the fluorescent signaling entity comprises a dye.

66. The method of any one of claims 52-55, 57-59, or 61-65, wherein the fluorescent signaling entity comprises Alexa Fluor 750.

67. The method of any one of claims 52 52-55, 57-59, or 61-66, wherein the fluorescent signaling entity comprises Cy5.

68. The method of any one of claims 52-55, 57-59, or 61-67, wherein the fluorescent signaling entity comprises a nanoparticle.

69. The method of any one of claims 52-55, 57-59, or 61-68, comprising inactivating the fluorescent signaling entity between exposure of the sample to nucleic acid probes.

70. The method of claim 69, comprising inactivating the fluorescent signaling entity by removing the fluorescent signaling entity.

71. The method of any one of claims 69 or 70, comprising inactivating the fluorescent signaling entity using a reducing agent.

72. The method of any one of claims 69-71, comprising inactivating the fluorescent signaling entity using disulfide cleavage.

73. The method of any one of claims 69-72, comprising inactivating the fluorescent signaling entity by photobleaching the fluorescent signaling entity.

74. The method of any one of claims 69-73, comprising inactivating the fluorescent signaling entity by chemically bleaching the fluorescent signaling entity.

75. The method of any one of claims 69-74, comprising inactivating the fluorescent signaling entity by enzymatically cleaving the fluorescent signaling entity.

76. The method of any one of claims 1-75, wherein the plurality of nucleic acid probes defines a code space with a Hamming distance of at least 2.

77. The method of any one of claims 76, wherein the plurality of nucleic acid probes defines a code space with a Hamming distance of at least 3.

78. The method of any one of claims 76 or 77, wherein the code space is a Hamming(7, 4) code, a Hamming(15, 11) code, a Hamming(31, 26) code, a Hamming(63, 57) code, or a Hamming(127, 120) code.

79. The method of any one of claims 76-78, wherein the code space is a SECDED code.

80. The method of claim 79, wherein the code space is a SECDED(8,4) code, a SECDED(16,4) code, a SECDED(16, 11) code, a SECDED (22, 16) code, a SECDED (39, 32) code, or a SECDED (72, 64) code.

81. The method of any one of claims 76-80, wherein the nucleic acid probes defines a code space with only a constant number of ones.

82. The method of any one of claims 1-81, comprising determining the distribution of nucleic acid probes by imaging at least a portion of the sample.

83. The method of claim 82, comprising determining binding of the nucleic acid probes using an optical imaging technique.

84. The method of any one of claims 82 or 83, comprising determining binding of the nucleic acid probes using a fluorescence imaging technique.

85. The method of any one of claims 82-84, comprising determining binding of the nucleic acid probes using a multi-color fluorescence imaging technique.

86. The method of any one of claims 82-85, comprising determining binding of the nucleic acid probes using a super-resolution fluorescence imaging technique.

87. The method of claim 86, comprising determining binding of the nucleic acid probes using stochastic optical reconstruction microscopy (STORM).

88. The method of any one of claims 82-87, comprising determining the centroid of images of the secondary amplifier nucleic acids using an algorithm for determining non-overlapping single emitters.

89. The method of any one of claims 82-88, comprising determining the centroid of images of the secondary amplifier nucleic acids using an algorithm for determining partially overlapping single emitters.

90. The method of any one of claims 88 or 89, further comprising determining a confidence level for the secondary amplifier nucleic acids.

91. The method of claim 90, comprising determining the confidence level using a ratio of the number of exact matches to the number of matches having one or more one-bit errors to the codeword.

92. The method of any one of claims 90 or 91, comprising determining the confidence level using a ratio of the number of exact matches to the number of matches having exactly one one-bit error to the codeword.

93. The method of claims 82-92, comprising determining the distribution of nucleic acid probes within the sample at a resolution better than 300 nm.

94. The method of claims 82-93, comprising determining the distribution of nucleic acid probes within the sample at a resolution better than 100 nm.

95. The method of claims 82-94, comprising determining the distribution of nucleic acid probes within the sample at a resolution better than 50 nm.

96. The method of any one of claims 1-95, wherein the sample comprises a cell.

97. The method of claim 96, wherein the cell is a human cell.

98. The method of any one of claims 96-97, wherein the cell is fixed.

99. The method of any one of claims 1-98, wherein the sample comprises tissue.

100. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, wherein the binding of primary amplifier nucleic acids and secondary amplifier nucleic acids to a target is saturatable;

determining a distribution of the nucleic acid probes within the sample using fluorescence;

creating codewords based on the fluorescence distribution within the sample; and

for at least some of the codewords, matching the codeword to a valid codeword optionally wherein, if no match is found, applying error correction to the codeword to form a valid codeword.

