In [1]:
```
from matplotlib import image as mp_image
import matplotlib.pyplot as plt
import os
%matplotlib inline
```

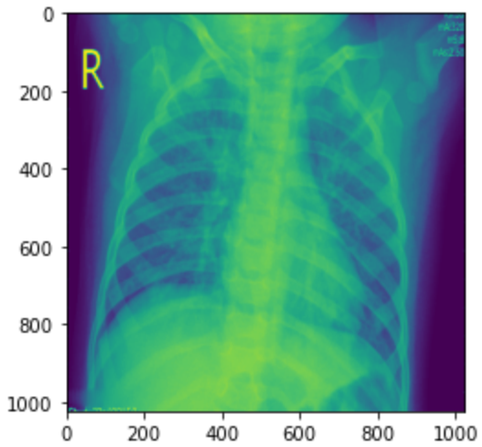
In [2]:
```
#second class
viral_pneumonia_images='XRays/Viral Pneumonia'
```

In [3]:
```
#Create a function to iterate through the images of each folder, check file type
#the size and the dimensions of each image also ensure that the file names are consistent,
#make sure there are no major irregularities or unwanted patterns
#to ensure correct learning from the upcoming model

def show_images(image_folder):
    fig = plt.figure()
    %matplotlib inline
    file_names = os.listdir(image_folder)
    print('this folder contains {} files'.format(len(file_names)))
    img_num = 0
    for file_name in file_names:
        file_path = os.path.join(image_folder, file_name)
        # Open the file using the matplotlib.image library
        image = mp_image.imread(file_path)
        # Add the image to the figure (which will have 1 row, a column for each filename, and a position based on its index in the file_names list)
        a=fig.add_subplot(1, len(file_names), file_names.index(file_name)+1)
        # Add the image to the plot
        image_plot = plt.imshow(image)
        # Add a caption with the file name
        a.set_title(file_name)
        # Show filenames
        print('file name: ',file_name)
        # Show image shape
        print('shape: ',image.shape)
        # Show images
        plt.show()
        # Show the plot
```

In [4]:
```
show_images(viral_pneumonia_images)
```

```
this folder contains 1345 files
file name:  Viral Pneumonia (1).png
shape:  (1024, 1024)
```



In [ ]:
```
# we can these images are not as heavily annotated as
# some of the covid images, you do see markings
# though, aside from
# top left hand markings like R which I saw repeatedly and
# could be different from the covid images in quantity
# and distribution
# I would like to run a test one day with no markings
# but these images are mostly clean, and on both sets there
# are images to be classified that have no markings
# so another aspect is recognizing features as the data set
# grows and grows maybe these things become less important as
# you gather higher numbers of both annotated and clean
# images.
# Also noted on these images is that the majority are much more
# concentrated on ventral or dorsal views (I'm not
# pretending to be a domain expert at all!
# but they are definitely more focused on the chest
# area,something that again might become less of a concern
# as the dataset grows, but considering this is the larger data
# set, it might be a good consideration to put more lateral images,
# just to make sure the image angles are consistent as well from
# data set to data set between the classes.
# Last notable difference in this set is that the images
# are grayscale(one dimension) mixed with three dimensional
# images.
```