101. The method of claim 100, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

102. The method of claim 101, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

103. The method of claim 101, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

104. The method of any one of claims 102 or 103, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

105. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, wherein the secondary amplifier nucleic acids bind to the primary amplifier nucleic acids within a fixed distance;

determining a distribution of the nucleic acid probes within the sample using fluorescence;

creating codewords based on the fluorescence distribution within the sample; and

for at least some of the codewords, matching the codeword to a valid codeword optionally wherein, if no match is found, applying error correction to the codeword to form a valid codeword.

106. The method of claim 105, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

107. The method of claim 106, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

108. The method of claim 106, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

109. The method of any one of claims 107 or 108, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

110. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein the primary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids;

determining a distribution of the nucleic acid probes within the sample using fluorescence;

creating codewords based on the fluorescence distribution within the sample; and

for at least some of the codewords, matching the codeword to a valid codeword optionally wherein, if no match is found, applying error correction to the codeword to form a valid codeword.

111. The method of claim 110, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

112. The method of claim 111, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

113. The method of claim 111, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

114. The method of any one of claims 112 or 113, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

115. The method of any one of claims 111-114, wherein at least one of the one or more additional rounds of amplifier nucleic acids is formed from only 3 of the 4 naturally occurring nucleotides.

116. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, wherein the secondary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides;

determining a distribution of the nucleic acid probes within the sample using fluorescence;

creating codewords based on the fluorescence distribution within the sample; and

for at least some of the codewords, matching the codeword to a valid codeword optionally wherein, if no match is found, applying error correction to the codeword to form a valid codeword.

117. The method of claim 116, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

118. The method of claim 117, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

119. The method of claim 117, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

120. The method of any one of claims 118 or 119, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

121. The method of any one of claims 117-120, wherein at least one of the one or more additional rounds of amplifier nucleic acids is formed from only 3 of the 4 naturally occurring nucleotides.

122. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein a maximum number of primary amplifier nucleic acids is able to bind to a nucleic acid probe;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

123. The method of claim 122, further comprising exposing the primary amplifier nucleic acids to a fluorescent signaling entity.

124. The method of claim 122, wherein the primary amplifier nucleic acids comprise a fluorescent signaling entity.

125. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein a maximum number of primary amplifier nucleic acids is able to bind to a nucleic acid probe;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample, and for at least some of the codewords, matching the codeword to a valid codeword wherein, if no match is found, potentially applying error correction to the codeword to form a valid codeword.

126. The method of claim 125, further comprising exposing the primary amplifier nucleic acids to a fluorescent signaling entity.

127. The method of claim 125, wherein the primary amplifier nucleic acids comprise a fluorescent signaling entity.

128. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein the binding of primary amplifier nucleic acids to a target is saturatable;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

129. The method of claim 128, further comprising exposing the primary amplifier nucleic acids to a fluorescent signaling entity.

130. The method of claim 128, wherein the primary amplifier nucleic acids comprise a fluorescent signaling entity.

131. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein the binding of primary amplifier nucleic acids to a target is saturatable;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample, and for at least some of the codewords, matching the codeword to a valid codeword wherein, if no match is found, potentially applying error correction to the codeword to form a valid codeword.

132. The method of claim 131, further comprising exposing the primary amplifier nucleic acids to a fluorescent signaling entity.

133. The method of claim 131, wherein the primary amplifier nucleic acids comprise a fluorescent signaling entity.

134. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein the primary amplifier nucleic acids bind to the nucleic acid probes within a fixed distance;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

135. The method of claim 134, wherein for at least some of the codewords, matching the codeword to a valid codeword.

136. The method of claim 135, wherein, if no match is found, applying error correction to the codeword to form a valid codeword.

137. The method of any one of claims 134-136, further comprising exposing the primary amplifier nucleic acids to a fluorescent signaling entity.

138. The method of any one of claims 134-136, wherein the primary amplifier nucleic acids comprise a fluorescent signaling entity.

139. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein the primary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

140. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein the primary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample and for at least some of the codewords, matching the codeword to a valid codeword wherein, if no match is found, potentially applying error correction to the codeword to form a valid codeword.

141. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein a maximum number of primary amplifier nucleic acids is able to bind to a nucleic acid probe;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, wherein a maximum number of secondary amplifier nucleic acids is able to bind to the primary amplifier nucleic acids;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

142. The method of claim 141, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

143. The method of claim 142, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

144. The method of claim 142, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

145. The method of any one of claims 143 or 144, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

146. The method of any one of claims 141-145, wherein for at least some of the codewords, matching the codeword to a valid codeword.

147. The method of claim 146, wherein, if no match is found, potentially applying error correction to the codeword to form a valid codeword

148. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, wherein the binding of primary amplifier nucleic acids and secondary amplifier nucleic acids to a target is saturatable;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

149. The method of claim 148, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

150. The method of claim 149, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

151. The method of claim 149, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

152. The method of any one of claims 150 or 151, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

153. The method of any one of claims 148-152, wherein for at least some of the codewords, matching the codeword to a valid codeword.

154. The method of claim 153, wherein, if no match is found, potentially applying error correction to the codeword to form a valid codeword

155. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, wherein the secondary amplifier nucleic acids bind to the primary amplifier nucleic acids within a fixed distance;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

156. The method of claim 155, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

157. The method of claim 156, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

158. The method of claim 156, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

159. The method of any one of claims 157 or 158, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

160. The method of any one of claims 155-159, wherein for at least some of the codewords, matching the codeword to a valid codeword.

161. The method of claim 160, wherein, if no match is found, potentially applying error correction to the codeword to form a valid codeword

162. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein the primary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

163. The method of claim 162, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

164. The method of claim 163, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

165. The method of claim 163, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

166. The method of any one of claims 164 or 165, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

167. The method of any one of claims 162 to 166, wherein for at least some of the codewords, matching the codeword to a valid codeword.

168. The method of claim 167, wherein, if no match is found, potentially applying error correction to the codeword to form a valid codeword

169. The method of any one of claims 162-168, wherein at least one of the one or more additional rounds of amplifier nucleic acids is formed from only 3 of the 4 naturally occurring nucleotides.

170. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, wherein the secondary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

171. The method of claim 170, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

172. The method of claim 171, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

173. The method of claim 171, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

174. The method of any one of claims 172 or 173, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

175. The method of any one of claims 170-174, wherein for at least some of the codewords, matching the codeword to a valid codeword.

176. The method of claim 175, wherein, if no match is found, potentially applying error correction to the codeword to form a valid codeword

177. The method of any one of claims 170-176, wherein at least one of the one or more additional rounds of amplifier nucleic acids is formed from only 3 of the 4 naturally occurring nucleotides.

178. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

179. A method, comprising:

exposing a sample to nucleic acid probes;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

180. The method of claim 179, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an

ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

181. The method of claim 180, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

182. The method of claim 180, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

183. The method of any one of claims 181 or 182, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

184. A method, comprising:

exposing a sample to a binding entity conjugated to a nucleic acid probe;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

185. A method, comprising:

exposing a sample to a binding entity conjugated to a nucleic acid probe;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein the binding of primary amplifier nucleic acids to a target is saturatable;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

186. A method, comprising:

exposing a sample to a targeting entity conjugated to a nucleic acid probe;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic

acids able to bind to the primary amplifier nucleic acids;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

187. The method of claim 186, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

188. The method of claim 187, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

189. The method of claim 187, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

190. The method of any one of claims 188 or 189, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

191. A method, comprising:

exposing a sample to a targeting entity conjugated to a nucleic acid probe;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, wherein the binding of primary amplifier nucleic acids and secondary amplifier nucleic acids to a target is saturatable;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

192. The method of claim 191, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an

ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

193. The method of claim 192, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

194. The method of claim 192, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

195. The method of any one of claims 193 or 194, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

196. A method, comprising:

exposing a sample to a targeting entity conjugated to a nucleic acid probe;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes, wherein the primary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

197. The method of claim 196, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

198. The method of claim 197, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

199. The method of claim 197, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

200. The method of any one of claims 198 or 199, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

201. The method of any one of claims 196-200, wherein for at least some of the codewords, matching the codeword to a valid codeword.

202. The method of claim 201, wherein, if no match is found, potentially applying error correction to the codeword to form a valid codeword

203. The method of any one of claims 196-202, wherein at least one of the one or more additional rounds of amplifier nucleic acids is formed from only 3 of the 4 naturally occurring nucleotides.

204. A method, comprising:

exposing a sample to a targeting entity conjugated to a nucleic acid probe;

exposing the nucleic acid probes to primary amplifier nucleic acids able to bind to the nucleic acid probes;

exposing the primary amplifier nucleic acids to secondary amplifier nucleic acids able to bind to the primary amplifier nucleic acids, wherein the secondary amplifier nucleic acids are formed from only 3 of the 4 naturally-occurring nucleotides;

determining a distribution of the nucleic acid probes within the sample using fluorescence; and

creating codewords based on the fluorescence distribution within the sample.

205. The method of claim 204, further comprising exposing the secondary amplifier nucleic acids to one or more additional rounds of amplifier nucleic acids, including an ending round of amplifier nucleic acids, each able to bind the amplifier nucleic acids of the previous round.

206. The method of claim 205, wherein the ending amplifier nucleic acids comprise a fluorescent signaling entity.

207. The method of claim 205, further comprising exposing the ending round of amplifier nucleic acids to a fluorescent signaling entity.

208. The method of any one of claims 206 or 207, wherein the ending amplifier nucleic acids have a repeating sequence that the fluorescent signaling entity is able to bind.

209. The method of any one of claims 204-208, wherein for at least some of the codewords, matching the codeword to a valid codeword.

210. The method of claim 209, wherein, if no match is found, potentially applying error correction to the codeword to form a valid codeword

211. The method of any one of claims 204-210, wherein at least one of the one or more additional rounds of amplifier nucleic acids is formed from only 3 of the 4 naturally occurring nucleotides.